

Chapter 1

Vector Spaces

1.1 Fields

Definition 1.1. A **field** is a set F with two operations, called **addition** (denoted by $+$) and **multiplication** (denoted by \cdot), which satisfy the following axioms.

- (A 1) If $a \in F$ and $b \in F$, then $a + b \in F$.
- (A 2) $a + b = b + a$ for all $a, b \in F$.
- (A 3) $(a + b) + c = a + (b + c)$ for all $a, b, c \in F$.
- (A 4) There is an element 0_F in F such that $0_F + a = a$ for all $a \in F$.
- (A 5) For each $a \in F$ there is an element $-a$ in F such that $a + (-a) = 0_F$.
- (M 1) If $a \in F$ and $b \in F$, then $a \cdot b \in F$.
- (M 2) $a \cdot b = b \cdot a$ for all $a, b \in F$.
- (M 3) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in F$.
- (M 4) There is an element 1_F in $F \setminus \{0_F\}$ such that $1_F \cdot a = a$ for all $a \in F$.
- (M 5) For each $a \in F \setminus \{0_F\}$ there is an element a^{-1} in F such that $a \cdot a^{-1} = 1_F$.
- (D) $a \cdot (b + c) = a \cdot b + a \cdot c$ for all $a, b, c \in F$.

Remark.

- For simplification, we usually write ab instead of $a \cdot b$.
- The axioms labeled with “A” and “M” are usually called the **axioms of addition** and the **axioms of multiplication**, respectively. The axiom labeled with “D” is the **distributive law**.
- The elements 0_F and 1_F are usually called the **additive identity** and the **multiplicative identity** of F , respectively. Also, $-a$ and a^{-1} are called the **additive inverse** and the **multiplicative inverse** of a , respectively.
- **Subtraction** and **division** can be defined using additive and multiplicative inverses.

Example. \mathbb{Q} , \mathbb{R} and \mathbb{C} are fields.

Example. Let $\mathbb{B} = \{0, 1\}$ and the operations \oplus and \odot are defined as follows.

$$\begin{array}{c|cc} \oplus & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \quad \begin{array}{c|cc} \odot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

Then \mathbb{B} is a field with \oplus and \odot as addition and multiplication, respectively.

Proposition 1.2. Let F be a field with $a, b, c \in F$.

- (a) If $a + b = a + c$, then $b = c$.
- (b) If $a + b = a$, then $b = 0_F$.
- (c) If $a + b = 0_F$, then $b = -a$.
- (d) $-(-a) = a$.

Proof.

- (a) It can be proved by

$$\begin{aligned} b &= 0_F + b \\ &= (-a + a) + b \\ &= -a + (a + b) \\ &= -a + (a + c) \\ &= (-a + a) + c \\ &= 0_F + c \\ &= c. \end{aligned}$$

- (b) By applying (a), it follows from $a + b = a + 0_F$ that $b = 0_F$.
- (c) By applying (a), it follows from $a + b = a + (-a)$ that $b = -a$.
- (d) Since $-a + a = 0_F$, we have $a = -(-a)$ by (c). □

Proposition 1.3. Let F be a field with $a, b, c \in F$ and $a \neq 0_F$.

- (a) If $a \cdot b = a \cdot c$, then $b = c$.
- (b) If $a \cdot b = a$, then $b = 1_F$.
- (c) If $a \cdot b = 1_F$, then $b = a^{-1}$.
- (d) $(a^{-1})^{-1} = a$.

Proof. The proof is omitted since it is similar to that of Proposition 1.2. □

Proposition 1.4. Let F be a field with $a, b \in F$.

- (a) $0_F \cdot a = 0_F$.

$$(b) \quad (-a) \cdot b = -(a \cdot b) = a \cdot (-b).$$

$$(c) \quad (-a) \cdot (-b) = a \cdot b.$$

Proof.

(a) Since

$$0_F \cdot a + 0_F \cdot a = (0_F + 0_F) \cdot a = 0_F \cdot a,$$

we have $0_F \cdot a = 0_F$ by Proposition 1.2 (b).

(b) Since

$$(-a) \cdot b + a \cdot b = (-a + a) \cdot b = 0_F \cdot b = 0_F,$$

we have $(-a) \cdot b = -(a \cdot b)$ by Proposition 1.2 (c). The other half can be proved similarly.

(c) By applying (b) twice, we have

$$(-a) \cdot (-b) = -(a \cdot (-b)) = -(-(a \cdot b)) = a \cdot b. \quad \square$$

1.2 Vector Spaces

Definition 1.5. A **vector space** over a field F is a set V with two operations, called **addition** (denoted by $+$) and **scalar multiplication** (denoted by \cdot), which satisfy the following axioms.

(V 1) If $x \in V$ and $y \in V$, then $x + y \in V$.

(V 2) $x + y = y + x$ for all $x, y \in V$.

(V 3) $(x + y) + z = x + (y + z)$ for all $x, y, z \in V$.

(V 4) There is an element 0_V in V such that $0_V + x = x$ for all $x \in V$.

(V 5) For each $x \in V$ there is an element $-x$ such that $x + (-x) = 0_V$.

(V 6) If $a \in F$ and $x \in V$, then $a \cdot x \in V$.

(V 7) $(a \cdot b) \cdot x = a \cdot (b \cdot x)$ for all $a, b \in F$ and $x \in V$.

(V 8) $1_F \cdot x = x$ for all $x \in V$.

(V 9) $a \cdot (x + y) = a \cdot x + a \cdot y$ for all $a \in F$ and $x, y \in V$.

(V 10) $(a + b) \cdot x = a \cdot x + b \cdot x$ for all $a, b \in F$ and $x \in V$.

Remark.

- For simplification, we usually write ax instead of $a \cdot x$.
- The element 0_V is usually called the **additive identity** of V , and $-x$ is called the **additive inverse** of x in V .
- **Subtraction** can be defined using additive inverses.

Examples.

- A field is a vector space over itself, e.g., \mathbb{R} is a vector space over \mathbb{R} .
- \mathbb{C} is a vector space over \mathbb{R} .
- \mathbb{R} is a vector space over \mathbb{Q} .

Examples.

- The set of **n -tuples** with elements from a field F is denoted by F^n . For $x = (x_1, \dots, x_n) \in F^n$, $y = (y_1, \dots, y_n) \in F^n$, and $c \in F$, we define the operations of addition and scalar multiplication by

$$x + y = (x_1 + y_1, \dots, x_n + y_n) \quad \text{and} \quad c \cdot x = (c \cdot x_1, \dots, c \cdot x_n).$$

Then F^n is a vector space over F .

- The set of all $m \times n$ **matrices** with elements from a field F is denoted by $F^{m \times n}$. For $A, B \in F^{m \times n}$ and $c \in F$, we define the operations of addition and scalar multiplication by

$$(A + B)_{ij} = A_{ij} + B_{ij} \quad \text{and} \quad (c \cdot A)_{ij} = c \cdot A_{ij}$$

for $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$. Then $F^{m \times n}$ is a vector space over F .

- The set of **functions** from a nonempty set S to a field F is denoted by $\mathcal{F}(S, F)$. For $f, g \in \mathcal{F}(S, F)$ and $c \in F$, we define the operations of addition and scalar multiplication by

$$(f + g)(s) = f(s) + g(s) \quad \text{and} \quad (c \cdot f)(s) = c \cdot f(s)$$

for all $s \in S$. Then $\mathcal{F}(S, F)$ is a vector space over F .

- The set of **polynomials** with coefficients from a field F is denoted by $\mathcal{P}(F)$. For $f, g \in \mathcal{P}(F)$ and $c \in F$ with

$$f(t) = \sum_{i=0}^n a_i t^i \quad \text{and} \quad g(t) = \sum_{i=0}^n b_i t^i,$$

we define the operations of addition and scalar multiplication by

$$(f + g)(t) = \sum_{i=0}^n (a_i + b_i) t^i \quad \text{and} \quad (c \cdot f)(t) = \sum_{i=0}^n (c \cdot a_i) t^i.$$

Then $\mathcal{P}(F)$ is a vector space over F .

Proposition 1.6. Let V be a vector space with $x, y, z \in V$.

- (a) If $x + y = x + z$, then $y = z$.
- (b) If $x + y = x$, then $y = 0_V$.
- (c) If $x + y = 0_V$, then $y = -x$.
- (d) $-(-x) = x$.

Proof. The proof is omitted since it is similar to that of Proposition 1.2. □

Proposition 1.7. Let V be a vector space over a field F with $x \in V$ and $a \in F$.

- (a) $0_F \cdot x = 0_V$.
- (b) $a \cdot 0_V = 0_V$.
- (c) $(-a) \cdot x = -(a \cdot x) = a \cdot (-x)$.

Proof. The proof is omitted since it is similar to that of Proposition 1.4. □

1.3 Subspaces

Definition 1.8. Let V be a vector space over a field F . Then a subset W of V is called a **subspace** of V if W is a vector space over F with the operations of addition and scalar multiplication defined on V .

Theorem 1.9. Let V be a vector space over a field F and $W \subseteq V$. Then W is a subspace of V if the following conditions hold.

- (a) $0_V \in W$.
- (b) $x + y \in W$ for all $x, y \in W$.
- (c) $ax \in W$ for all $x \in W$ and $a \in F$.

Proof. Since a vector in W is also in V , (V 2), (V 3), (V 7), (V 8), (V 9) and (V 10) in Definition 1.5 hold trivially. Furthermore, (a) implies (V 4), (b) implies (V 1), (c) implies (V 6), and (V 5) is also true since

$$-x = -(1_F x) = (-1_F)x \in W$$

holds for all $x \in W$. Thus, W is a vector space over F . □

Corollary 1.10. Let V be a vector space over a field F and $W \subseteq V$. Then W is a subspace of V if and only if the following conditions hold.

- (a) $0_V \in W$.
- (b) $ax + y \in W$ for all $x, y \in W$ and $a \in F$.

Proof. (\Rightarrow) Straightforward. (\Leftarrow) For all $x, y \in W$ and $a \in F$, we have

$$x + y = 1_F x + y \in W \quad \text{and} \quad ax = ax + 0_V \in W.$$

Thus, W is a subspace of V by Theorem 1.9. □

Example. The set of polynomials in $\mathcal{P}(F)$ with degree not greater than n is denoted by $\mathcal{P}_n(F)$, where the **degree** of a nonzero polynomial

$$f(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_m t^m$$

is defined to be the largest integer n such that $a_n \neq 0_F$, and the degree of zero polynomial is defined to be -1 . Then one can verify that $\mathcal{P}_n(F)$ is a subspace of $\mathcal{P}(F)$.

Examples.

- An $n \times n$ matrix A is called **diagonal** if $A_{ij} = 0_F$ for all $i, j \in \{1, \dots, n\}$ with $i \neq j$. Then one can verify that the set of $n \times n$ diagonal matrices is a subspace of $F^{n \times n}$.
- The **trace** of an $n \times n$ matrix A , denoted by $\text{tr}(A)$, is defined by

$$\text{tr}(A) = \sum_{i=1}^n A_{ii}.$$

Then one can verify that the set of $n \times n$ matrices that have trace equal to 0_F is a subspace of $F^{n \times n}$.

Proposition 1.11. Let V be a vector space and let W_1 and W_2 be subspaces of V . Then $W_1 \cap W_2$ is a subspace of V .

Proof. Since W_1 and W_2 are subspaces of V , we have $0_V \in W_1 \cap W_2$. Furthermore, for each $x, y \in W_1 \cap W_2$ and for each $a \in F$, we have $ax + y \in W_1 \cap W_2$ by Corollary 1.10. Thus, $W_1 \cap W_2$ is a subspace of V . \square

Example. Let W_1 be the set of $n \times n$ diagonal matrices. Let W_2 be the set of $n \times n$ matrices that have trace equal to 0_F . Then since both W_1 and W_2 are subspaces of $F^{n \times n}$, we can conclude that $W_1 \cap W_2$ is also a subspace of $F^{n \times n}$.

Definition 1.12. Let V be a vector space and let $S_1, S_2 \subseteq V$. Then the **sum** of S_1 and S_2 , denoted by $S_1 + S_2$, is the set

$$\{x + y : x \in S_1 \text{ and } y \in S_2\}.$$

Proposition 1.13. Let V be a vector space and let W_1 and W_2 be subspaces of V . Then the following statements are true.

- (a) $W_1 + W_2$ is a subspace of V .
- (b) If U is a subspace of V with $W_1 \cup W_2 \subseteq U$, then $W_1 + W_2 \subseteq U$.

Proof.

- (a) We have $0_V = 0_V + 0_V \in W_1 + W_2$. For each $x, y \in W_1 + W_2$ and for each $a \in F$, by Definition 1.12 there exist $x_1, y_1 \in W_1$ and $x_2, y_2 \in W_2$ such that $x = x_1 + x_2$ and $y = y_1 + y_2$. Thus,

$$\begin{aligned} ax + y &= a(x_1 + x_2) + (y_1 + y_2) \\ &= (ax_1 + ax_2) + (y_1 + y_2) \\ &= (ax_1 + y_1) + (ax_2 + y_2) \\ &\in W_1 + W_2. \end{aligned}$$

- (b) Let x be a vector in $W_1 + W_2$. Then by Definition 1.12 there exists $x_1 \in W_1$ and $x_2 \in W_2$ such that $x = x_1 + x_2$. We have $x_1 \in U$ since $W_1 \subseteq U$. Also, we have $x_2 \in U$ since $W_2 \subseteq U$. It follows that $x = x_1 + x_2 \in U$, and thus $W_1 + W_2 \subseteq U$. \square

1.4 Spanning Sets

Definition 1.14. Let V be a vector space over a field F and let $S \subseteq V$. Then a vector $x \in V$ is called a **linear combination** of S if there exist scalars $a_1, \dots, a_n \in F$ and vectors $x_1, \dots, x_n \in S$ for some nonnegative integer n such that

$$x = \sum_{i=1}^n a_i x_i.$$

Remark.

- If $n = 0$, then the sum in the right hand side is 0_V since nothing are added up. Thus, 0_V is a linear combination of any subset of V .
- Note that n should be finite. Thus, in the vector space \mathbb{R} over the field \mathbb{Q} , e is not a linear combination of \mathbb{Q} even if we have

$$e = \sum_{i=0}^{\infty} \frac{1}{i!}.$$

Definition 1.15. Let V be a vector space over a field F and let $S \subseteq V$. Then the **span** of S , denoted $\text{span}(S)$, is defined as the set of all linear combinations of S .

Theorem 1.16. Let V be a vector space over F and let $S \subseteq V$. Then the following statements are true.

- $\text{span}(S)$ is a subspace of V .
- If U is a subspace of V such that $S \subseteq U$, then $\text{span}(S) \subseteq U$.

Proof.

- Let $c \in F$ and $x, y \in \text{span}(S)$. Then there exist scalars $a_1, \dots, a_n \in F$ and vectors $x_1, \dots, x_n \in S$ such that

$$x = a_1 x_1 + \dots + a_n x_n.$$

Also, there exist scalars $b_1, \dots, b_m \in F$ and vectors $y_1, \dots, y_m \in S$ such that

$$y = b_1 y_1 + \dots + b_m y_m.$$

Thus, we have

$$\begin{aligned} cx + y &= c(x_1 + \dots + x_n) + (y_1 + \dots + y_m) \\ &= cx_1 + \dots + cx_n + y_1 + \dots + y_m \\ &\in \text{span}(S). \end{aligned}$$

Furthermore, $0_V \in \text{span}(S)$. Hence, $\text{span}(S)$ is a subspace of V by Corollary 1.10.

- Let $x \in \text{span}(S)$. Then there exist scalars $a_1, \dots, a_n \in F$ and vectors $x_1, \dots, x_n \in S$ such that

$$x = a_1 x_1 + \dots + a_n x_n.$$

Since $S \subseteq U$, we have $x_1, \dots, x_n \in U$, and it follows that $x = a_1 x_1 + \dots + a_n x_n \in U$ due to the closeness of U . Thus, $\text{span}(S) \subseteq U$. \square

Definition 1.17. Let V be a vector space and let $S \subseteq V$. If $\text{span}(S) = V$, then S is called a **spanning set** of V , and we also say S **spans** V .

Example. $\{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$ is a spanning set of \mathbb{R}^3 since for any $x, y, z \in \mathbb{R}$,

$$(x, y, z) = \frac{-x + y + z}{2} \cdot (0, 1, 1) + \frac{x - y + z}{2} \cdot (1, 0, 1) + \frac{x + y - z}{2} \cdot (1, 1, 0).$$

Proposition 1.18. Let V be a vector space and let $R, S \subseteq V$.

- (a) $S \subseteq \text{span}(S)$.
- (b) If $R \subseteq S$, then $\text{span}(R) \subseteq \text{span}(S)$.
- (c) $S = \text{span}(S)$ if and only if S is a subspace of V .
- (d) $\text{span}(R \cup S) = \text{span}(R) + \text{span}(S)$.

Proof.

- (a) Straightforward.
- (b) It is true since a linear combination of a subset of S is also a linear combination of S .
- (c) (\Rightarrow) Straightforward from Theorem 1.16 (a).
 (\Leftarrow) Note that any linear combination of S is in S due to closeness of addition and scalar multiplication in S . Thus, $\text{span}(S) \subseteq S$, and it follows that $S = \text{span}(S)$.
- (d) Since $R \subseteq \text{span}(R)$ and $S \subseteq \text{span}(S)$, we have $R \cup S \subseteq \text{span}(R) + \text{span}(S)$. Thus, by Theorem 1.16, we have $\text{span}(R \cup S) \subseteq \text{span}(R) + \text{span}(S)$. On the other side, since

$$\text{span}(R) \subseteq \text{span}(R \cup S) \quad \text{and} \quad \text{span}(S) \subseteq \text{span}(R \cup S),$$

we can conclude that $\text{span}(R) + \text{span}(S) \subseteq \text{span}(R \cup S)$. Thus, $\text{span}(R) + \text{span}(S) \subseteq \text{span}(R \cup S)$ by Proposition 1.13. \square

1.5 Linearly Independent Sets

Definition 1.19. Let V be a vector space over a field F and let $S \subseteq V$.

- S is **linearly dependent** if there exist scalars $a_1, a_2, \dots, a_n \in F \setminus \{0_F\}$ and distinct vectors $x_1, x_2, \dots, x_n \in S$ for some positive integer n such that

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0_V.$$

- S is **linearly independent** if it is not linearly dependent.

Remark.

- Note that \emptyset is linearly independent.

Theorem 1.20. Let V be a vector space over a field F and let $S \subseteq V$. Then the following statements are equivalent.

- (a) S is linearly dependent.
- (b) There exists $x \in S$ with $x \in \text{span}(S \setminus \{x\})$.
- (c) There exists $x \in S$ with $\text{span}(S) = \text{span}(S \setminus \{x\})$.

Proof.

- (i) First we assume (a) and prove (b). Suppose that

$$a_0x_0 + a_1x_1 + \cdots + a_nx_n = 0_V,$$

where a_0, a_1, \dots, a_n are nonzero scalars and x_0, x_1, \dots, x_n are distinct vectors. Then

$$\begin{aligned} x_0 &= (-a_0)^{-1}(a_1x_1 + \cdots + a_nx_n) \\ &= ((-a_0)^{-1}a_1)x_1 + \cdots + ((-a_0)^{-1}a_n)x_n \\ &\in \text{span}(S \setminus \{x_0\}). \end{aligned}$$

- (ii) Then we assume (b) and prove (c). Since

$$x \in \text{span}(S \setminus \{x\}) \quad \text{and} \quad S \setminus \{x\} \subseteq \text{span}(S \setminus \{x\}),$$

we have $S \subseteq \text{span}(S \setminus \{x\})$. Thus, $\text{span}(S) \subseteq \text{span}(S \setminus \{x\})$ by Theorem 1.16, and we can conclude that $\text{span}(S) = \text{span}(S \setminus \{x\})$.

- (iii) Then we assume (c) and prove (b). It is straightforward since $x \in S \subseteq \text{span}(S) = \text{span}(S \setminus \{x\})$.
- (iv) Finally we assume (b) and prove (a). Without loss of generality, let $a_1, \dots, a_n \in F$ be nonzero scalars and $x_1, \dots, x_n \in S \setminus \{x\}$ be distinct vectors such that $x = a_1x_1 + \cdots + a_nx_n$. Then we have

$$(-1_F)x + a_1x_1 + \cdots + a_nx_n = 0_V,$$

which completes the proof. □

Example. Let $S = \{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$ be a subset of \mathbb{R}^3 . Suppose that $a_1, a_2, a_3 \in \mathbb{R}$ are scalars such that

$$a_1(0, 1, 1) + a_2(1, 0, 1) + a_3(1, 1, 0) = (0, 0, 0).$$

Then we have the following system of equations.

$$\begin{aligned} a_2 + a_3 &= 0 \\ a_1 + a_3 &= 0 \\ a_1 + a_2 &= 0 \end{aligned}$$

Since the only solution to this system of equations is $a_1 = a_2 = a_3 = 0$, we can conclude that S is linearly independent by Definition 1.19.

Example. Let $S = \{(1, 1, 1), (0, 1, 1), (1, 0, 1), (1, 1, 0)\}$ be a subset of \mathbb{R}^3 . We can conclude that S is linearly dependent since

$$(1, 1, 1) = \frac{1}{2} \cdot (0, 1, 1) + \frac{1}{2} \cdot (1, 0, 1) + \frac{1}{2} \cdot (1, 1, 0).$$

Proposition 1.21. Let V be a vector space and let R, S be subsets of V with $R \subseteq S$.

- (a) If R is linearly dependent, then so is S .
- (b) If S is linearly independent, then so is R .

Proof.

- (a) Suppose that R is linearly dependent. Then by Definition 1.19 there exists $x \in R$ such that $x \in \text{span}(R \setminus \{x\})$. Also, we have $R \setminus \{x\} \subseteq S \setminus \{x\}$ since $R \subseteq S$. Thus, $x \in \text{span } S \setminus \{x\}$, and it follows that S is linearly dependent.
- (b) Straightforward from (a). □

1.6 Bases and Dimension

Definition 1.22. A **basis** for a vector space V is a linearly independent subset of V that spans V .

Examples.

- \emptyset is a basis for $\{0_V\}$.
- $\{e_1, \dots, e_n\}$ is a basis for F^n , where e_i is the n -tuple whose i -th component is 1_F and the other components are all 0_F .
- $\{E_{ij} : 1 \leq i \leq m \text{ and } 1 \leq j \leq n\}$ is a basis for $F^{m \times n}$, where E_{ij} is the matrix whose (i, j) -entry is 1_F and the other entries are all 0_F .
- $\{t^0, t^1, t^2, \dots, t^n\}$ is a basis for $\mathcal{P}_n(F)$.
- $\{t^0, t^1, t^2, \dots\}$ is a basis for $\mathcal{P}(F)$.

Proposition 1.23. Let V be a vector space. If there exists a finite set S that spans V , then there is a subset Q of S that is a finite basis of V .

Proof. The proof is by induction on $|S|$. For the induction basis, suppose that $|S| = 0$, i.e., $S = \emptyset$. Then the proposition holds since one can choose $Q = \emptyset$ as a basis for V .

Now assume the induction hypothesis that the proposition holds for $|S| = n$ with $n \geq 0$. If S is linearly independent, then we can choose $Q = S$ as a basis for V . Otherwise, there exists $x \in S$ with $\text{span}(S \setminus \{x\}) = \text{span}(S)$, i.e., $S \setminus \{x\}$ spans V . Thus, by induction hypothesis there is a subset Q of $S \setminus \{x\}$ that is a basis for V , which completes the proof. \square

Theorem 1.24 (Replacement Theorem). Let V be a vector space over a field F . Let S be a finite set that spans V , and let $Q \subseteq S$ be a finite linearly independent set. Then $|Q| \leq |S|$, and there exists $R \subseteq S \setminus Q$ such that both $|Q \cup R| = |S|$ and $\text{span}(Q \cup R) = V$ hold.

Proof. The proof is based on induction on $|Q|$. The theorem holds for $|Q| = 0$, i.e., $Q = \emptyset$, since we have $|\emptyset| \leq |S|$, $|\emptyset \cup S| = |S|$ and $\text{span}(\emptyset \cup S) = V$.

Now suppose that the theorem is true for $|Q| = m$ with $m \geq 0$, and we prove that the theorem holds for $|Q| = m + 1$. Let $Q = \{x_1, \dots, x_{m+1}\}$ and let $Q' = \{x_1, \dots, x_m\}$. By induction hypothesis, there exists $R' = \{y_1, \dots, y_k\} \subseteq S \setminus Q'$ such that $|Q'| + |R'| = |S|$ and $\text{span}(Q' \cup R') = V$. Since $Q' \cup R'$ spans V , there exists $a_1, \dots, a_m, b_1, \dots, b_k \in F$ such that

$$x_{m+1} = \sum_{i=1}^m a_i x_i + \sum_{j=1}^k b_j y_j.$$

If $b_j = 0_F$ for all $j \in \{1, \dots, k\}$, then $x_{m+1} \in \text{span}(Q') = \text{span}(Q \setminus \{x_{m+1}\})$, implying that Q is linearly dependent, contradiction. Thus, there must exist some $j \in \{1, \dots, k\}$ such that $b_j \neq 0_F$. Without loss of generality, suppose that $b_k \neq 0_F$ with $k \geq 1$. Also, let $R = \{y_1, \dots, y_{k-1}\}$. Then $|Q \cup R| = (m+1) + (k-1) = |S|$, and we have $|Q| \leq |S|$. It follows that

$$Q' \cup R' \subseteq Q \cup R \cup \{y_k\} \subseteq \text{span}(Q \cup R),$$

where the second inclusion holds because

$$y_k = (-b_k)^{-1} \left(\sum_{i=1}^m a_i x_i + (-1_F) x_{m+1} + \sum_{j=1}^{k-1} b_j y_j \right) \in \text{span}(Q \cup R).$$

Then, we have

$$V = \text{span}(Q' \cup R') \subseteq \text{span}(Q \cup R) \subseteq V.$$

by Theorem 1.16. Thus, $\text{span}(Q \cup R) = V$, which completes the proof. \square

Corollary 1.25. Let V be a vector space and Q be a linearly independent subset of V that is infinite. Then each spanning set of V is infinite.

Proof. Suppose that there is a finite set S that spans V . Let Q' be a subset of Q with $|Q'| = |S| + 1$. By Proposition 1.21, we can conclude that Q' is also linearly independent. Thus, we have $|Q'| \leq |S|$ by replacement theorem (Theorem 1.24), contradiction. \square

Corollary 1.26. Let V be a vector space. If V has a finite basis, then each basis for V has the same size.

Proof. Let S be a finite basis for V and Q an arbitrary basis for V . Since $V = \text{span}(S)$ and Q is linearly independent, it follows that Q is finite by Corollary 1.25, and thus we have $|Q| \leq |S|$. Also, since $V = \text{span}(Q)$ and S is linearly independent, we have $|S| \leq |Q|$. Thus, $|Q| = |S|$. \square

Definition 1.27. Let V be a vector space.

- V is **finite-dimensional** if it has a finite basis. In this case, the number of vectors in each basis for V is called the **dimension** of V , denoted by $\dim(V)$.
- V is **infinite-dimensional** if it is not finite-dimensional.

Remark.

- If a vector space has a linearly independent subset that is infinite, we can conclude that it is infinite-dimensional by Corollary 1.25.

Examples. One can find the dimension of a vector space by any basis it admits.

- $\dim(\{0_V\}) = 0$.
- $\dim(F^n) = n$.
- $\dim(F^{m \times n}) = mn$.
- $\dim(\mathcal{P}_n(F)) = n + 1$.
- $\mathcal{P}(F)$ is infinite-dimensional.

Examples. Note that the dimension of a vector space depends on its field of scalars.

- Let $V = \mathbb{C}$ be a vector space over \mathbb{R} . Then we have $\dim(V) = 2$ since $\{1, i\}$ is a basis for V .
- Let $W = \mathbb{C}$ be a vector space over \mathbb{C} . Then we have $\dim(W) = 1$ since $\{1\}$ is a basis for V .

Proposition 1.28. Let V be a vector space. Then a subset of V of $n = \dim(V)$ vectors is linearly independent if and only if it is a spanning set of V .

Proof. (\Rightarrow) Suppose that Q is linearly independent with $|Q| = n$. By replacement theorem (Theorem 1.24), there exists $R \subseteq S \setminus Q$ such that $|Q \cup R| = |S|$ and $\text{span}(Q \cup R) = V$. Since $|Q| = |S|$, we have $|R| = 0$, i.e., $R = \emptyset$. Thus, $\text{span}(Q) = V$.

(\Leftarrow) Suppose that S spans V with $|S| = n$. By Proposition 1.23, there is a subset Q of S that is a basis of V . Then we have $|Q| = n$, implying $Q = S$. Thus, S is a basis for V . \square

Proposition 1.29. Let V be a finite-dimensional vector space. Let $S = \{x_1, \dots, x_n\}$ be a basis for V . Then for each $x \in V$, there exist a unique n -tuple $(a_1, \dots, a_n) \in F^n$ with

$$x = a_1x_1 + \dots + a_nx_n.$$

Proof. Since $x \in \text{span}(S)$, there exist scalars $a_1, \dots, a_n \in F$ such that

$$x = a_1x_1 + \dots + a_nx_n.$$

Now we prove the uniqueness. Let $b_1, \dots, b_n \in F$ be scalars with

$$x = b_1x_1 + \dots + b_nx_n.$$

Then we have

$$0_V = (a_1 - b_1)x_1 + \dots + (a_n - b_n)x_n,$$

and it follows that $(a_1 - b_1, a_2 - b_2, \dots, a_n - b_n) = 0_{F^n}$ since S is linearly independent. Thus, $(a_1, \dots, a_n) = (b_1, \dots, b_n)$. \square

Proposition 1.30. Let V be a finite-dimensional vector space. Let V' be a subspace of V . Then the following statements are true.

- (a) $\dim(V') \leq \dim(V)$.
- (b) If $\dim(V') = \dim(V)$, then $V' = V$.

Proof. Let S and S' be bases for V and V' , respectively.

- (a) Since S' is linearly independent and $V = \text{span}(S)$, we have $|S'| \leq |S|$ by replacement theorem (Theorem 1.24). Thus, $\dim(V') \leq \dim(V)$.
- (b) Since S' is linearly independent and $|S'| = \dim(V)$, we have $\text{span}(S') = V$ by Proposition 1.28. Thus, $V' = \text{span}(S') = V$. \square

Example. Let W be the set of $n \times n$ diagonal matrices, which is a subspace of $F^{n \times n}$. Then one can verify that $\{E_{ii} : 1 \leq i \leq n\}$ is a basis for W , where E_{ij} is the matrix whose (i, j) -entry is 1_F and the other entries are 0_F . Thus, $\dim(W) = n$.

Chapter 2

Linear Transformations

2.1 Linear Transformations

Definition 2.1. Let X and Y be sets. Let $f : X \rightarrow Y$ be a function.

- f is **injective** if $T(x) = T(x')$ implies $x = x'$ for all $x, x' \in X$.
- f is **surjective** if there exists $x \in X$ with $T(x) = y$ for each $y \in Y$.
- f is **bijective** if f is injective and surjective.

Definition 2.2. Let V and W be vector spaces over a field F . A transformation $T : V \rightarrow W$ is said to be **linear** if

$$T(ax + y) = aT(x) + T(y)$$

holds for any scalar $a \in F$ and any vectors $x, y \in V$. The set of all linear transformations from V to W is denoted by $\mathcal{L}(V, W)$, and in the case that $V = W$ we write $\mathcal{L}(V)$ for short.

Theorem 2.3. If V and W are vector spaces over a field F , then $\mathcal{L}(V, W)$ is also a vector space over F .

Proof. For any $c \in F$ and $T_1, T_2 \in \mathcal{L}(V, W)$, since

$$\begin{aligned}(cT_1 + T_2)(ax + y) &= cT_1(ax + y) + T_2(ax + y) && \text{(linearity of } cT_1 + T_2\text{)} \\ &= c(aT_1(x) + T_1(y)) + (aT_2(x) + T_2(y)) && \text{(linearity of } T_1 \text{ and } T_2\text{)} \\ &= acT_1(x) + cT_1(y) + aT_2(x) + T_2(y) \\ &= a(cT_1(x) + T_2(x)) + (cT_1(y) + T_2(y)) \\ &= a(cT_1 + T_2)(x) + (cT_1 + T_2)(y) && \text{(linearity of } cT_1 + T_2\text{)}\end{aligned}$$

holds for each $a \in F$ and $x, y \in V$, we have $cT_1 + T_2 \in \mathcal{L}(V, W)$. Furthermore, $0_{\mathcal{F}(V, W)} \in \mathcal{L}(V, W)$. Thus, $\mathcal{L}(V, W)$ is a subspace of $\mathcal{F}(V, W)$. \square

Theorem 2.4. Let V and W be vector spaces over F , and let $T : V \rightarrow W$ be linear. Let S be a subset of V and let U be a subspace of V . Then the following statements are true.

(a) If n is a nonnegative integer, then for $a_1, \dots, a_n \in F$ and $x_1, \dots, x_n \in V$, we have

$$T\left(\sum_{i=1}^n a_i x_i\right) = a_i \sum_{i=1}^n T(x_i).$$

(b) If S spans U , then $T(S)$ spans $T(U)$.

Proof.

(a) The proof is by induction on n . For $n = 0$, it holds trivially. If the statement is true for some $n \geq 0$, then we have

$$\begin{aligned} T(a_1 x_1 + \dots + a_n x_n + a_{n+1} x_{n+1}) &= T(a_1 x_1 + \dots + a_n x_n) + T(a_{n+1} x_{n+1}) \\ &= a_1 T(x_1) + \dots + a_n T(x_n) + a_{n+1} T(x_{n+1}). \end{aligned}$$

Thus, the statement is true for nonnegative integer n .

(b) We prove that $\text{span}(T(S)) = T(U)$. If $y \in \text{span}(T(S))$, then there exist $a_i \in F$, $x_i \in S$ for $i \in \{1, \dots, n\}$ such that

$$y = \sum_{i=1}^n a_i T(x_i) = T\left(\sum_{i=1}^n a_i x_i\right) \in T(U),$$

so $\text{span}(T(S)) \subseteq T(U)$.

If $y \in T(U)$, then there exist $a_i \in F$, $x_i \in S$ for $i \in \{1, \dots, n\}$ such that

$$y = T\left(\sum_{i=1}^n a_i x_i\right) = \sum_{i=1}^n a_i T(x_i) \in \text{span}(T(S)),$$

so $T(U) \subseteq \text{span}(T(S))$. Thus, $\text{span}(T(S)) = T(U)$. □

2.2 Rank and Nullity

Definition 2.5. Let V and W be vector spaces. The **range** of a transformation $T : V \rightarrow W$, denoted by $\mathcal{R}(T)$, is defined by

$$\mathcal{R}(T) = \{y \in W : y = T(x) \text{ for some } x \in V\}.$$

Proposition 2.6. Let V and W be vector spaces over a field F . If $T : V \rightarrow W$ is linear, then $\mathcal{R}(T)$ is a subspace of W .

Proof. For each $a \in F$ and $y, y' \in \mathcal{R}(T)$, there exist $x, x' \in V$ such that $y = T(x)$ and $y' = T(x')$. Since

$$ay + y' = aT(x) + T(x') = T(ax + x'),$$

we have $ay + y' \in \mathcal{R}(T)$. Furthermore, $0_W = T(0_V) \in \mathcal{R}(T)$. Thus, $\mathcal{R}(T)$ is a subspace of W . \square

Definition 2.7. Let V and W be vector spaces. The **null space** of a transformation $T : V \rightarrow W$, denoted by $\mathcal{N}(T)$, is defined by

$$\mathcal{N}(T) = \{x \in V : T(x) = 0_W\}.$$

Proposition 2.8. Let V and W be vector spaces over a field F . If $T : V \rightarrow W$ is linear, then $\mathcal{N}(T)$ is a subspace of V .

Proof. For each $a \in F$ and $x, x' \in \mathcal{N}(T)$, we have

$$T(ax + x') = aT(x) + T(x') = a0_W + 0_W = 0_W.$$

Thus, $ax + x' \in \mathcal{N}(T)$. Furthermore, $0_V \in \mathcal{N}(T)$ since $T(0_V) = 0_W$. Thus, $\mathcal{N}(T)$ is a subspace of V . \square

Definition 2.9. Let V and W be vector spaces. Let $T : V \rightarrow W$ be linear.

- The **rank** of T , denoted by $\text{rank}(T)$, is the dimension of $\mathcal{R}(T)$.
- The **nullity** of T , denoted by $\text{nullity}(T)$, is the dimension of $\mathcal{N}(T)$.

Theorem 2.10 (Rank-nullity Theorem). Let V and W be finite-dimensional vector spaces over F . Let $T : V \rightarrow W$ be linear. Then we have

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

Proof. Let S be a basis for V and Q a basis for $\mathcal{N}(T)$. By replacement theorem (Theorem 1.24), there is $R \subseteq S \setminus Q$ such that $Q \cup R$ is a basis for V . Since

$$|R| = |Q \cup R| - |Q| = \dim(V) - \text{nullity}(T),$$

it suffices to prove that $|R| = \dim(\mathcal{R}(T))$.

If there exist different $x, x' \in R$ with $T(x) = T(x')$, then we have $T(x - x') = T(x) - T(x') = 0_W$, and thus $x - x' \in \mathcal{N}(T) = \text{span}(Q)$. It follows that $x \in \text{span}(Q \cup \{x'\})$, contradiction to the fact that $Q \cup R$ is linearly independent. Thus, $T(x) \neq T(x')$ for any different $x, x' \in R$, and it follows that $|R| = |T(R)|$. We claim that $T(R)$ is a basis for $\mathcal{R}(T)$.

First we prove that $T(R)$ spans $\mathcal{R}(T)$. By Theorem 2.4 (b) and the fact that $T(Q) = \{0_V\}$, we have

$$\begin{aligned}\mathcal{R}(T) &= T(\text{span}(Q \cup R)) \\ &= \text{span}(T(Q \cup R)) \\ &= \text{span}(T(Q)) + \text{span}(T(R)) \\ &= \text{span}(T(R)).\end{aligned}$$

Then we prove that $T(R)$ is linearly independent. Suppose that

$$a_1T(x_1) + \cdots + a_nT(x_n) = 0_W$$

holds for some $a_1, \dots, a_n \in F$ and some different $x_1, \dots, x_n \in R$ with $n \geq 1$. Then by Theorem 2.4 we have $T(a_1x_1 + \cdots + a_nx_n) = 0_W$, and thus $a_1x_1 + \cdots + a_nx_n \in \mathcal{N}(T)$. Hence, there exist some $b_1, \dots, b_m \in F$ and some different $y_1, \dots, y_m \in Q$ such that

$$a_1x_1 + \cdots + a_nx_n = b_1y_1 + \cdots + b_my_m.$$

That is,

$$a_1x_1 + \cdots + a_nx_n + (-b_1)y_1 + \cdots + (-b_m)y_m = 0_V.$$

Since $Q \cup R$ is linearly independent, we have $a_1 = \cdots = a_n = b_1 = \cdots = b_m = 0_F$, implying that $T(R)$ is linearly independent. Thus, $T(R)$ is a basis for $\mathcal{R}(T)$, which completes the proof. \square

2.3 Isomorphisms

Definition 2.11. Let X, Y, Z be sets. Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be functions. Then the **composition** of f and g is the function $gf : X \rightarrow Z$ such that

$$(gf)(x) = g(f(x))$$

for all $x \in X$.

Definition 2.12. The **identity function** over a set X is a function $I_X : X \rightarrow X$ with $I_X(x) = x$ for all $x \in X$.

Definition 2.13. Let X and Y be sets. A function $f : X \rightarrow Y$ is said to be **invertible** if there exists a function $f^{-1} : Y \rightarrow X$, called the **inverse** of f , such that

$$f^{-1}f = I_X \quad \text{and} \quad ff^{-1} = I_Y.$$

Proposition 2.14. Let X and Y be sets. Let $f : X \rightarrow Y$ and $g : Y \rightarrow X$ be functions.

- (a) If f is invertible, then f^{-1} is invertible.
- (b) If f is invertible, then f^{-1} is linear.
- (c) If f is invertible, then either $gf = I_X$ or $fg = I_Y$ implies $g = f^{-1}$.
- (d) f is invertible if and only if f is bijective.

Proof.

- (a) Straightforward from Definition 2.13.
- (b) For $a \in F$ and $y, y' \in Y$, we have

$$\begin{aligned} f^{-1}(ay + y') &= f^{-1}(af(f^{-1}(y)) + f(f^{-1}(y'))) && (ff^{-1} = I_Y) \\ &= f^{-1}(f(af^{-1}(y) + f^{-1}(y'))) && (\text{linearity of } f) \\ &= af^{-1}(y) + f^{-1}(y'). && (f^{-1}f = I_X) \end{aligned}$$

Thus, f^{-1} is linear.

- (c) If $gf = I_X$, then

$$g = gI_Y = g(ff^{-1}) = (gf)f^{-1} = I_X f^{-1} = f^{-1}.$$

If $fg = I_Y$, then

$$g = I_X g = (f^{-1}f)g = f^{-1}(fg) = f^{-1}I_Y = f^{-1}.$$

- (d) (\Rightarrow) Suppose that f is invertible. Then f is injective since for each $x, x' \in X$ with $f(x) = f(x')$, we have

$$x = (f^{-1}f)(x) = f^{-1}(f(x)) = f^{-1}(f(x')) = (f^{-1}f)(x') = x'.$$

Also, f is surjective since for each $y \in Y$, we have

$$y = (ff^{-1})y = f(f^{-1}(y)).$$

(\Leftarrow) If f is bijective, then for each $y \in Y$ there exists a unique element $x \in X$ with $f(x) = y$. Thus, there exists a function $g : Y \rightarrow X$ such that

$$g(f(x)) = x$$

for each $x \in X$. For any $y \in Y$, if $x \in X$ is the element such that $f(x) = y$, then we have

$$f(g(y)) = f(g(f(x))) = f(x) = y.$$

Thus, f is invertible since $gf = I_X$ and $fg = I_Y$. \square

Definition 2.15. Let V and W be vector spaces. An **isomorphism** from V onto W is a invertible linear transformation from V to W . If there is an isomorphism from V onto W , then V and W are said to be **isomorphic**, denoted by $V \cong W$.

Lemma 2.16. Let V and W be finite-dimensional vector spaces with $\dim(V) = \dim(W)$. Let $T : V \rightarrow W$ be linear. Then T is injective if and only if T is surjective.

Proof. (\Rightarrow) If T is injective, then $\mathcal{N}(T) = \{0_V\}$, implying $\text{nullity}(T) = 0$. Then we have

$$\dim(\mathcal{R}(T)) = \text{rank}(T) = \dim(V) - \text{nullity}(T) = \dim(W) - 0 = \dim(W).$$

Since $\mathcal{R}(T)$ is a subspace of W with $\dim(\mathcal{R}(T)) = \dim(W)$, we can conclude that $\mathcal{R}(T) = W$ by Proposition 1.30.

(\Leftarrow) If T is surjective, then $\mathcal{R}(T) = W$. Thus,

$$\text{nullity}(T) = \dim(V) - \text{rank}(T) = \dim(W) - \dim(W) = 0,$$

implying $\mathcal{N}(T) = \{0_V\}$. It follows that T is injective. \square

Lemma 2.17. Let V and W be finite-dimensional vector spaces over a field F . Let $S = \{x_1, x_2, \dots, x_n\}$ be a basis for V and let y_1, y_2, \dots, y_n be vectors in W . Then there exists a unique $T \in \mathcal{L}(V, W)$ with $T(x_i) = y_i$ for each $i \in \{1, \dots, n\}$.

Proof. Let T be the transformation that satisfies

$$T(a_1x_1 + a_2x_2 + \dots + a_nx_n) = a_1y_1 + a_2y_2 + \dots + a_ny_n$$

for any $a_1, a_2, \dots, a_n \in F$. It is obvious that $T(x_i) = y_i$ for each $i \in \{1, \dots, n\}$, and T is linear since

$$\begin{aligned} T\left(c \sum_{i=1}^n a_i x_i + \sum_{i=1}^n b_i x_i\right) &= T\left(\sum_{i=1}^n (ca_i + b_i) x_i\right) \\ &= \sum_{i=1}^n (ca_i + b_i) y_i \\ &= c \sum_{i=1}^n a_i y_i + \sum_{i=1}^n b_i y_i \\ &= cT\left(\sum_{i=1}^n a_i x_i\right) + T\left(\sum_{i=1}^n b_i x_i\right) \end{aligned}$$

holds for any scalars $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n, c \in F$. To see the uniqueness, if $T' \in \mathcal{L}(V, W)$ satisfies $T'(x_i) = y_i$ for each $i \in \{1, \dots, n\}$, then we have

$$\begin{aligned} T'(a_1x_1 + \dots + a_nx_n) &= a_1T'(x_1) + \dots + a_nT'(x_n) \\ &= a_1T(x_1) + \dots + a_nT(x_n) \\ &= T(a_1x_1 + \dots + a_nx_n). \end{aligned}$$

for any $a_1, \dots, a_n \in F$. Thus, $T' = T$. □

Theorem 2.18. Let V and W be finite-dimensional vector spaces over a field F . Then $V \cong W$ if and only if $\dim(V) = \dim(W)$.

Proof. (\Rightarrow) Let T be an isomorphism from V onto W . Since T is invertible, T is bijective. Then we have $\text{rank}(T) = \dim(W)$ since $\mathcal{R}(T) = W$. Furthermore, since T is injective, we have $\text{nullity}(T) = 0$, and it follows that $\text{rank}(T) = \dim(V)$ by rank-nullity theorem (Theorem 2.10). Thus, $\dim(V) = \text{rank}(T) = \dim(W)$.

(\Leftarrow) Suppose that $S = \{x_1, x_2, \dots, x_n\}$ is a basis for V and $R = \{y_1, y_2, \dots, y_n\}$ is a basis for W . Then by Lemma 2.17 there exists $T \in \mathcal{L}(V, W)$ such that $T(x_i) = y_i$ for each $i \in \{1, \dots, n\}$. Since R is a basis for W , for each $y \in W$ there exist scalars $a_1, \dots, a_n \in F$ such that

$$y = \sum_{i=1}^n a_i y_i = \sum_{i=1}^n a_i T(x_i) = T\left(\sum_{i=1}^n a_i x_i\right).$$

It follows that T is surjective, and we can conclude that T is bijective by Lemma 2.16. Thus, T is an isomorphism from V onto W , implying $V \cong W$. □

2.4 Coordinates and Matrix Representations

Definition 2.19. Let V be a finite-dimensional vector space over a field F with $\dim(V) = n$. An **ordered basis** for V is a finite sequence

$$\beta = (x_1, x_2, \dots, x_n)$$

of vectors in V such that the set $S = \{x_1, x_2, \dots, x_n\}$ is a basis for V .

Examples.

- The **standard ordered basis** for F^n is (e_1, \dots, e_n) , where e_i is the n -tuple whose i -th component is 1_F and the other components are all 0_F .
- The **standard ordered basis** for $\mathcal{P}_n(F)$ is (t^0, t^1, \dots, t^n) .

Definition 2.20. Let V be a finite-dimensional vector space over a field F . Let $\beta = (x_1, \dots, x_n)$ be an ordered basis for V . Then we define $\phi_\beta : V \rightarrow F^n$ such that

$$\phi_\beta(x) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad \text{for each } x \in V \text{ with } x = \sum_{i=1}^n a_i x_i,$$

where $a_1, a_2, \dots, a_n \in F$. For each vector x in V , $\phi_\beta(x)$ is called the **coordinate** of x with respect to β , denoted by $[x]_\beta$.

Proposition 2.21. Let $\beta = (x_1, \dots, x_n)$ be an ordered basis for a vector space V over F . Then ϕ_β is an isomorphism from V onto F^n .

Proof. ϕ_β is linear since

$$\begin{aligned} \phi_\beta \left(c \sum_{i=1}^n a_i x_i + \sum_{i=1}^n b_i x_i \right) &= \phi_\beta \left(\sum_{i=1}^n (ca_i + b_i) x_i \right) = \begin{pmatrix} ca_1 + b_1 \\ \vdots \\ ca_n + b_n \end{pmatrix} = c \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \\ &= c \cdot \phi_\beta \left(\sum_{i=1}^n a_i x_i \right) + \phi_\beta \left(\sum_{i=1}^n b_i x_i \right) \end{aligned}$$

holds for any $a_1, \dots, a_n, b_1, \dots, b_n, c \in F$. Also, ϕ_β is invertible since there exists $\phi_\beta^{-1} : F^n \rightarrow V$ with

$$\phi_\beta^{-1} \left(\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \right) = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$$

for any $a_1, a_2, \dots, a_n \in F$. Thus, ϕ_β is an isomorphism. □

Definition 2.22. Let V and W be finite-dimensional vector spaces over a field F . Let

$$\beta = (x_1, \dots, x_n) \quad \text{and} \quad \gamma = (y_1, \dots, y_m)$$

be ordered basis for V and W , respectively. Then we define $\Phi_\beta^\gamma : \mathcal{L}(V, W) \rightarrow F^{m \times n}$ by

$$\Phi_\beta^\gamma(T) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

for each $T \in \mathcal{L}(V, W)$, where

$$\begin{aligned} T(x_1) &= a_{11}y_1 + a_{21}y_2 + \cdots + a_{m1}y_m \\ T(x_2) &= a_{12}y_1 + a_{22}y_2 + \cdots + a_{m2}y_m \\ &\vdots \\ T(x_n) &= a_{1n}y_1 + a_{2n}y_2 + \cdots + a_{mn}y_m \end{aligned}$$

hold. For each linear $T : V \rightarrow W$, the matrix $\Phi_\beta^\gamma(T)$ is called the **matrix representation** of T with respect to β and γ , denoted by $[T]_\beta^\gamma$.

Proposition 2.23. Let $\beta = (x_1, \dots, x_n)$ and $\gamma = (y_1, \dots, y_m)$ be ordered bases for a vector spaces V and W over F , respectively. Then for any $T \in \mathcal{L}(V, W)$, we have

$$\left([T]_\beta^\gamma\right)_{ij} = \left([T(x_j)]_\gamma\right)_i$$

for any $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$.

Proof. Let

$$[T]_\beta^\gamma = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Since $T(x_j) = a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m$, we have

$$[T(x_j)]_\gamma = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Thus,

$$\left([T(x_j)]_\gamma\right)_i = a_{ij}$$

holds, which completes the proof. \square

Theorem 2.24. Let β and γ be ordered bases for a vector spaces V and W over F , respectively. Then Φ_β^γ is an isomorphism from $\mathcal{L}(V, W)$ onto $F^{m \times n}$.

Proof. Let $\beta = (x_1, \dots, x_n)$ and $\gamma = (y_1, \dots, y_m)$. Note that Φ_β^γ is linear since for any $c \in F$ and $T_1, T_2 \in \mathcal{L}(V, W)$, we have

$$\begin{aligned}
\left([cT_1 + T_2]_\beta^\gamma\right)_{ij} &= \left([(cT_1 + T_2)(x_j)]_\gamma\right)_i && \text{(Proposition 2.23)} \\
&= \left([cT_1(x_j) + T_2(x_j)]_\gamma\right)_i \\
&= \left(c[T_1(x_j)]_\gamma + [T_2(x_j)]_\gamma\right)_i && (\phi_\gamma \text{ is linear}) \\
&= c\left([T_1(x_j)]_\gamma\right)_i + \left([T_2(x_j)]_\gamma\right)_i \\
&= c\left([T_1]_\beta^\gamma\right)_{ij} + \left([T_2]_\beta^\gamma\right)_{ij} && \text{(Proposition 2.23)} \\
&= \left(c[T_1]_\beta^\gamma + [T_2]_\beta^\gamma\right)_{ij}
\end{aligned}$$

for any $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$. To prove that Φ_β^γ is invertible, let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

be an arbitrary matrix in $F^{m \times n}$. By Lemma 2.17, there exists a unique linear transformation $T : V \rightarrow W$ such that

$$T(x_j) = \sum_{i=1}^m a_{ij} y_i$$

for each $j \in \{1, \dots, n\}$, and it follows that $[T]_\beta^\gamma = A$. Thus, there exists $(\Phi_\beta^\gamma)^{-1} : F^{m \times n} \rightarrow \mathcal{L}(V, W)$ such that $(\Phi_\beta^\gamma)^{-1}(A) = T$ with $[T]_\beta^\gamma = A$ for each $A \in F^{m \times n}$, which completes the proof. \square

Corollary 2.25. If V and W are finite-dimensional vector spaces over F with $\dim(V) = n$ and $\dim(W) = m$, then $\mathcal{L}(V, W)$ is finite-dimensional with $\dim(\mathcal{L}(V, W)) = mn$.

Proof. Straightforward from Theorem 2.18 and Theorem 2.24. \square

2.5 Matrix Multiplication

Definition 2.26. Let F be a field and let $A \in F^{\ell \times m}$ and $B \in F^{m \times n}$ be matrices. The **product** of A and B , denoted by AB , is a matrix in $F^{\ell \times n}$ that satisfies

$$(AB)_{ik} = \sum_{j=1}^m A_{ij}B_{jk}$$

for $i \in \{1, \dots, \ell\}$ and $k \in \{1, \dots, n\}$.

Proposition 2.27. Let U, V, W be vector spaces over F . If $T_1 : U \rightarrow V$ and $T_2 : V \rightarrow W$ are linear, then so is T_2T_1 .

Proof. For $a \in F$ and $x, y \in U$, we have

$$\begin{aligned} (T_2T_1)(ax + y) &= T_2(T_1(ax + y)) && \text{(composition of } T_1 \text{ and } T_2) \\ &= T_2(aT_1(x) + T_1(y)) && \text{(linearity of } T_1) \\ &= aT_2(T_1(x)) + T_2(T_1(y)) && \text{(linearity of } T_2) \\ &= a(T_2T_1)(x) + (T_2T_1)(y). && \text{(composition of } T_1 \text{ and } T_2) \end{aligned}$$

Thus, T_2T_1 is linear. \square

Theorem 2.28. Let U, V, W be finite-dimensional vector spaces with ordered bases

$$\alpha = (x_1, \dots, x_n), \quad \beta = (y_1, \dots, y_m), \quad \text{and} \quad \gamma = (z_1, \dots, z_\ell),$$

respectively. If $T_1 : U \rightarrow V$ and $T_2 : V \rightarrow W$ are linear, then

$$[T_2T_1]_\alpha^\gamma = [T_2]_\beta^\gamma [T_1]_\alpha^\beta.$$

Proof. Let $A = [T_2]_\beta^\gamma$, $B = [T_1]_\alpha^\beta$ and $C = [T_2T_1]_\alpha^\gamma$. Then

$$T_2(y_j) = \sum_{i=1}^{\ell} A_{ij}z_i, \quad T_1(x_k) = \sum_{j=1}^m B_{jk}y_j, \quad \text{and} \quad (T_2T_1)(x_k) = \sum_{i=1}^{\ell} C_{ik}z_i$$

hold for any $j \in \{1, \dots, m\}$ and $k \in \{1, \dots, n\}$. Since for each $k \in \{1, \dots, n\}$,

$$\begin{aligned} \sum_{i=1}^{\ell} C_{ik}z_i &= (T_2T_1)(x_k) \\ &= T_2(T_1(x_k)) \\ &= T_2\left(\sum_{j=1}^m B_{jk}y_j\right) \\ &= \sum_{j=1}^m B_{jk}T_2(y_j) \\ &= \sum_{j=1}^m B_{jk} \sum_{i=1}^{\ell} A_{ij}z_i \\ &= \sum_{i=1}^{\ell} \left(\sum_{j=1}^m A_{ij}B_{jk}\right) z_i, \end{aligned}$$

we have

$$C_{ik} = \sum_{j=1}^m A_{ij} B_{jk}$$

for each $i \in \{1, \dots, \ell\}$ and $k \in \{1, \dots, n\}$. Thus, $C = AB$. \square

Corollary 2.29. Let V and W be finite-dimensional vector spaces with ordered bases β and γ over a field F , respectively. If $T : V \rightarrow W$ is linear, then

$$[T(x)]_\gamma = [T]_\beta^\gamma [x]_\beta$$

for each $x \in V$.

Proof. Let $\alpha = (1_F)$ be an ordered basis for F . For each $x \in V$, let $\varphi : F \rightarrow V$ be the linear transformation with $\varphi(c) = cx$ for each $c \in F$. By Definition 2.22, we have

$$[\varphi]_\alpha^\beta = [\varphi(1_F)]_\beta \quad \text{and} \quad [T\varphi]_\alpha^\gamma = [(T\varphi)(1_F)]_\gamma.$$

Thus, it follows that

$$\begin{aligned} [T(x)]_\gamma &= [T(\varphi(1_F))]_\gamma \\ &= [(T\varphi)(1_F)]_\gamma \\ &= [T\varphi]_\alpha^\gamma \\ &= [T]_\beta^\gamma [\varphi]_\alpha^\beta && \text{(Theorem 2.28)} \\ &= [T]_\beta^\gamma [\varphi(1_F)]_\beta \\ &= [T]_\beta^\gamma [x]_\beta. \end{aligned} \quad \square$$

2.6 Left-Multiplication Transformations

Definition 2.30. Let $A \in F^{m \times n}$ be a matrix. The **left-multiplication transformation** of A , denoted by L_A , is the transformation from F^n to F^m with

$$L_A(x) = Ax$$

for each $x \in F^n$.

Proposition 2.31. Let α , β and γ be standard ordered bases for F^n , F^m and F^ℓ , respectively. Then the following statements are true.

- (a) L_A is linear for each $A \in F^{m \times n}$.
- (b) $[L_A]_\alpha^\beta = A$ for each $A \in F^{m \times n}$.
- (c) $L_{cA+B} = cL_A + L_B$ for each $c \in F$ and $A, B \in F^{m \times n}$.
- (d) $L_{AB} = L_AL_B$ for each $A \in F^{\ell \times m}$ and $B \in F^{m \times n}$.
- (e) $L_{I_n} = I_{F^n}$.

Proof.

- (a) L_A is linear since for any $c \in F$ and $x, y \in F^n$,

$$\begin{aligned} [L_A(cx + y)]_i &= [A(cx + y)]_i \\ &= \sum_{j=1}^n A_{ij} [cx + y]_j \\ &= \sum_{j=1}^n A_{ij} (cx_j + y_j) \\ &= c \sum_{j=1}^n A_{ij} x_j + \sum_{j=1}^n A_{ij} y_j \\ &= c[Ax]_i + [Ay]_i \\ &= [cAx + Ay]_i \\ &= [cL_A(x) + L_A(y)]_i \end{aligned}$$

holds for each $i \in \{1, \dots, m\}$.

- (b) Let $T \in \mathcal{L}(V, W)$ be the transformation with $[T]_\alpha^\beta = A$. Then we have

$$T(x) = [T(x)]_\beta = [T]_\alpha^\beta [x]_\alpha = Ax$$

for each $x \in F^n$ since α and β are standard ordered bases. Thus, $T = L_A$.

- (c) It is proved by

$$[L_{cA+B}]_\alpha^\beta = cA + B = c[L_A]_\alpha^\beta + [L_B]_\alpha^\beta = [cL_A + L_B]_\alpha^\beta.$$

(d) It is proved by

$$[L_{AB}]_{\alpha}^{\gamma} = AB = [L_A]_{\beta}^{\gamma}[L_B]_{\alpha}^{\beta} = [L_AL_B]_{\alpha}^{\gamma}.$$

(e) Since

$$L_{I_n}(x) = I_n x = x = I_{F^n}(x)$$

holds for each $x \in F^n$, $L_{I_n} = I_{F^n}$. □

Lemma 2.32. Let U, V, W, X be vector spaces. Let

$$T_1, T'_1 \in \mathcal{L}(U, V), \quad T_2, T'_2 \in \mathcal{L}(V, W), \quad \text{and} \quad T_3 \in \mathcal{L}(W, X)$$

be linear transformations and let $c \in F$ be a scalar. Then the following statements are true.

(a) $T_1 I_U = T_1 = I_V T_1$.

(b) $T_2(T_1 + T'_1) = T_2 T_1 + T_2 T'_1$.

(c) $(T_2 + T'_2)T_1 = T_2 T_1 + T'_2 T_1$.

(d) $c(T_2 T_1) = (cT_2)T_1 = T_2(cT_1)$.

(e) $T_3(T_2 T_1) = (T_3 T_2)T_1$.

Proof.

(a) Since

$$(T_1 I_U)(x) = T_1(I_U(x)) = T_1(x) = I_V(T_1(x)) = (I_V T_1)(x)$$

holds for each $x \in U$, we have $T_1 I_U = T_1 = I_V T_1$.

(b) Since

$$\begin{aligned} (T_2(T_1 + T'_1))(x) &= T_2((T_1 + T'_1)(x)) && \text{(composition)} \\ &= T_2(T_1(x) + T'_1(x)) && \text{(addition)} \\ &= T_2(T_1(x)) + T_2(T'_1(x)) && \text{(linearity)} \\ &= (T_2 T_1)(x) + (T_2 T'_1)(x) && \text{(composition)} \\ &= (T_2 T_1 + T_2 T'_1)(x) && \text{(addition)} \end{aligned}$$

holds for each $x \in U$, we have $T_2(T_1 + T'_1) = T_2 T_1 + T_2 T'_1$.

(c) Since

$$\begin{aligned} ((T_2 + T'_2)T_1)(x) &= (T_2 + T'_2)(T_1(x)) && \text{(composition)} \\ &= T_2(T_1(x)) + T'_2(T_1(x)) && \text{(addition)} \\ &= (T_2 T_1)(x) + (T'_2 T_1)(x) && \text{(composition)} \\ &= (T_2 T_1 + T'_2 T_1)(x) && \text{(addition)} \end{aligned}$$

holds for each $x \in U$, we have $(T_2 + T'_2)T_1 = T_2 T_1 + T'_2 T_1$.

(d) Since

$$\begin{aligned} (c(T_2T_1))(x) &= c(T_2T_1)(x) = cT_2(T_1(x)) \\ ((cT_2)T_1)(x) &= (cT_2)(T_1(x)) = cT_2(T_1(x)) \\ (T_2(cT_1))(x) &= T_2(cT_1(x)) = cT_2(T_1(x)) \end{aligned}$$

hold for each $x \in U$, we have $c(T_2T_1) = (cT_2)T_1 = T_2(cT_1)$.

(e) Since

$$\begin{aligned} (T_3(T_2T_1))(x) &= T_3((T_2T_1)(x)) && \text{(composition of } T_3 \text{ and } T_2T_1) \\ &= T_3(T_2(T_1(x))) && \text{(composition of } T_2 \text{ and } T_1) \\ &= (T_3T_2)(T_1(x)) && \text{(composition of } T_3 \text{ and } T_2) \\ &= ((T_3T_2)T_1)(x) && \text{(composition of } T_3T_2 \text{ and } T_1) \end{aligned}$$

holds for each $x \in U$, we have $T_3(T_2T_1) = (T_3T_2)T_1$. \square

Theorem 2.33. Let $A, A' \in F^{k \times \ell}$, $B, B' \in F^{\ell \times m}$ and $C \in F^{m \times n}$ be matrices and let $c \in F$ be a scalar. Then the following statements are true.

- (a) $AI_\ell = A = I_kA$.
- (b) $A(B + B') = AB + AB'$.
- (c) $(A + A')B = AB + A'B$.
- (d) $c(AB) = (cA)B = A(cB)$.
- (e) $A(BC) = (AB)C$.

Proof. Straightforward from Lemma 2.32. \square

2.7 Invertible Matrices

Definition 2.34. A matrix $A \in F^{n \times n}$ is **invertible** if L_A is invertible. If A is invertible, then it has an **inverse**, denoted by A^{-1} , which is the matrix in $F^{n \times n}$ such that

$$L_{A^{-1}} = (L_A)^{-1}.$$

Proposition 2.35. The following statements are true for matrices $A, B \in F^{n \times n}$.

- (a) If A is invertible, then $AA^{-1} = I_n = A^{-1}A$.
- (b) If $AB = I_n = BA$, then A is invertible and $B = A^{-1}$.

Proof.

- (a) We have

$$L_{AA^{-1}} = L_AL_{A^{-1}} = L_A(L_A)^{-1} = I_{F^n} = L_{I_n}$$

and

$$L_{A^{-1}A} = L_{A^{-1}}L_A = (L_A)^{-1}L_A = I_{F^n} = L_{I_n},$$

implying $AA^{-1} = I_n = A^{-1}A$.

- (b) Since we have

$$L_AL_B = L_{AB} = L_{I_n} = I_{F^n}$$

and

$$L_BL_A = L_{BA} = L_{I_n} = I_{F^n},$$

we can conclude that $L_B = (L_A)^{-1}$. It follows that

$$B = BI_n = B(AA^{-1}) = (BA)A^{-1} = I_nA^{-1} = A^{-1}.$$

□