Chapter 1

Vector Spaces

1.1 Fields

Definition 1.1.1. 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) $(a \cdot b) + c = a + (b \cdot c)$ 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 labeld 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.

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

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

Proposition 1.1.2. 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). $\square$

Proposition 1.1.3. 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 1.1.2. $\square$

Proposition 1.1.4. Let F be a field with $a, b \in F$.

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

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

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

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 1.1.2 (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 1.1.2 (c). The other half can be proved similarly.

(c) By applying (b) twice, we have
$$(-a) \cdot (-b) = -(a \cdot (-b)) = -(-(a \cdot b)) = a \cdot b. \qquad \square$$

1.2 Vector Spaces

Definition 1.2.1. 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 elements 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.

Example. A field is a vector space over itself.

Example. $\mathbb{C}$ is a vector space over $\mathbb{R}$.

Example. $\mathbb{R}$ is a vector space over $\mathbb{Q}$.

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

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

Then F^n is a vector space over F.

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

Example. 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.

Example. 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 1.2.2. Let V be a vector space with $x, y, z \in F$.

(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 1.1.2. □

Proposition 1.2.3. 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 1.1.4. □

1.3 Subspaces

Definition 1.3.1. 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 1.3.2. 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 1.2.1 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 1.3.3. 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 1.3.2. □

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 $\mathcal{P}(F)$.

Example. An $n \times n$ matrix A is called **diagonal** if $i \neq j$ impies $A_{ij} = 0_F$ for all $i, j \in \{1, \ldots, n\}$. Then one can verify that the set of $n \times n$ diagonal matrices is a subspace of $F^{n \times n}$.

Example. The **trace** of an $n \times n$ matrix A, denoted by $\operatorname{tr}(A)$, is defined by

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

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 1.3.4. 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 1.3.3. 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 1.3.5. 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 1.3.6. 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 1.3.5 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 1.3.5 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 1.4.1. Let $(G, +)$ be an Abelian group. Then we define

$$\sum_{i=m}^{n} a_i = \begin{cases} \displaystyle\sum_{i=m}^{n-1} a_i + a_n & \text{if } m \le n \\ 0_G & \text{if } m > n, \end{cases}$$

where $a_i \in G$ for each integer i with $m \le i \le n$.

Definition 1.4.2. Let $(V, +, \cdot)$ be a vector space over F. Let S be a subset of V. Then a vector $x \in V$ is called a **linear combination** of S if there exist some nonnegative integer n, scalars $a_1, \ldots, a_n \in F$, and vectors $x_1, \ldots, x_n \in S$ such that

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

Remark. Since n can be zero, 0_V is a linear combination for all $S \subseteq V$.

Remark. Although S can be infinite, the number of terms in the summation must be finite. For example, in the vector space $\mathbb{R}$ over $\mathbb{Q}$, although we have

$$e = \sum_{k=0}^{\infty} \frac{1}{k!} = \frac{1}{1} + \frac{1}{1} + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \cdots,$$

e is still not a linear combination of $\mathbb{Q}$.

Definition 1.4.3. Let $(V, +, \cdot)$ is a vector space over F. The **span** of S, denoted $\operatorname{span}(S)$, is the set that consists of all linear combinations of S.

Theorem 1.4.4. Let $(V, +, \cdot)$ be a vector space over F. Let $S \subseteq V$. Then the following statements are true.

(a) $\operatorname{span}(S)$ is a subspace of V.

(b) If W is a subspace of V such that $S \subseteq W$, then $\operatorname{span}(S) \subseteq W$.

Proof.

(a) If $c \in F$ and $x, y \in \operatorname{span}(S)$, then there exist nonnegative integers m, n, scalars $a_1, \ldots, a_m, b_1, \ldots, b_n \in F$ and vectors $x_1, \ldots, x_m, y_1, \ldots, y_n \in S$ such that

$$x = \sum_{i=1}^{m} a_i x_i \quad \text{and} \quad y = \sum_{j=1}^{n} b_j y_j.$$

Thus, we have

$$\begin{aligned} cx &= c(a_1x_1 + \cdots + a_m x_m) \\ &= c(a_1x_1) + \cdots + c(a_m x_m) \\ &= (ca_1)x_1 + \cdots + (ca_m)x_m \in \operatorname{span}(S) \end{aligned}$$

and

$$x + y = a_1x_1 + \cdots a_m x_m + b_1 y_1 + \cdots + b_n y_n \in \operatorname{span}(S).$$

Also, $0_V \in \operatorname{span}(S)$. Hence, $\operatorname{span}(S)$ is a subspace of V.

(b) If $x \in \operatorname{span}(S)$, then there exists an nonnegative integer n, scalars $a_1, \ldots, a_n \in F$ and vectors $x_1, \ldots, x_n \in S$ such that

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

Thus, since $x_1, \ldots, x_n \in W$, we have $x = a_1 x_1 + \cdots + a_n x_n \in W$. $\square$

Definition 1.4.5. A subset S of a vector space $(V, +, \cdot)$ **spans** V if $\operatorname{span}(S) = V$. In this case, we also say that S is a **spanning set** of V.

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

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

1.5 Linearly Independent Sets

Definition 1.5.1. Let $(V, +, \cdot)$ be a vector space over F. Let S be a subset of V. For scalars $a_1, \ldots, a_n \in F$ and distinct vectors $x_1, \ldots, x_n \in S$, we say that

$$\sum_{i=1}^{n} a_i x_i = 0_V$$

is a **trivial representation** of 0_V as a linear combination of S if $a_1 = \cdots = a_n = 0_F$.

Definition 1.5.2. Let $(V, +, \cdot)$ be a vector space over F.

- A subset S of V is called **linearly dependent** if there exists a nontrivial representation of 0_V as a linear combination of S.
- A subset S of V is called **linearly independent** if it is not linear dependent.

Theorem 1.5.3. Let $(V, +, \cdot)$ be a vector space over F and let $S \subseteq V$. Then S is linearly independent if and only if there exists $x \in S$ such that $x \in \operatorname{span}(S \setminus \{x\})$.

Proof. $(\Rightarrow)$ Because S is linearly dependent, it follows that there exists a nontrivial representation

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

as a linear combination of S, where $a_1, \ldots, a_n \in F$ are scalars and $x_1, \ldots, x_n \in S$ are distinct vectors. Without loss of generality, let $a_1 \neq 0_F$. Then we have

$$\begin{aligned} x_1 &= (-a_1)^{-1}(a_2x_2 + \cdots + a_nx_n) \\ &= (-a_1)^{-1}a_2x_2 + \cdots + (-a_1)^{-1}a_nx_n \\ &\in \operatorname{span}(S \setminus \{x_1\}). \end{aligned}$$

$(\Leftarrow)$ Since $x \in \operatorname{span}(S \setminus \{x\})$, there exists scalars $a_1, \ldots, a_n \in F$ and distinct vectors $x_1, \ldots, x_n \in S \setminus \{x\}$ such that

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

Then

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

is a nontrivial representation of 0_V as a linear combination of S. $\square$

Theorem 1.5.4. Let $(V, +, \cdot)$ be a vector space over F. Let S be a subset of V and let x be an element of S. Then $x \in \operatorname{span}(S \setminus \{x\})$ if and only if $\operatorname{span}(S) = \operatorname{span}(S \setminus \{x\})$.

Proof. $(\Rightarrow)$ Since $x \in \operatorname{span}(S \setminus \{x\})$ and $S \setminus \{x\} \subseteq \operatorname{span}(S \setminus \{x\})$, we have

$$S \subseteq \operatorname{span}(S \setminus \{x\}) \quad \Rightarrow \quad \operatorname{span}(S) \subseteq \operatorname{span}(S \setminus \{x\})$$

by Theorem 1.4.4. Also, $\operatorname{span}(S \setminus \{x\}) \subseteq \operatorname{span}(S)$ because $S \setminus \{x\} \subseteq S$. Thus, we can conclude that $\operatorname{span}(S \setminus \{x\}) = \operatorname{span}(S)$.

$(\Leftarrow)$ Since $x \in S \subseteq \operatorname{span}(S) = \operatorname{span}(S \setminus \{x\})$, we have $x \in \operatorname{span}(S \setminus \{x\})$. $\square$

Example. Let $S = \{1, 1+2x, 1+2x+3x^2, 1+2x+3x^2+4x^3\}$ be a subset of $\mathcal{P}_3(\mathbb{R})$. Then S is linearly independent since the only solution to the following system of linear equations

$$\begin{aligned} a_1 \qquad\qquad\qquad &= 0 \\ a_1 + 2a_2 \qquad\qquad &= 0 \\ a_1 + 2a_2 + 3a_3 \qquad &= 0 \\ a_1 + 2a_2 + 3a_3 + 4a_4 &= 0 \end{aligned}$$

is $a_1 = a_2 = a_3 = a_4 = 0$.

Theorem 1.5.5. Let $(V, +, \cdot)$ be a vector space, and let $R \subseteq S \subseteq V$. If R is linearly dependent, then S is linearly dependent.

Proof. If R is linearly dependent, then there exists $x \in R$ such that $x \in \operatorname{span}(R \setminus \{x\})$. By $R \subseteq S$, we have $R \setminus \{x\} \subseteq S \setminus \{x\}$. Since $x \in S$ and $x \in \operatorname{span}(S \setminus \{x\})$, S is linearly dependent. □

Corollary 1.5.6. Let $(V, +, \cdot)$ be a vector space, and let $R \subseteq S \subseteq V$. If S is linearly independent, then R is linearly independent.

Proof. Suppose that S is linearly independent. If R is linearly dependent, then so is S by Theorem 1.5.5, contradiction. Thus, R is linearly independent. □

Theorem 1.5.7. Let $(V, +, \cdot)$ be a vector space. For each finite set $S \subseteq V$, there exists a linearly independent set $Q \subseteq S$ such that $\operatorname{span}(Q) = \operatorname{span}(S)$.

Proof. The proof is by induction on $n = |S|$. The induction begins with $n = 0$, i.e., $S = \varnothing$. Since $\varnothing$ is linearly independent, we can choose $R = \varnothing$, and thus the theorem holds.

Now suppose that the theorem is true for some integer $n \geq 0$, and we prove that the theorem holds for $n + 1$. If S is linearly independent, then we can choose $Q = S$. Otherwise, there exists $x \in S$ with $\operatorname{span}(S \setminus \{x\}) = \operatorname{span}(S)$ because S is linearly dependent. Let $S' = S \setminus \{x\}$. Then there exists a linearly independent set $Q \subseteq S'$ such that $\operatorname{span}(Q) = \operatorname{span}(S')$ by induction hypothesis, implying $Q \subseteq S$ and $\operatorname{span}(Q) = \operatorname{span}(S)$. □

1.6 Bases and Dimension

Definition 1.6.1. Let $(V, +, \cdot)$ be a vector space. A subset S of V is a **basis** of V if S is not only a spanning set but also a linearly independent set of V.

Example. Following are some examples of bases.

- Since $\operatorname{span}(\varnothing) = \{0_V\}$ and $\varnothing$ is linearly independent, $\varnothing$ is a basis of $\{0_V\}$.
- Let $S = \{x_1, \ldots, x_n\}$ be a subset of F^n with $(x_i)_j = [\![i = j]\!]$ for all $i, j \in \{1, \ldots, n\}$. Then S is called the **standard basis** of F^n.
- The set $S = \{1_F, x, x^2, \ldots, x^n\}$ is the called the **standard basis** of $\mathcal{P}_n(F)$.

Theorem 1.6.2. Let $(V, +, \cdot)$ be a vector space over F. 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. By Theorem 1.5.7, there exists a linearly independent set $Q \subseteq S$ such that $\operatorname{span}(Q) = \operatorname{span}(S) = V$. Thus, Q is a finite basis of V. $\square$

Theorem 1.6.3 (Replacement Theorem)**.** Let $(V, +, \cdot)$ be a vector space over F. Let S be a finite set that spans V, and let $Q \subseteq V$ 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 $\operatorname{span}(Q \cup R) = V$ hold.

Proof. The proof is based on induction on $|Q|$. The induction begins with $|Q| = 0$, i.e., $Q = \varnothing$. Choosing $R = S$, we have $Q \cup R = S$, and thus both $|Q \cup R| = S$ and $\operatorname{span}(Q \cup R) = V$ hold.

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 $Q = \{x_1, \ldots, x_{m+1}\}$ and let $Q' = Q \setminus \{x_{m+1}\}$. By induction hypothesis, there exists $R' = \{y_1, \ldots, y_k\} \subseteq S \setminus Q'$ such that $m + k = |S|$ and $\operatorname{span}(Q' \cup R') = V$. Since $Q' \cup R'$ spans V, there exists $a_1, \ldots, a_m, b_1, \ldots, b_k \in F$ 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 $j \in \{1, \ldots, k\}$, then x_{m+1} is a linear combination of Q, implying that Q is linearly dependent, contradiction. Thus, there must exist some $j \in \{1, \ldots, k\}$ such that $b_j \neq 0_F$. Without loss of generality let $b_k \neq 0_F$. Also, let $R = \{y_1, \ldots, y_{k-1}\}$. Then $|Q \cup R| = (m + 1) + (k - 1) = |S|$. Since $k \geq 1$, we have $|Q| \leq |S|$. Note that $(Q' \cup R') \setminus (Q \cup R) = \{y_k\}$. By

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

we have

$$Q' \cup R' \subseteq Q \cup R \cup \{y_k\} \subseteq \operatorname{span}(Q \cup R).$$

Thus, by Theorem 1.4.4 we have

$$V = \operatorname{span}(Q' \cup R') \subseteq \operatorname{span}(Q \cup R) \subseteq V,$$

implying $\operatorname{span}(Q \cup R) = V$. $\square$

Corollary 1.6.4. Let $(V, +, \cdot)$ be a vector space over F that is spanned by a finite set. Then every linearly independent subset of V is finite.

Proof. Suppose that S is a finite spanning set of V and that Q is linearly independent. If Q is infinite, then there exists $Q' \subseteq Q$ with $|Q'| = |S|+1$. It follows that Q' is linearly independent by Theorem 1.5.5, and thus $|Q'| \leq |S|$ by Theorem 1.6.3, contradiction to $|Q'| = |S| + 1$. Therefore, Q is finite. $\square$

Theorem 1.6.5. Let $(V, +, \cdot)$ be a vector space over F. If V has a finite basis, then all bases of V have the same size.

Proof. Let S be a finite basis of V and let Q be an arbitrary basis of V. Since $V = \text{span}(S)$ and Q is linearly independent, it follows that Q is finite, and thus $|Q| \leq |S|$ by replacement theorem (Theorem 1.6.3).

Also, since $V = \text{span}(Q)$ and S is linearly independent, we have $|S| \leq |Q|$ by replacement theorem (Theorem 1.6.3). Thus, $|Q| = |S|$. $\square$

Definition 1.6.6. A vector space $(V, +, \cdot)$ over F is called **finite-dimensional** if it has a finite basis. A vector space that is not finite-dimensional is called **infinite-dimensional**.

Definition 1.6.7. The number of vectors in each basis of a finite-dimensional vector space V is called the **dimension** of V and is denoted by $\dim(V)$.

Example. We have $\dim(\{0_V\}) = 0$, $\dim(F^n) = n$, and $\dim(\mathcal{P}_n(F)) = n + 1$.

Example. The dimension of a vector space depends on its field of scalars.

- If $V = \mathbb{C}$ is a vector space over $\mathbb{R}$, then $\dim(V) = 2$ since $\{1, i\}$ is a basis of V.
- If $W = \mathbb{C}$ is a vector space over $\mathbb{C}$, then $\dim(W) = 1$ since $\{1\}$ is a basis of W.

Theorem 1.6.8. Let $(V, +, \cdot)$ be a vector space over F. 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 1.6.3), 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., $R = \varnothing$. Thus, $\text{span}(Q) = V$.

($\Leftarrow$) Suppose that S spans V with $|S| = n$. By Theorem 1.6.2, 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 of V. $\square$

Theorem 1.6.9. Let $(V, +, \cdot)$ be a finite-dimensional vector space over F, and let V' be a subspace of V. Then the following statements hold.

(a) $\dim(V') \leq \dim(V)$.

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

Proof. Let S be a basis of V and let S' be a basis of V'.

(a) Since S' is linearly independent and $V = \text{span}(S)$, we have $|S'| \leq |S|$ by replacement theorem (Theorem 1.6.3). Thus, $\dim(V') \leq \dim(V)$.

(b) Since S' is linearly independent and $|S'| = \dim(V)$, we have $\text{span}(S') = V$ by Theorem 1.6.8. Thus, $V' = \text{span}(S') = V$. $\square$

Chapter 2

Linear Transformations

2.1 Linear Transformations, Null Spaces and Ranges

Definition 2.1.1. Let $f : X \to Y$ be a function.

- f is **injective** (i.e., f is an **injection**) if $T(x) = T(x')$ implies $x = x'$ for $x, x' \in X$.
- f is **surjective** (i.e., f is a **surjection**) if for each $y \in Y$, there exists some $x \in X$ with $T(x) = y$.
- f is **bijective** (i.e., f is a **bijection**) if f is injective and surjective.

Remark. If both domain and codomain of a function are vector spaces, then the function is usually said to be a **transformation**. Furthermore, it is said to be an **operator** if its domain and codomain are the same.

Definition 2.1.2. Let V and W be vector spaces over F. A transformation $T : V \to W$ is **linear** if the following statements hold.

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

(b) $T(ax) = aT(x)$ for all $a \in F$ and $x \in V$.

The set of all linear transformations from V to W is denoted by $\mathcal{L}(V, W)$. In the case that $V = W$, we write $\mathcal{L}(V)$ for short.

Example. The **zero transformation** from V to W is the transformation $O_{V,W} : V \to W$ that satisfies $O_{V,W}(x) = 0_W$ for all $x \in V$. It is clear that $O_{V,W} \in \mathcal{L}(V, W)$.

Example. The **identity transformation** on V is the transformation $I_V : V \to V$ that satisfies $I_V(x) = x$ for all $x \in V$. It is clear that $I_V \in \mathcal{L}(V)$.

Example. Recall that $\mathcal{P}(F)$ is the set of polynomials with coefficients in F.

- The differential operator $D : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ with $D(f) = f'$ for $f \in \mathcal{P}(\mathbb{R})$, where f' is the derivative of f, is linear.
- The operator $T : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ such that for $f \in \mathcal{P}(\mathbb{R})$,

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

  for all $x \in \mathbb{R}$, is linear.

Theorem 2.1.3. If V and W are vector spaces over F, then $\mathcal{L}(V, W)$ is also a vector space over F.

Proof. $\mathcal{L}(V, W)$ is a vector space because it is a subspace of $\mathcal{F}(V, W)$, which is proved as follows.

(a) If $T_1, T_2 \in \mathcal{L}(V, W)$, then $T_1 + T_2$ is linear because

$$\begin{aligned}(T_1 + T_2)(x + y) &= T_1(x + y) + T_2(x + y) \\ &= T_1(x) + T_1(y) + T_2(x) + T_2(y) \\ &= T_1(x) + T_2(x) + T_1(y) + T_2(y) \\ &= (T_1 + T_2)(x) + (T_1 + T_2)(y)\end{aligned}$$

and

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

hold for $x, y \in V$ and $c \in F$.

(b) If $T \in \mathcal{L}(V, W)$ and $a \in F$, then aT is linear because

$$\begin{aligned}(aT)(x + y) &= aT(x + y) \\ &= a(T(x) + T(y)) \\ &= aT(x) + aT(y) \\ &= (aT)(x) + (aT)(y)\end{aligned}$$

and

$$(aT)(cx) = aT(cx) = a(cT(x)) = c(aT(x)) = c(aT)(x)$$

hold for $x, y \in V$ and $c \in F$.

(c) It is clear that $O_{V,W} \in \mathcal{L}(V, W)$. □

Theorem 2.1.4. Let V and W be vector spaces over F, and let $T : V \to W$ 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

$$\begin{aligned} T(a_1x_1 + \cdots + a_nx_n + a_{n+1}x_{n+1}) &= T(a_1x_1 + \cdots + a_nx_n) + T(a_{n+1}x_{n+1}) \\ &= a_1T(x_1) + \cdots + a_nT(x_n) + a_{n+1}T(x_{n+1}). \end{aligned}$$

Thus, the statement is true for nonnegative integer n.

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

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

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

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

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

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

Definition 2.1.5. Let V and W be vector spaces over F, and let $T : V \to W$ be linear.

- The **null space** $\mathcal{N}(T)$ of T is the set of vectors $x \in V$ with $T(x) = 0_W$; that is,

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

- The **range** $\mathcal{R}(T)$ of T is the image of V under T; that is,

$$\mathcal{R}(T) = \{T(x) : x \in V\}.$$

Example. Let $D : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ be the differential operator. Then

$$\mathcal{N}(D) = \{a_0 : a_0 \in \mathbb{R}\} \quad \text{and} \quad \mathcal{R}(D) = \mathcal{P}(\mathbb{R}).$$

Theorem 2.1.6. Let V and W be vector spaces over F, and let $T : V \to W$ be linear. Then $\mathcal{N}(T)$ and $\mathcal{R}(T)$ are subspaces of V and W, respectively.

Proof.

(a) Let $x, x' \in \mathcal{N}(T)$ and $a \in F$. Then we have $T(x+x') = T(x)+T(x') = 0_W+0_W = 0_W$, $T(ax) = aT(x) = a0_W = 0_W$ and $T(0_V) = 0_W$. Thus, $\mathcal{N}(T)$ is a subspace of V.

(b) Let $y, y' \in \mathcal{R}(T)$ and $a \in F$. There exist $x, x' \in V$ with $y = T(x)$ and $y' = T(x')$. Then we have $y + y' = T(x) + T(x') = T(x + x')$, $ay = aT(x) = T(ax)$ and $0_W = T(0_V)$. Thus, $\mathcal{R}(T)$ is a subspace of W. □

Definition 2.1.7. Let V and W be vector spaces over F, and let $T : V \to W$ be linear.

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

Theorem 2.1.8 (Rank-nullity Theorem)**.** Let V and W be vector spaces over F, and let $T : V \to W$ be linear. If V is finite-dimensional, then $\text{nullity}(T) + \text{rank}(T) = \dim(V)$.

Proof. Let S be a basis for V and Q a basis for $\mathcal{N}(T)$. By corollary to replacement theorem (Theorem 1.6.3), there is $R \subseteq S \setminus Q$ such that $Q \cup R$ is a basis for V. Since $|R| = |Q \cup R| - |Q| = \dim(V) - \text{nullity}(T)$, the theorem holds if $|R| = \dim(\mathcal{R}(T))$.

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 S is linearly independent. Thus, $|R| = |T(R)|$. We claim that $T(R)$ is a basis for $\mathcal{R}(T)$.

First we prove that $T(R)$ spans $\mathcal{R}(T)$. By Theorem 2.1.4 (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

$$a_1 T(x_1) + \cdots + a_n T(x_n) = 0_W$$

holds for some $a_1, \ldots, a_n \in F$ and some different $x_1, \ldots, x_n \in R$ with $n \geq 1$. Then by Theorem 2.1.4 we have $T(a_1 x_1 + \cdots + a_n x_n) = 0_W$, and thus $a_1 x_1 + \cdots a_n x_n \in \mathcal{N}(T)$. Hence, there exist some $b_1, \ldots, b_m \in F$ and some different $y_1, \ldots, y_m \in Q$ such that

$$a_1 x_1 + \cdots + a_n x_n = b_1 y_1 + \cdots + b_m y_m.$$

That is,

$$a_1 x_1 + \cdots + a_n x_n + (-b_1) y_1 + \cdots + (-b_m) y_m = 0_V.$$

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 $\mathcal{R}(T)$, and we can conclude that $\text{rank}(T) = |T(R)| = |R| = |Q \cup R| - |Q|$, which completes the proof. $\square$

2.2 Invertibility and Isomorphisms

Definition 2.2.1. Let X and Y be sets and let $f : X \to Y$ be a function.

- A function $g : Y \to X$ is a **left inverse** of f if $g \circ f = I_X$. We say that f is **left invertible** if it has a left inverse.
- A function $g : Y \to X$ is a **right inverse** of f if $f \circ g = I_Y$. We say that f is **right invertible** if it has a right inverse.
- A function $g : R \to S$ is an **inverse** of f if it is a left inverse and a right inverse of f. We say that f is **invertible** if it has an inverse.

Proposition 2.2.2. The following statements are true.

(a) A function is left invertible if and only if it is injective.

(b) A function is right invertible if and only if it is surjective.

(c) A function is invertible if and only if it is bijective.

Proof.

(a) ($\Rightarrow$) Suppose that $f : X \to Y$ is left invertible. Let $g : Y \to X$ be an left inverse of f. Then for each $x, x' \in X$ that satisfy $f(x) = f(x')$, we have $x = g(f(x)) = g(f(x')) = x'$.

($\Leftarrow$) Suppose that $f : X \to Y$ is injective. Then there exists a function $g : Y \to X$ such that $g(f(x)) = x$ holds for all $x \in X$, implying g is a left inverse of f.

(b) ($\Rightarrow$) Suppose that $f : X \to Y$ is right invertible. Let $g : Y \to X$ be an right inverse of f. Then $y = f(g(y))$ for all $y \in Y$, and thus f is surjective.

($\Leftarrow$) Suppose that $f : X \to Y$ is surjective. Then there exists a function $g : Y \to X$ such that $f(g(y)) = y$ for all $y \in Y$, implying g is a right inverse of g.

(c) Straightforward from (a) and (b). $\square$

Definition 2.2.3. Let V and W be vector spaces over F.

- A linear transformation $T : V \to W$ is called an **isomorphism** from V onto W if it is invertible.
- We say that V is **isomorphic** to W, denoted by $V \cong W$, if there is an isomorphism from V onto W.

Proposition 2.2.4. Let V and W be vector spaces over F. Then $V \cong W$ if and only if $W \cong V$.

Proof. If $V \cong W$, then there exists $T \in \mathcal{L}(V, W)$ that is invertible. Because T^{-1} is linear and invertible, it is an isomorphism from W onto V, and thus $W \cong V$. The other side can be proved similarly. $\square$

Theorem 2.2.5. Let V and W be finite-dimensional vector spaces over 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, we have $\text{nullity}(T) = 0$. Thus, by rank-nullity theorem (Theorem 2.1.8) we have $\text{rank}(T) = \dim(V)$. Furthermore, we have $\mathcal{R}(T) = W$ since T is bijective by Proposition 2.2.2. Therefore, $\dim(V) = \dim(W)$.

($\Leftarrow$) To be completed. $\square$