MAS242 선형대수학 Notes

한승우

October 24, 2023

CONTENTS

CHAPTER	Linear Equations	PAGE 2
CHAPTER	VECTOR SPACES	
2.1	Bases and Dimension	3
CHAPTER	Linear Transformations	Page 6
3.1	Linear Transformations	6
3.2	The Algebra of Linear Transformations	7
3.3	Isomorphism	9
3.4	Representation of Transformation by Matrices	9
3.5	Linear Functionals	10
3.6	The Double Dual	12
3.7	The Transpose of a Linear Transformation	13
CHAPTER	POLYNOMIALS	Page 14
4.1	Algebras	14
CHAPTER	Determinants	Page 18
5.1	Determinant Functions	18
Снартек	Elementary Canonical Forms	Page 22
	Eigenvalues	22

Chapter 1 Linear Equations

Vector Spaces

2.1 Bases and Dimension

Theorem 2.1.1

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

Lemma 2.1.1

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

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

Theorem 2.1.2 Inclusion/Exclusion Principle for Vector Spaces

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

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

We now claim that

$$B \triangleq \left\{ \alpha_{1}, \cdots, \alpha_{d}, \beta_{d+1}, \cdots, \beta_{a}, \gamma_{d+1}, \cdots, \gamma_{b} \right\}$$

is a basis of $W_1 + W_2$.

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

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

Definition 2.1.1: Ordered Basis

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

Note:-

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

Lemma 2.1.2

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

$$x = x_1 v_2 + x_2 v_2 + \cdots + x_n v_n$$

for some $x_i \in F$.

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

(Uniqueness) Suppose we have two such expressions:

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

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

Definition 2.1.2: Coordinate Matrix

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

$$\begin{bmatrix} x \end{bmatrix}_{B} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

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

Theorem 2.1.3

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

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

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

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

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \cdots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix} \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}$$

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

Linear Transformations

Linear Transformations 3.1

Definition 3.1.1: Linear Transformation

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

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

Theorem 3.1.1

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

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

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

This is a linear transformation indeed.

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

Definition 3.1.2: Null Space and Range Space

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

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

🛉 Note:- 🛉

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

Definition 3.1.3

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

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

Theorem 3.1.2 Rank-Nullity Theorem

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

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

Suppose $\sum_{i=k+1}^n c_i T(\nu_i) = 0$ for some $c_i \in F$. Then, we have $T\left(\sum_{i=k+1}^n c_i \nu_i\right) = 0$; hence $\sum_{i=k+1}^n c_i \nu_i \in \ker T$. Since $\{\nu_1, \dots, \nu_k\}$ is a basis of $\ker T$, we have $\sum_{i=k+1}^n c_i \nu_i = \sum_{i=1}^k a_i \nu_i$ for some a_i 's. Therefore, since $\{\nu_1, \dots, \nu_n\}$ is linearly independent, all c_i 's and a_i 's are zero. This implies that $\{T(\nu_i)\}_{i=k+1}^n$ is linearly independent.

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

The two paragraphs imply that rank T = n - k.

Theorem 3.1.3

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

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

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

3.2 The Algebra of Linear Transformations

Definition 3.2.1

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

Theorem 3.2.1

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

Theorem 3.2.2

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

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

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

Then,

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

• They span L(V, W).

Lemma 3.2.1

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

Definition 3.2.2: Linear Operator (Endomorphism)

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

Note:-

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

Definition 3.2.3: Injectivity and Surjectivity

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

- injective (or, nonsingular) if $T(v) = 0 \implies v = 0$.
- *surjective* if T(V) = W.
- *invertible* if \exists linear transform $U: W \to V$, $U \circ T = id_V \wedge T \circ U = id_W$.

Exercise 3.2.1

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

Exercise 3.2.2

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

Exercise 3.2.3

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

Theorem 3.2.3

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

- (i) *T* is invertible.
- (ii) *T* is injective.
- (iii) T is surjective.

Proof. T is injective \iff nullity T=0 \iff rank T=n \iff Im T=W \iff T is onto

Definition 3.2.4: General Linear Group

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

Note:-

The general linear group is actually a group.

3.3 Isomorphism

Definition 3.3.1: Isomorphism

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

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

Theorem 3.3.1

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

Proof. Let $B = \{\alpha_1; \dots; \alpha_n\}$ be a basis of V. Define $T: V \to F^n$ by $v \mapsto [v]_B$. Suppose T(v) = 0. Then, $v = 0 \cdot \alpha_1 + \dots + 0 \cdot \alpha_n = 0$. Hence, T is injective. By Theorem 3.2.3, T is isomorphism.

3.4 Representation of Transformation by Matrices

Theorem 3.4.1

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

Proof. $A = [[T(v_1)]_{B'} \ [T(v_2)]_{B'} \ \cdots \ [T(v_n)]_{B'}]$ where v_i is the i^{th} basis vector of B.

Theorem 3.4.2

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

Theorem 3.4.3

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

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

Definition 3.4.1: Similar Matrices

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

Proof.
$$[T(v)]_{B_1} = [T]_{B_1}[v]_{B_1} = [T]_{B_1}P[v]_{B_2}$$
. $[T(v)]_{B_1} = P[T(v)]_{B_2} = P[T]_{B_2}[v]_{B_2}$. Since v was arbitrary, $P[T]_{B_2} = [T]_{B_1}P$.

• Note:- •

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

3.5 **Linear Functionals**

Definition 3.5.1: Linear Functional

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

Definition 3.5.2: Dual Vector Space

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

Note:-

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

Lemma 3.5.1

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

Proof. Since dim $V^* = \dim V = n$, we only need to show that the set is linearly independent. Suppose $\sum_{i=1}^{n} c_i f_i = 0$ for some $c_i \in F$. Then, for each $j \in [n]$, as $f_i(\alpha_j) = \delta_{ij}$, 0 = 0 $\left(\sum_{i=1}^{n} c_i f_i\right)(\alpha_j) = c_j f_j(\alpha_j) = c_j$. Hence, they are linearly independent.

Definition 3.5.3: Dual Basis

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

Lemma 3.5.2

Let *V* be a finite-dimensional vector space over *F* and let $n = \dim V$. Let $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ be a basis for V. Let $\{f_1, \dots, f_n\} \subseteq V^*$ be the dual basis of it. (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) There exists $x_i \in F$ such that $f = \sum_{i=1}^n x_i f_i$. Evaluating at α_j for each $j \in [n]$, we get $f(\alpha_i) = x_i$.
- (ii) There exists $y_i \in F$ such that $v = \sum_{i=1}^n y_i \alpha_i$. Applying f_j for each $j \in [n]$, we get $f_i(v) = y_i$.

Definition 3.5.4: Hyperspace

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

Example 3.5.1

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

Definition 3.5.5: Annihilator

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

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

🛉 Note:- 🛉

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

Theorem 3.5.1

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

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

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

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

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

Corollary 3.5.1

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

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

Corollary 3.5.2

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

3.6 The Double Dual

Note:-

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

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

Then, define \mathcal{L} by

$$\mathcal{L}: V \longrightarrow V^{**}$$
$$\alpha \longmapsto L_{\alpha}.$$

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

Theorem 3.6.1

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

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

Definition 3.6.1: Proper Subspace

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

Definition 3.6.2: Maximal Subspace

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

Definition 3.6.3: Hyperspace

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

Note:-

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

Theorem 3.6.2

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

Proof. ker f is proper, since, otherwise, f = 0.

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

Theorem 3.6.3

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

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

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

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

🛉 Note:- 🛉

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

The Transpose of a Linear Transformation 3.7

Definition 3.7.1: Transpose

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

Lemma 3.7.1

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

Theorem 3.7.1

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

Proof. Let $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$, $\mathcal{B}' = \{\beta_1, \dots, \beta_m\}$, $\mathcal{B}^* = \{f_1, \dots, f_n\}$, and $\mathcal{B}'^* = \{g_1, \dots, g_m\}$. Then, we have $T\alpha_j = \sum_{i=1}^m a_{ij}\beta_i$ for each $j \in [n]$ and $T^t g_j = \sum_{i=1}^n b_{ij}f_i$ for each $j \in [m]$. For each $i \in [n]$ and $j \in [m]$, $(T^t g_j)(\alpha_i) = g_j T\alpha_i = g_j \left(\sum_{k=1}^m a_{ki}\beta_k\right) = \sum_{k=1}^m a_{ki}g_j(\beta_k) = \alpha_{ji}$. Hence, since $T^t g_j$ is a linear functional on V, $T^t g_j = \sum_{i=1}^n (T^t g_j)(\alpha_i)f_i = \sum_{i=1}^n \alpha_{ji}f_i$. Therefore, $a_{ij} = b_{ji}$ for each $i \in [n]$ and $j \in [m]$.

Theorem 3.7.2

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

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

Proof.

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

Polynomials

4.1 Algebras

Definition 4.1.1: *F*-algebra

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

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

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

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

Definition 4.1.2: Polynomial

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

Definition 4.1.3: Degree

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

Theorem 4.1.1

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

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

Theorem 4.1.2 Euclidean Algorithm

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

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

Definition 4.1.4: Divisibility

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

Lemma 4.1.1

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

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

Definition 4.1.5: Evaluation

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

Definition 4.1.6: Ideal

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

Definition 4.1.7: Principal Ideal

An ideal of the form

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

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

Theorem 4.1.3

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

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

We shall claim that $M = (g_0)$. Take any $f \in M$. By the Euclidean algorithm, $\exists q, r \in F[x]$, $f = g_0q + r$ with either r = 0 or $\deg r < \deg g_0$. If $r \neq 0$, then $r = f - g_0q \in M$ but $\deg r < \deg g_0$, which contradicts the minimality of $\deg g_0$. Hence, r = 0, and thus $f = g_0q \in (g_0)$.

🛉 Note:- 🛉

By putting "monic" assumption, such g_0 is unique as well.

Corollary 4.1.1

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

- (i) $p_1F[x] + p_2F[x] + \cdots + p_nF[x] = (g_0)$
- (ii) $\forall i \in [n], g_0 \mid p_i$
- (iii) $(\forall i \in [n], f \mid p_i) \Longrightarrow f \mid g_0$

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

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

Definition 4.1.8: Relatively Prime

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

Definition 4.1.9: Reducibility

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

Definition 4.1.10: Prime Element

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

Example 4.1.1

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

Theorem 4.1.4

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

Proof.

- (⇒) Suppose $p \mid gh$ for some $g,h \in F[x]$. If g or h is zero, then it is done. Hence, WMA that $g,h \neq 0$. Let (p,g) = (d). $d \mid p$ implies that d=1 or d=p since p is irreducible. If d=p, then $d \mid g$, i.e., $p \mid g$. If d=1, then there exists p_0, g_0 such that $pp_0 + gg_0 = 1$. Hence, $php_0 + ghg_0 = h$. Hence, $p \mid h$.
- (⇐) Suppose p is reducible. Then, p = gh for some g, h with nonzero degrees. Since p is prime, $p \mid g$ or $p \mid h$. This implies $\deg p \leq \deg g$ or $\deg p \leq \deg h$. This is a contradiction since $\deg p = \deg g + \deg h \leq 2 \deg p$ arises.

Theorem 4.1.5 Unique Factorization of Polynomials

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

Proof. WLOG, *f* is monic.

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

Suppose $\deg f > 1$. Suppose the theorem holds for all $g \in F[x]$ with $\deg g < \deg f$. If f is itself irreducible, then done. Otherwise, there are $g_1, g_2 \in F[x]$ with $\deg g_i \ge 1$ such that $f = g_1g_2$. Then, $\deg g_1$ and $\deg g_2$ are less than f. Hence, $g_1 = p_1p_2\cdots p_r$ and $g_2 = q_1q_2\cdots q_s$ where p_i and q_i are irreducible, yielding $f = p_1\cdots p_rq_1\cdots q_s$.

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

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

Definition 4.1.11: (Formal) Derivative

For $f(x) = a_0 + a_1 x + \dots + a_n x^n \in F[x]$, we define

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

Note:-

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

Theorem 4.1.6

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

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

Definition 4.1.12: Algebraically Closed

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

Note:-

F is algebraically closed.

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

Note:-

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

Determinants

5.1 Determinant Functions

Definition 5.1.1: *n*-linear and Iterating

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

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

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

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

Note:-

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

Definition 5.1.2: Determinant

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

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

Definition 5.1.3: Minor Matrix

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

Theorem 5.1.1

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

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

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

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

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

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

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

$$E_{j}(A) = (-1)^{k+j} A_{kj} D_{kj}(A) + (-1)^{k+j+1} A_{(k+1),j} D_{(k+1),j}(A)$$
$$= (-1)^{k+j} (A_{kj} D_{kj}(A) - A_{kj} D_{kj}(A)) = 0$$

Claim.
$$E_j(I_n) = 1$$
. $I_n(i \mid j) = I_{n-1}$.

Corollary 5.1.1

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

Definition 5.1.4: Permutation

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

Definition 5.1.5: Transposition

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

Definition 5.1.6: Cycle

A cycle is like:

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

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

Note:-

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

Theorem 5.1.2

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

Definition 5.1.7: Sign of Permutation

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

Corollary 5.1.2

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

Theorem 5.1.3

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

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

Proof. Let e_1, \dots, e_n be the rows of I_n . For $A \in K^{n \times n}$, let α_i be the *i*-th rows of A. Then,

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

$$\mathcal{D}(A) = \mathcal{D}(\alpha_1, \alpha_2, \dots, \alpha_n)$$

$$= \mathcal{D}\left(\sum_{j=1}^n A_{1j} e_j, \alpha_2, \dots, \alpha_n\right)$$

$$= \sum_{j=1}^n A_{1j} \mathcal{D}(e_j, \alpha_2, \dots, \alpha_n)$$

$$= \dots = \sum_{j=1}^n \sum_{j=1}^n \sum_{j=1}^n \dots \sum_{j_n=1}^n A_{1j_1} A_{2j_2} \dots A_{n,j_n} \mathcal{D}(e_{j_1}, \dots, e_{j_n})$$

$$= \sum_{\alpha \in S} \operatorname{sign}(\sigma) \prod_{i \in [n]} A_{i,\sigma(i)}$$

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

Corollary 5.1.3

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

Corollary 5.1.4

Any cofactor expansion gives the same value.

Corollary 5.1.5

 $\det A^t = \det A$

Proof. Theorem 3.7.1 and Theorem 5.1.3.

Exercise 5.1.1

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

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

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

Definition 5.1.8: Adjoint Matrix

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

Corollary 5.1.6

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

Elementary Canonical Forms

6.1 Eigenvalues

Definition 6.1.1: Eigenvalue

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

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

Theorem 6.1.1

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

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

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

Definition 6.1.2: Characteristic Polynomial

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

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

Note:-

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

Note:-

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

Corollary 6.1.1

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

Proof. Every polynomial of degree n has at most n solutions.

Definition 6.1.3: Diagonalizability

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

Note:-

- If B = {v₁, ···, vₙ} and Tvᵢ = cᵢvᵢ for each i ∈ [n], then [T]_B = diag(c₁, c₂, ···, cₙ).
 If T ∈ L(V) is diagonalizable, then the characteristic polynomial can be completely decomposed into a product of linear factors.