MATH 2022
Linear Algebra II

Joe Tran

April 24, 2025

Preface

These are the first edition of these lecture notes for MATH 2022 (Linear Algebra II). Consequently, there may be several typographical errors, missing exposition on necessary background, and more advanced topics for which there will not be time in class to cover. Future iterations of these notes will hopefully be fairly self-contained provided one has the necessary background. If you come across any typos, errors, omissions, or unclear expositions, please feel free to contact me so that I may continually improve these notes.

Contents

Chapter 1

Vector Spaces

Linear algebra is fundamentally concerned with the study of vector spaces and linear transformations between them. A vector space is a mathematical structure formed by a collection of objects, called vectors, which can be added together and multiplied ("scaled") by elements of a field, such as the real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$.

The abstraction of vector spaces allows us to unify and generalize many familiar mathematical objects—such as geometric vectors, sequences, matrices, polynomials, and functions—under a single framework. Despite their varied appearances, these objects all obey a shared set of algebraic rules.

Understanding vector spaces is crucial not only in pure mathematics but also in applied fields such as physics, engineering, computer science, and data science. Many problems in these domains can be reduced to questions about vectors, subspaces, and transformations.

In this chapter, we begin by defining vector spaces formally, exploring key examples and counterexamples, and laying the groundwork for deeper study of linear structure, spanning sets, linear independence, basis, and dimension.

1.1 Vector Spaces

A vector space consists of a nonempty set V and two operations, one called vector addition, and the other called scalar multiplication. In these notes, we will be more focused on vector spaces over $\mathbb{F}$, where $\mathbb{F} = \mathbb{Q}$, $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. Later on, we will be more focused on scalars of real or complex numbers, so we will denote $\mathbb{K}$ to be either $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

Definition 1.1.1. A *vector space* (or *linear space*) over a field $\mathbb{F}$ is a set V equipped with:

- A binary operation $+ : V \times V \to V$ (called *vector addition*).
- An operation $\cdot : \mathbb{F} \times V \to V$ (called *scalar multiplication*)

such that the following axioms hold for all $\vec{u}, \vec{v}, \vec{w} \in V$ and $\alpha, \beta \in \mathbb{F}$:

1. (Associativity of Vector Addition) $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$.
2. (Commutativity of Vector Addition) $\vec{u} + \vec{v} = \vec{v} + \vec{u}$.
3. (Additive Identity) There exists a $\vec{0}_V \in V$ such that for all $\vec{v} \in V$, $\vec{v} + \vec{0}_V = \vec{0}_V + \vec{v} = \vec{v}$.
4. (Additive Inverse) For every $\vec{v} \in V$, there exists a $-\vec{v} \in V$ such that $\vec{v} + (-\vec{v}) = -\vec{v} + \vec{v} = \vec{0}_V$.
5. (Multiplicative Identity) $\vec{v} \in V$, $1 \cdot \vec{v} = \vec{v}$.
6. (Distributivity over Vector Addition) $\alpha \cdot (\vec{u} + \vec{v}) = \alpha \cdot \vec{u} + \alpha \cdot \vec{v}$.
7. (Distributivity over Scalar Addition) $(\alpha + \beta) \cdot \vec{v} = \alpha \cdot \vec{v} + \beta \cdot \vec{v}$.
8. (Compatibility of Scalar Multiplication) $(\alpha\beta) \cdot \vec{v} = \alpha \cdot (\beta\vec{v})$.

Example 1.1.2. For $n \in \mathbb{N}$, we define the n-dimensional Euclidean space $\mathbb{R}^n$ by

$$\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} : x_1, x_2, ..., x_n \in \mathbb{R} \right\}$$

We equip $\mathbb{R}^n$ with the usual vector addition and scalar multiplication defined by

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} \quad \text{and} \quad \alpha \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix}$$

It is elementary to verify that $\mathbb{R}^n$ with the defined operations is a vector space. Note that instead of $\mathbb{R}^n$, we can also have $\mathbb{C}^n$ or $\mathbb{Q}^n$ with the same operations defined.

Example 1.1.3. For $m, n \in \mathbb{N}$, we define the set of all $m \times n$ matrices with entries in $\mathbb{F}$ by

$$\mathcal{M}_{mn}(\mathbb{F}) = \left\{ \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} : x_{ij} \in \mathbb{F} \text{ for all } 1 \leq i \leq m,\ 1 \leq j \leq n \right\}$$

We equip $\mathcal{M}_{mn}(\mathbb{F})$ with the usual matrix addition defined by

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} + \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} = \begin{bmatrix} x_{11}+y_{11} & x_{12}+y_{12} & \cdots & x_{1n}+y_{1n} \\ x_{21}+y_{21} & x_{22}+y_{22} & \cdots & x_{2n}+y_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1}+y_{m1} & x_{m2}+y_{m2} & \cdots & x_{mn}+y_{mn} \end{bmatrix}$$

and scalar multiplication defined by

$$\alpha \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} = \begin{bmatrix} \alpha x_{11} & \alpha x_{12} & \cdots & \alpha x_{1n} \\ \alpha x_{21} & \alpha x_{22} & \cdots & \alpha x_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha x_{m1} & \alpha x_{m2} & \cdots & \alpha x_{mn} \end{bmatrix}$$

It is elementary to verify that $\mathcal{M}_{mn}(\mathbb{F})$ with the defined operations is a vector space over $\mathbb{F}$. Note that in the case where we are working with $n \times n$ matrices over $\mathbb{F}$, we will simply denote it as $\mathcal{M}_n(\mathbb{F})$.

Example 1.1.4. For $n \in \mathbb{N}$, we define the set of all functions on the interval $[0, 1]$ over $\mathbb{F}$ by

$$\mathcal{F}[0,1] = \{f : [0,1] \to \mathbb{F} : f \text{ is a function}\}$$

We equip $\mathcal{F}[0, 1]$ with pointwise addition and scalar multiplication defined by

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \\ (\alpha f)(x) &= \alpha f(x) \end{aligned}$$

It is elementary to verify that $\mathcal{F}[0, 1]$ with the defined operations is a vector space over $\mathbb{F}$.

Example 1.1.5. For $n \in \mathbb{N}$, we define the set of all polynomials over $\mathbb{F}$ of degree at most n on the interval $[0,1]$ by

$$\mathcal{P}_n[0,1] = \left\{ p : [0,1] \to \mathbb{F} : n \in \mathbb{N}, a_0, a_1, ..., a_n \in \mathbb{F}, p(x) = \sum_{k=0}^{n} a_k x^k \right\}$$

We equip $\mathcal{P}_n[0,1]$ with pointwise addition and scalar multiplication as in Example 1.1.4. It is elementary to verify that $\mathcal{P}_n[0,1]$ with the defined operations is a vector space over $\mathbb{F}$.

Example 1.1.6. The set of all $\mathbb{F}$-valued sequences defined by

$$\mathbb{F}^{\mathbb{N}} = \{(x_n)_{n=1}^{\infty} : \text{for all } n \in \mathbb{N},\, x_n \in F\}$$

We equip $\mathbb{F}^{\mathbb{N}}$ with the usual vector addition and scalar multiplication defined by

$$\begin{aligned} (x_n)_{n=1}^{\infty} + (y_n)_{n=1}^{\infty} &= (x_n + y_n)_{n=1}^{\infty} \\ \alpha(x_n)_{n=1}^{\infty} &= (\alpha x_n)_{n=1}^{\infty} \end{aligned}$$

It is elementary to verify that $\mathbb{F}^{\mathbb{N}}$ with the defined operations is a vector space over $\mathbb{F}$.

We now present some examples of non-vector spaces.

Example 1.1.7. Let $V = \{x \in \mathbb{R} : x \geq 0\}$. We claim that V is not a vector space over $\mathbb{R}$. To see this, note that for any $x > 0$, there is no additive inverse $-x \in \mathbb{R}$ such that $x + (-x) = -x + x = 0$. For example, $2 \in V$, but $-2 \notin V$.

Example 1.1.8. Consider the 2-dimensional Euclidean space $\mathbb{R}^2$, but say we equip it with addition defined by

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \boxplus \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and usual scalar multiplication. We claim that $\mathbb{R}^2$ with this form of addition and scalar multiplication, is not a vector space over $\mathbb{R}$. To see this, we show that $\mathbb{R}^2$ does not satisfy the associativity property. Observe that

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \boxplus \left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \boxplus \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \right) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \boxplus \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}$$

On the other hand, we have

$$\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \boxplus \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right) \boxplus \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \boxplus \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} x_1 + z_1 \\ x_2 + z_2 \end{bmatrix}$$

Clearly, the two resulting vectors are different, so $\mathbb{R}^n$ with such addition and scalar multiplication is not a vector space over $\mathbb{R}$.

1.2 Subspaces of a Vector Space

Subspaces form the foundational building blocks within a vector space. Just as subsets of a set inherit structure from the parent set, subspaces inherit the algebraic structure of a vector space. A subspace is a subset that is itself a vector space under the same operations of addition and scalar multiplication.

Understanding subspaces is central to linear algebra because many important constructs—such as the null space and column space of a matrix, the space of solutions to a homogeneous system, and spaces of polynomials or continuous functions—are all subspaces of some ambient vector space. Much of the power of linear algebra lies in analyzing the structure and relationships between these subspaces.

An essential aspect of studying subspaces involves understanding how they are generated. Given a subset of vectors, we are often interested in the smallest subspace that contains them. This leads naturally to the concept of the *span* of a set of vectors: the set of all linear combinations of those vectors. A spanning set provides a way to describe the entirety of a subspace in terms of simpler components.

In this section, we explore the definition and properties of subspaces, techniques to verify whether a given subset is a subspace, important examples and counterexamples, the concept of the span of a set of vectors, and the relationship between spanning sets and subspaces. These ideas form the theoretical foundation for understanding linear dependence, bases, and dimension in later sections.

Definition 1.2.1. Let V be a vector space over a field $\mathbb{F}$. A subset $W \subseteq V$ is called a *subspace* of V if W is itself a vector space under the same operations of vector addition and scalar multiplication defined on V.

To verify that a nonempty subset $W \subseteq V$ is a subspace, we require the following proposition.

Theorem 1.2.2 (Subspace Test). *Let V be a vector space over $\mathbb{F}$ and let $W \subseteq V$. Then W is a subspace of V if and only if*

1. *(Zero element)* $\vec{0}_V \in W$
2. *(Closed under vector addition) If $\vec{u}, \vec{v} \in W$, then $\vec{v} + \vec{w} \in W$.*
3. *(Closed under scalar multiplication) If $\vec{v} \in W$ and $\alpha \in \mathbb{F}$, then $\alpha\vec{v} \in W$.*

Proof. Follows directly from checking the vector space axioms restricted to the subset W. □

Example 1.2.3. If V is a vector space, then $\{\vec{0}_V\}$ and V are subspaces of V.

Example 1.2.4. Let $V = \mathbb{R}^3$ be a vector space over $\mathbb{R}$, and define

$$W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in V : x + y + z = 0 \right\}$$

We claim that W is a subspace of V. To see this, we need to verify the conditions of Theorem 1.2.2.

To see that (1) holds, note that the zero vector $\vec{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ satisfies $x+y+z = 0+0+0 = 0$, so $\vec{0} \in W$.

To see that (2) holds, let $\vec{u} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$ be vectors in W. Then we have $x_1 + y_1 + z_1 = 0$ and $x_2 + y_2 + z_2 = 0$. Then observe that

$$(x_1 + x_2) + (y_1 + y_2) + (z_1 + z_2) = (x_1 + y_1 + z_1) + (x_2 + y_2 + z_2) = 0$$

Therefore, the vector

$$\vec{u} + \vec{v} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{bmatrix}$$

belongs to W.

Finally, to see that (3) holds, let $\vec{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, and let $\alpha \in \mathbb{R}$. Then we have $x + y + z = 0$. Then observe that

$$\alpha x + \alpha y + \alpha z = \alpha(x + y + z) = 0$$

Therefore, the vector

$$\alpha\vec{v} = \begin{bmatrix} \alpha x \\ \alpha y \\ \alpha z \end{bmatrix}$$

belongs to W.

Therefore, by Theorem 1.2.2, we have shown that W is a subspace of V.

Example 1.2.5. Let $V = \mathcal{F}[0,1]$ be a vector space over $\mathbb{F}$, and let $W = \{f \in V : f(0) = 0\}$. We claim that W is a subspace of V. To see this, we need to verify the conditions of Theorem 1.2.2.

To see that (1) holds, it is easy to note that the zero function is a member of W.

To see that (2) holds, let $f, g \in W$. Then $f(0) = 0$ and $g(0) = 0$ and so

$$(f+g)(0) = f(0) + g(0) = 0 + 0 = 0$$

Therefore, $f + g \in W$.

Finally, to see that (3) holds, let $f \in W$ and $\alpha \in \mathbb{F}$. Then $f(0) = 0$ and so

$$(\alpha f)(0) = \alpha f(0) = \alpha \cdot 0 = 0$$

Therefore, $\alpha f \in W$.

Therefore, by Theorem 1.2.2, we have shown that W is a subspace of V.

Example 1.2.6. Let $V = \mathcal{M}_n(\mathbb{R})$, and for $A \in \mathcal{M}_n(\mathbb{R})$, define

$$Z_n(A) = \{B \in \mathcal{M}_n(\mathbb{R}) : AB = BA\}$$

Then $Z_n(A)$ is a subspace of $\mathcal{M}_n(\mathbb{R})$. To see this, we need to verify the conditions of Theorem 1.2.2.

To see that (1) holds, first note that the zero matrix $O_n \in \mathcal{M}_n(\mathbb{R})$ satisfies

$$AO_n = O_nA = O_n$$

so $O_n \in Z_n(A)$.

To see that (2) holds, let $B, C \in Z_n(\mathbb{R})$. Then we have $AB = BA$ and $AC = CA$. Consequently,

$$A(B+C) = AB + AC = BA + CA = (B+C)A$$

which shows that $B + C \in Z_n(\mathbb{R})$.

Finally, to see that (3) holds, let $B \in Z_n(\mathbb{R})$ and $\alpha \in \mathbb{R}$. Then we have $AB = BA$. Consequently,

$$A(\alpha B) = \alpha AB = \alpha BA = (\alpha B)A$$

which shows that $\alpha B \in Z_n(\mathbb{R})$.

Therefore, by Theorem 1.2.2, we have shown that $Z_n(\mathbb{R})$ is a subspace of $\mathcal{M}_n(\mathbb{R})$.

Example 1.2.7. Let $V = \mathcal{C}[0,1]$ be the space of all continuous functions from the unit interval $[0,1]$ to $\mathbb{F}$. It is elementary to verify that $\mathcal{C}[0,1]$ with pointwise addition and scalar multiplication is a vector space over $\mathbb{F}$. Define

$$\mathcal{D}[0,1] = \{f \in \mathcal{C}[0,1] : f \text{ is differentiable on } [0,1]\}$$

We claim that $\mathcal{D}[0,1]$ is a subspace of $\mathcal{C}[0,1]$. To see this, we need to verify the conditions of Theorem 1.2.2.

To see that (1) holds, note that clearly, the zero function $O(x) = 0$ is differentiable on $[0,1]$, so $O \in \mathcal{D}[0,1]$.

To see that (2) holds, let $f, g \in \mathcal{D}[0,1]$, so f and g are differentiable on $[0,1]$, so f', g' exists. Then from calculus, it follows that $f+g$ is differentiable on $[0,1]$ and

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

so $f+g \in \mathcal{D}[0,1]$.

To see that (3) holds, let $f \in \mathcal{D}[0,1]$ and $\alpha \in \mathbb{F}$, so f is differentiable on $[0,1]$, so f' exists. From calculus, again, it follows that αf is differentiable on $[0,1]$, and

$$(\alpha f)' = \alpha f'$$

so $\alpha f \in \mathcal{D}[0,1]$.

Therefore, by Theorem 1.2.2, we have shown that $\mathcal{D}[0,1]$ is a subspace of $\mathcal{C}[0,1]$.

In linear algebra, one of the central goals is to describe vector spaces using simpler, more manageable building blocks. A key concept in this process is the idea of a *spanning set*.

Definition 1.2.8. Let V be a vector space over $\mathbb{F}$ and let $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ be a collection of vectors in V.

1. A vector $\vec{v}$ is said to be a *linear combination* of the vectors $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ if there are coefficients $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}$ such that

$$\vec{v} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$$

2. The set of all linear combinations of $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ is the *span*, and is denoted by

$$\text{span}\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\} = \left\{\sum_{k=1}^{n} \alpha_k\vec{v}_k : \alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}\right\}$$

3. We say that a nonempty set $S \subseteq V$ is a *spanning set* for V if

$$\operatorname{span}(S) = V$$

Example 1.2.9. Let $V = \mathbb{R}^2$, and consider the vectors $\vec{v}_1 = \begin{bmatrix}1\\2\end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix}3\\1\end{bmatrix}$. Then the vector $\vec{v} = \begin{bmatrix}9\\8\end{bmatrix}$ is a linear combination of $\vec{v}_1$ and $\vec{v}2$ because

$$\vec{v} = 3\vec{v}_1 + 2\vec{v}_2 = 3\begin{bmatrix}1\\2\end{bmatrix} + 2\begin{bmatrix}3\\1\end{bmatrix}$$

Example 1.2.10. Let $V = \mathcal{P}_2(\mathbb{R})$ and let $p(x) = 1 + x + 4x^2$. We claim that $p(x)$ belongs to

$$\operatorname{span}\{1 + 2x - x^2, 3 + 5x + 2x^2\}$$

To see this, let $a, b \in \mathbb{R}$ be such that

$$p(x) = a(1 + 2x - x^2) + b(3 + 5x + 2x^2)$$

Then

$$1 + x + 4x^2 = (a + 3b) + (2a + 5b)x + (-a + 2b)x^2$$

Equating coefficients of powers of x, we have

$$\begin{cases} a + 3b = 1 \\ 2a + 5b = 1 \\ -a + 2b = 4 \end{cases}$$

These equations have the solutions $a = -2$ and $b = 1$, so $p(x)$ is indeed in $\operatorname{span}\{1 + 2x - x^2, 3 + 5x + 2x^2\}$.

Example 1.2.11. Let $V = \mathbb{R}^3$ and let

$$S = \left\{ \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \right\} \subseteq \mathbb{R}^3$$

Then $\operatorname{span}(S) = \mathbb{R}^3$ since every vector $\vec{v} = \begin{bmatrix}x\\y\\z\end{bmatrix}$ in $\mathbb{R}^3$ can be written as

$$\vec{v} = x\begin{bmatrix}1\\0\\0\end{bmatrix} + y\begin{bmatrix}0\\1\\0\end{bmatrix} + z\begin{bmatrix}0\\0\\1\end{bmatrix}$$

In this case, S is a spanning set for $\mathbb{R}^3$. In general, this can be extended to n-dimensional. That is, if

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, ..., \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\} \subseteq \mathbb{R}^n$$

Then $\operatorname{span}(S) = \mathbb{R}^n$.

Example 1.2.12. Let $V = \mathcal{M}_2(\mathbb{R})$ and let

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\} \subseteq \mathcal{M}_2(\mathbb{R})$$

Then $\operatorname{span}(S) = \mathcal{M}_2(\mathbb{R})$ since every vector $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in $\mathcal{M}_2(\mathbb{R})$ can be written as

$$A = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

In this case, S is a spanning set for $\mathcal{M}_2(\mathbb{R})$. In general, this can be extended to $\mathcal{M}_n(\mathbb{R})$. In this case, if

$$S = \left\{ \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, ..., \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \right\} \subseteq \mathcal{M}_n(\mathbb{R})$$

Then $\operatorname{span}(S) = \mathcal{M}_n(\mathbb{R})$.

Example 1.2.13. Let $V = \mathcal{P}_2(\mathbb{R})$ and let

$$S = \{1, x, x^2\} \subseteq \mathcal{P}_2(\mathbb{R})$$

Then $\operatorname{span}(S) = \mathcal{P}_2(\mathbb{R})$ since every vector $p(x) = a + bx + cx^2$ in $\mathcal{P}_2(\mathbb{R})$ can be written as

$$p(x) = a \cdot 1 + b \cdot x + c \cdot x^2$$

In this case, S is a spanning set for $\mathcal{P}_2(\mathbb{R})$. In general, this can be extended to $\mathcal{P}_n(\mathbb{R})$. In this case, if

$$S = \{1, x, x^2, ..., x^n\} \subseteq \mathcal{P}_n(\mathbb{R})$$

Then $\operatorname{span}(S) = \mathcal{P}_n(\mathbb{R})$.

We now present a theorem that says that the span of any set of vectors is a subspace, and in fact, it is the smallest subspace containing such vectors.

Theorem 1.2.14. *Let V be a vector space over $\mathbb{F}$ and let $W = \operatorname{span}\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$. Then*

1. *W is a subspace of V containing each of $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$.*
2. *W is the smallest subspace containing these vectors in the sense that any subspace that contains each of $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ must contain W.*

The following illustrates how (2) of Theorem 1.2.14 works.

Example 1.2.15. Let $V = \mathcal{P}_3(\mathbb{R})$ and let

$$S = \{x^2 + x^3, x, 2x^2 + 1, 3\}$$

We claim that $\operatorname{span}(S) = \mathcal{P}_3(\mathbb{R})$. To see this, we need to show that $\operatorname{span}(S) \subseteq \mathcal{P}_3(\mathbb{R})$ and $\mathcal{P}_3(\mathbb{R}) \subseteq \operatorname{span}(S)$.

For the first, note that any linear combination of vectors in S belong to $\mathcal{P}_3(\mathbb{R})$ because if $a, b, c, d \in \mathbb{R}$ are such that

$$p(x) = a(x^2 + x^3) + bx + c(2x^2 + 1) + 3d$$

Then see that

$$\begin{aligned} p(x) &= ax^2 + ax^3 + bx + 2cx^2 + c + 3d \\ &= ax^3 + (a + c)x^2 + bx + (c + 3d) \end{aligned}$$

and we note from Example 1.2.13, $p(x)$ is a linear combination of $\{1, x, x^2, x^3\}$, so we have shown that $\operatorname{span}(S) \subseteq \mathcal{P}_3(\mathbb{R})$.

For the second, let $p(x) = ax^3 + bx^2 + cx + d \in \mathcal{P}_3(\mathbb{R})$ be arbitrary. We need to show that $p(x)$ can be expressed as a linear combination of vectors in S. So we have for some $e, f, g, h \in \mathbb{R}$,

$$e(x^2 + x^3) + fx + g(2x^2 + 1) + 3h = ax^3 + bx^2 + cx + d$$

Using the above, we have on the left side

$$ex^3 + (e + g)x^2 + fx + (g + 3h) = ax^3 + bx^2 + cx + d$$

So comparing the coefficients, we have $a = e$, $b = e + g$, $c = f$, and $d = g + 3h$. This shows that it is possible to express any arbitrary polynomial $p(x) = ax^3 + bx^2 + cx + d$ as a linear combination of vectors in S, so we have $\mathcal{P}_3(\mathbb{R}) \subseteq \operatorname{span}(S)$.

Therefore, since $\operatorname{span}(S) \subseteq \mathcal{P}_3(\mathbb{R})$ and $\mathcal{P}_3(\mathbb{R}) \subseteq \operatorname{span}(S)$, then $\operatorname{span}(S) = \mathcal{P}_3(\mathbb{R})$, as desired.

Example 1.2.16. Let V be a vector space over $\mathbb{F}$, and let $\vec{u}, \vec{v} \in V$. We claim that

$$\text{span}\{\vec{u}, \vec{v}\} = \text{span}\{\vec{u} + 2\vec{v}, \vec{u} - \vec{v}\}$$

To see this we need to show "$\subseteq$" and "$\supseteq$".

To see the latter, it is obvious that $\text{span}\{\vec{u} + 2\vec{v}, \vec{u} - \vec{v}\} \subseteq \text{span}\{\vec{u}, \vec{v}\}$ by Theorem 1.2.14 since both $\vec{u} + 2\vec{v}$ and $\vec{u}, \vec{v}$ lie in $\text{span}\{\vec{u}, \vec{v}\}$.

To see the former, note that

$$\vec{u} = \frac{1}{3}(\vec{u} + 2\vec{v}) + \frac{2}{3}(\vec{u} - \vec{v})$$

and

$$\vec{v} = \frac{1}{3}(\vec{u} + 2\vec{v}) - \frac{1}{3}(\vec{u} - \vec{v})$$

so $\text{span}\{\vec{u}, \vec{v}\} \subseteq \text{span}\{\vec{u} + 2\vec{v}, \vec{u} - \vec{v}\}$ by Theorem 1.2.14, as desired.

1.3 Linear Independence and Bases

In the previous sections, we studied how vector spaces can be constructed from smaller subsets using linear combinations and spanning sets. However, not all spanning sets are equally efficient or informative. Some may contain redundant vectors—vectors that can be written as combinations of others in the set. To identify the most efficient way to describe a vector space, we turn to the concept of *linear independence*.

Definition 1.3.1. Let V be a vector space over $\mathbb{F}$ and let $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\} \subseteq V$. We say that $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ is *linearly independent* if for $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}$ are such that

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n = \vec{0}_V \implies \alpha_k = 0 \text{ for all } k = 1, 2, ..., n$$

If a set of vectors are not linearly independent, then we say that the set is *linearly dependent*.

Example 1.3.2. Let $V = \mathbb{R}^2$, and let $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. We claim that $\{\vec{v}_1, \vec{v}_2\}$ is linearly independent. To see this, assume that for $\alpha_1, \alpha_2 \in \mathbb{R}$ we have

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 = \vec{0}$$

Then

$$\begin{bmatrix} \alpha_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Comparing entries, we see that $\alpha_1 = 0$ and $\alpha_2 = 0$. In general, the set

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\} \subseteq \mathbb{R}^n$$

is linearly independent.

Example 1.3.3. Let $V = \mathcal{M}_2(\mathbb{R})$ and let

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\} \subseteq \mathcal{M}_2(\mathbb{R})$$

We claim that S is a linearly independent subset of $\mathcal{M}_2(\mathbb{R})$. To see this, assume that for $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$, we have that

$$\alpha_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \alpha_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Then

$$\begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

So comparing entries, we see that $\alpha_1 = 0$, $\alpha_2 = 0$, $\alpha_3 = 0$, and $\alpha_4 = 0$. In general, the set

$$S = \left\{ \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \right\} \subseteq \mathcal{M}_n(\mathbb{R})$$

is linearly independent.

Example 1.3.4. Let $V = \mathcal{P}_2(\mathbb{R})$ and let

$$S = \{1, x, x^2\} \subseteq \mathcal{P}_2(\mathbb{R})$$

We claim that S is a linearly independent subset of $\mathcal{P}_2(\mathbb{R})$. To see this, assume that for $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$, we have

$$\alpha_1 \cdot 1 + \alpha_2 \cdot x + \alpha_3 \cdot x^2 = 0$$

Then observe that

$$\alpha_1 + \alpha_2 x + \alpha_3 x^2 = 0 + 0x + 0x^2$$

So comparing coefficients, we see that $\alpha_1 = 0$, $\alpha_2 = 0$, and $\alpha_3 = 0$. In general, the set

$$S = \{1, x, x^2, ..., x^n\} \subseteq \mathcal{P}_n(\mathbb{R})$$

is linearly independent.

Example 1.3.5. Let $V = \mathcal{P}_2(\mathbb{R})$, and let $S = \{1 + x, 3x + x^2, 2 + x - x^2\} \subseteq \mathcal{P}_2(\mathbb{R})$. We claim that S is a linearly independent subset of $\mathcal{P}_2(\mathbb{R})$. To see this, assume that for $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$, we have

$$\alpha_1(1 + x) + \alpha_2(3x + x^2) + \alpha_3(2 + x - x^2) = 0$$

Then rearranging the equation so that

$$(\alpha_2 - \alpha_3)x^2 + (\alpha_1 + 3\alpha_2 + \alpha_3)x + (\alpha_1 + 2\alpha_3) = 0x^2 + 0x + 0$$

Now comparing the coefficients, we have the system of equations given by

$$\begin{cases} \alpha_2 - \alpha_3 = 0 \\ \alpha_1 + 3\alpha_2 + \alpha_3 = 0 \\ \alpha_1 + 2\alpha_3 = 0 \end{cases}$$

One can check that the unique solution to this system is when $\alpha_1 = 0$, $\alpha_2 = 0$, and $\alpha_3 = 0$. So the set S is linearly independent.

Example 1.3.6. Let $V = \mathcal{C}[0, 2\pi]$, and let $S = \{\sin(x), \cos(x)\} \subseteq \mathcal{C}[0, 2\pi]$. We claim that S is linearly independent. To see this, assume that for $\alpha_1, \alpha_2 \in \mathbb{R}$,

$$\alpha_1 \sin(x) + \alpha_2 \cos(x) = 0$$

This must hold for all values of $x \in [0, 2\pi]$. Taking $x = 0$ yields $\alpha_2 = 0$ and taking $x = \frac{\pi}{2}$ yields $\alpha_1 = 0$.

Example 1.3.7. Let $V = \mathbb{R}^2$, and let $S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right\}$. We claim that S is linearly dependent. To see this, assume that for $\alpha_1, \alpha_2 \in \mathbb{R}$, we have

$$\alpha_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Then

$$\begin{bmatrix} \alpha_1 + 2\alpha_2 \\ 2\alpha_1 + 4\alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So comparing entries, we see that $\alpha_1 + 2\alpha_2 = 0$ and $2\alpha_1 + 4\alpha_2 = 0$, but then $\alpha_1 = -2\alpha_2$ and not always zero, so S has to be linearly dependent.

Example 1.3.8. We claim that the set of polynomials of distinct degrees is independent. To see this, let $p_1, p_2, ..., p_n$ be polynomials such that $\deg(p_k) = d_k$ for each k. By relabelling if necessary, we may assume that $d_1 > d_2 > \cdots > d_n$. Suppose that for $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{R}$,

$$\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_n p_n = 0$$

Since $\deg(p_1) = d_1$, let ax^{d_1} be the term in p_1 of highest degree with $a \neq 0$. Since $d_1 > d_2 > \cdots > d_n$, it follows that $\alpha_1 a x^{d_1}$ is the only term of degree d_1 in the linear combination $\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_n p_n = 0$. This implies that $\alpha_1 a x^{d_1} = 0$, so $\alpha_1 = 0$. But then $\alpha_2 p_2 + \alpha_3 p_3 + \cdots + \alpha_n p_n = 0$, so we can repeat the argument to show that $\alpha_2 = 0$, and so on.

A set of vectors is linearly independent if $\vec{0}$ is a linear combination in a unique way. The following proposition shows that every linear combination of these vectors has uniquely determined coefficients.

Proposition 1.3.9. *Let V be a vector space over $\mathbb{F}$ and let $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\} \subseteq V$ be linearly independent. If $\vec{v}$ has two representations*

$$\vec{v} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_n \vec{v}_n$$
$$\vec{v} = \beta_1 \vec{v}_1 + \beta_2 \vec{v}_2 + \cdots + \beta_n \vec{v}_n$$

as a linear combination of these vectors, then for all $k = 1, 2, ..., n$, we have $\alpha_k = \beta_k$.

Proof. Note that by subtracting both equations we yield

$$\vec{0} = (\alpha_1 - \beta_1)\vec{v}_1 + (\alpha_2 - \beta_2)\vec{v}_2 + \cdots + (\alpha_n - \beta_n)\vec{v}_n$$

Then by linear independence, we have $\alpha_k - \beta_k = 0$ for each k, so $\alpha_k = \beta_k$ for each k, as desired. □

The following theorem is one of the most useful results in linear algebra.

Theorem 1.3.10. *Let V be a a vector space over $\mathbb{F}$. Assume that V can be spanned by n vectors. If any set of m vectors in V is linearly independent, then $m \leq n$.*

Proof. Since V can be spanned by n vectors, let $V = \text{span}\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$, and assume that $\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_m\}$ is a linearly independent set in V. Then we can write

$$\vec{u}_1 = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$$

where each $\alpha_k \in \mathbb{R}$. As $\vec{u}_1 \neq \vec{0}_V$, not all of the α_k are zero. Without loss of generality, say that $\alpha_1 \neq 0$, after relabelling the $\vec{v}_k$. Then $V = \text{span}\{\vec{u}_1, \vec{v}_2, ..., \vec{v}_n\}$. Similarly, write

$$\vec{u}_2 = \beta_1\vec{v}_1 + \gamma_2\vec{v}_2 + \cdots + \gamma_n\vec{v}_n$$

Then some $\gamma_k \neq 0$ since $\{\vec{u}_1, \vec{u}_2\}$ is linearly independent, so $V = \text{span}\{\vec{u}_1, \vec{u}_2, ..., \vec{v}_n\}$. Proceeding inductively, if $m > n$, this procedure continues until all the vectors $\vec{v}_k$ are replaced by the vectors $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n$. In particular, $V = \text{span}\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$. But then $\vec{u}_{n+1}$ is a linear combination of $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n$, contrary to the linear independence of the $\vec{u}_k$. Hence, we cannot have $m > n$, so $m \leq n$, as desired. $\square$

We now introduce the concept of a basis of a vector space.

Definition 1.3.11. Let V be a vector space over $\mathbb{F}$. A subset $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\} \subseteq V$ is called a *basis* for V if

1. $\mathcal{B}$ is linearly independent.
2. $V = \text{span}(\mathcal{B})$.

In other words, $\mathcal{B}$ is a basis of V if $\mathcal{B}$ is a linearly independent spanning set.

Example 1.3.12. Let $V = \mathbb{F}^n$. The set

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, ..., \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\} \subseteq \mathbb{R}^n$$

is a basis for $\mathbb{F}^n$. From Example 1.2.11, we have shown that $\mathcal{B}$ is a spanning set for $\mathbb{R}^n$, and in Example 1.3.2, we have shown that $\mathcal{B}$ is linearly independent. We call $\mathcal{B}$ the *standard basis* of $\mathbb{F}^n$.

Example 1.3.13. Let $V = \mathcal{M}_n(\mathbb{F})$. The set

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, ..., \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \right\} \subseteq \mathcal{M}_n(\mathbb{R})$$

is a basis for $\mathcal{M}_n(\mathbb{F})$. From Example 1.2.12, we have shown that $\mathcal{B}$ is a spanning set for $\mathcal{M}_n(\mathbb{F})$, and in Example 1.3.3, we have shown that $\mathcal{B}$ is linearly independent. We call $\mathcal{B}$ the *standard basis* of $\mathcal{M}_n(\mathbb{F})$.

Example 1.3.14. Let $V = \mathcal{P}_n(\mathbb{F})$. The set

$$\mathcal{B} = \{1, x, x^2, ..., x^n\} \subseteq \mathcal{P}_n(\mathbb{R})$$

is a basis for $\mathcal{P}_n(\mathbb{F})$. From Example 1.2.13, we have shown that $\mathcal{B}$ is a spanning set for $\mathcal{P}_n(\mathbb{F})$, and in Example 1.3.4, we have shown that $\mathcal{B}$ is linearly independent. We call $\mathcal{B}$ the *standard basis* of $\mathcal{P}_n(\mathbb{F})$.

Thus, based on the examples above, if $\mathcal{B}$ is a basis for V, then every vector in V can be written as a linear combination of these vectors in a unique way. But even more, any two bases of V must contain the same number of vectors.

Proposition 1.3.15. *Let V be a vector space over $\mathbb{F}$, and suppose that $\mathcal{B}_1$ is a basis for V that has n elements and $\mathcal{B}_2$ is a basis for V that has m elements. Then $n = m$.*

Proof. Since $\mathcal{B}_1$ is a basis for V that contains n elements, and $\mathcal{B}_2$ is a basis for V that contains m elements, it follows from Theorem 1.3.10 that $m \leq n$. Similarly, $n \leq m$, so $n = m$, as asserted. $\square$

Proposition 1.3.15 guarantees that no matter which basis V is chosen it contains the same number of vectors as any other basis. Hence, there is no ambiguity about the following definition.

Definition 1.3.16. Let V be a vector space over $\mathbb{F}$ and let $\mathcal{B}$ be a basis of V. The number of vectors in the basis $\mathcal{B}$ is called the *dimension* of V, and is denoted by

$$\dim(V) = |\mathcal{B}|$$

If $\mathcal{B}$ contains a finite number of elements, then we say that V is a *finite dimensional vector space*. Otherwise, if $\mathcal{B}$ contains an infinite number of elements, then we say that V is an *infinite dimensional vector space.*

Knowing more about infinite dimensional vector spaces would be nice, but it is not the focus of the course, but we mention it for our own understanding that if we have a vector space over $\mathbb{F}$, it is possible that we could have an infinite dimensional vector space. For example, the sequence space $\mathbb{F}^{\mathbb{N}}$ is an example of an infinite dimensional vector space.

Example 1.3.17. If $V = \mathbb{R}^n$, then the set $\mathcal{B}$ in Example 1.3.12 contains n elements, so $\dim(V) = n$.

Example 1.3.18. If $V = \mathcal{M}_n(\mathbb{R})$, then the set $\mathcal{B}$ in Example 1.3.13 contains n^2 elements, so $\dim(V) = n^2$.

Example 1.3.19. If $V = \mathcal{P}_n(\mathbb{R})$, then the set $\mathcal{B}$ in Example 1.3.14 contains $n+1$ elements, so $\dim(V) = n+1$.

The question that one may ask is, how do we construct a basis for a vector space? Proposition 1.3.15 tells us that we can have more than one basis for a vector space, on the condition that we make sure that this new basis contains the same number of elements. The following example illustrates how we can construct a basis for a given vector space.

Example 1.3.20. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and consider the subspace of $\mathcal{M}_2(\mathbb{R})$ given by

$$U(A) = \{B \in \mathcal{M}_2(\mathbb{R}) : AB = BA\} \subseteq \mathcal{M}_2(\mathbb{R})$$

We first need to find a basis for $\mathcal{M}_2(\mathbb{R})$. Start with an arbitrary 2×2 matrix in $U(A)$, so say $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in U(A)$. Then by the given condition that $AB = BA$, we have that

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+c & b+d \\ 0 & 0 \end{bmatrix}$$

and

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & a \\ c & c \end{bmatrix}$$

Thus,

$$\begin{bmatrix} a+c & b+d \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & a \\ c & c \end{bmatrix}$$

Now, this would imply that $c = 0$, $a = b+d$, so

$$\begin{aligned} B &= \begin{bmatrix} b+d & b \\ 0 & d \end{bmatrix} \\ &= \begin{bmatrix} b & b \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} d & 0 \\ 0 & d \end{bmatrix} \\ &= b \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + d \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

Therefore, as $B \in U(A)$ was arbitrary and B is a linear combination of matrices $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, we claim that

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

is a basis for $U(A)$. We have already showed that $\mathcal{B}$ is a spanning set for $U(A)$, so it suffices to show that $\mathcal{B}$ is linearly independent.

Assume that for $\alpha_1, \alpha_2 \in \mathbb{R}$, we have

$$\alpha_1 \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Then

$$\begin{bmatrix} \alpha_1 + \alpha_2 & \alpha_1 \\ 0 & \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Comparing entries, we easily see that $\alpha_1 = 0$ and $\alpha_2 = 0$, so $\mathcal{B}$ is indeed linearly independent. Therefore, $\mathcal{B}$ is indeed a basis of $U(A)$. Since $\mathcal{B}$ contains two elements, then $\dim(U(A)) = 2$.

Example 1.3.21. Let $V = \mathcal{M}_2(\mathbb{R})$, and let

$$W = \{A \in \mathcal{M}_2(\mathbb{R}) : A^T = A\}$$

i.e. W is the set of all 2×2 matrices such that A is symmetric. We want to find a basis for W and find its dimension, so let $A \in W$ be arbitrary, so $A^T = A$. Such matrices A takes the form

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

for $a, b, c \in \mathbb{R}$. Then observe that

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Therefore, A is a linear combination of the matrices $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, and $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Thus, define

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

We claim that $\mathcal{B}$ is a basis for W. Indeed, since any matrix in W is a linear combination of matrices in $\mathcal{B}$, $\mathcal{B}$ is a spanning set for W. To see that $\mathcal{B}$ is linearly independent, assume that for $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$, we have

$$\alpha_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

So then,

$$\begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_2 & \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and we easily see that by comparing the entries, $\alpha_1 = 0$, $\alpha_2 = 0$, and $\alpha_3 = 0$, so $\mathcal{B}$ is indeed linearly independent. Since $\mathcal{B}$ has three elements, $\dim(W) = 3$.

Up until this point, we only had a look at examples involving vector spaces over $\mathbb{R}$. Now let us have a look at some examples that involve vector spaces over $\mathbb{C}$.

Example 1.3.22. Let $V = \mathbb{C}^3$, and define a subspace

$$W = \left\{ \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \in \mathbb{C}^3 : z_1 + iz_2 = 0 \right\}$$

We want to find a basis for W and the dimension of W, so let $\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$ be arbitrary, so $z_1 + iz_2 = 0$, and thus, $z_1 = -iz_2$. Then observe that

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} -iz_2 \\ z_2 \\ z_3 \end{bmatrix} = z_2 \begin{bmatrix} -i \\ 1 \\ 0 \end{bmatrix} + z_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Therefore, $\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix} -i \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Thus, define

$$\mathcal{B} = \left\{ \begin{bmatrix} -i \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

We claim that $\mathcal{B}$ is a basis for W. Indeed, since any vector in W is a linear combination of the vectors in $\mathcal{B}$, $\mathcal{B}$ is a spanning set for W. To see that $\mathcal{B}$ is linearly independent, assume that for $\alpha_1, \alpha_2 \in \mathbb{C}$, we have

$$\alpha_1 \begin{bmatrix} -i \\ 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Then

$$\begin{bmatrix} -i\alpha_1 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus, comparing entries, we easily see that $\alpha_1 = 0$ and $\alpha_2 = 0$, and so $\mathcal{B}$ is linearly independent, as claimed. Since $\mathcal{B}$ has two elements, $\dim(W) = 2$.

Example 1.3.23. Let $V = \mathcal{P}_3(\mathbb{C})$ and consider the subspace

$$U = \{p(x) \in \mathcal{P}_3(\mathbb{C}) : p(1+i) = 0 \text{ and } p(1-i) = 0\}$$

We want to find a basis for U, and the dimension of U. Let $p(x) \in U$ be arbitrary, so $p(1+i) = 0$ and $p(1-i) = 0$. Then $1+i$ and $1-i$ are factors of the polynomial $p(x)$, so there is some polynomial $q(x)$ such that

$$p(x) = (x-(1+i))(x-(1-i))q(x)$$

where $\deg(q(x)) \leq 1$ as $\deg(p(x)) \leq 3$. In this case, we can write $q(x) = ax+b$ for some $a, b \in \mathbb{C}$, and so

$$\begin{aligned} p(x) &= (x^2-2x+2)(ax+b) \\ &= ax^3 - 2ax^2 + 2ax + bx^2 - 2bx + 2b \\ &= a(x^3-2x^2+2x) + b(x^2-2x+2) \end{aligned}$$

Therefore, $p(x)$ is a linear combination of the polynomials x^3-2x^2+2x and x^2-2x+2. Thus, define

$$\mathcal{B} = \{x^3-2x^2+2x, x^2-2x+2\}$$

We claim that $\mathcal{B}$ is a basis for U. Indeed, since any polynomial in U is a linear combination of polynomials in $\mathcal{B}$, then $\mathcal{B}$ is indeed a spanning set for U. To see that $\mathcal{B}$ is linearly independent, assume that for $\alpha_1, \alpha_2 \in \mathbb{C}$, we have

$$\alpha_1(x^3-2x^2+2x) + \alpha_2(x^2-2x+2) = 0$$

Then rearranging

$$\alpha_1 x^3 + (-2\alpha_1+\alpha_2)x^2 + (2\alpha_1 - 2\alpha_2)x + 2\alpha_2 = 0$$

So comparing coefficients, we have

$$\begin{cases} \alpha_1 = 0 \\ -2\alpha_1 + \alpha_2 = 0 \\ 2\alpha_1 - 2\alpha_2 = 0 \\ 2\alpha_2 = 0 \end{cases}$$

Therefore, we easily see that then $\alpha_1 = 0$ and $\alpha_2 = 0$, so $\mathcal{B}$ is indeed linearly independent. Since $\mathcal{B}$ has two elements, $\dim(U) = 2$.

1.4 Finite Dimensional Bases

We have introduced the definition of finite dimensional and infinite dimensional vector spaces in the previous section, but up to this point, we had no guarantee that an arbitrary vector space has a basis, and hence, no guarantee that one can speak at all of the dimension of V. The following theorem shows that any space that is spanned by a finite set of vectors has a finite basis.

Lemma 1.4.1. *Let V be a vector space over $\mathbb{F}$ and let $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ be an independent set of vectors in V. If $\vec{u} \in V \setminus \operatorname{span}\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$, then $\{\vec{u}, \vec{v}_1, ..., \vec{v}_n\}$ is linearly independent.*

Proof. Assume that for $\alpha, \alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}$, we have

$$\alpha\vec{u} + \alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n = \vec{0}_V$$

First, note that $\alpha = 0$ since otherwise,

$$\vec{u} = -\frac{\alpha_1}{\alpha}\vec{v}_1 - \frac{\alpha_2}{\alpha}\vec{v}_2 - \cdots - \frac{\alpha_n}{\alpha}\vec{v}_n \in \operatorname{span}\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$$

which is a contradiction. Therefore, we then have $\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n = \vec{0}_V$, and by assumption, $\alpha_k = 0$ for all $k = 1, 2, ..., n$. □

Remark 1.4.2. Note that the converse of Lemma 1.4.1 is also true, so if $\{\vec{u}, \vec{v}_1, ..., \vec{v}_n\}$ is linearly independent, then $\vec{u} \notin \operatorname{span}\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$.

The question that we may ask ourselves is: if we have a linearly independent set that contains fewer elements than the dimension of a vector space, can we keep adding vectors to this set so that it eventually becomes a basis of the space?

The answer is yes! In finite-dimensional vector spaces, every linearly independent set can be extended to a basis. This powerful and foundational result guarantees that any "partial basis" (a linearly independent set) can be completed into a full basis.

To do so, we require the following lemma.

Lemma 1.4.3. *Let V be a finite-dimensional vector space over $\mathbb{F}$. If W is a subspace of V, then any independent subset of W can be enlarged to a finite basis of W.*

Proof. Suppose that A is a linearly independent subset of W. If $\operatorname{span}(A) = W$, then A is already a basis of W. Otherwise, if $\operatorname{span}(A) \neq W$, then choose $\vec{w}_1 \in W$ such that $\vec{w}_1 \notin \operatorname{span}(A)$. Then the set $A \cup \{\vec{w}_1\}$ is linearly independent by Lemma 1.4.1.

Now if $\operatorname{span}(A \cup \{\vec{w}_1\}) = W$, then we are done. Otherwise, if $\operatorname{span}(A \cup \{\vec{w}_1\}) \neq W$, then choose $\vec{w}_2 \in W$ such that $\vec{w}_2 \notin \operatorname{span}(A \cup \{\vec{w}_1\})$. Then the set $A \cup \{\vec{w}_1, \vec{w}_2\}$ is linearly independent by Lemma 1.4.1.

By proceeding inductively, we claim that a basis of W will be reached eventually. To see this, if no basis of W is ever reached, then the process would create arbitrary large independent sets in V. This is not possible by Theorem 1.3.10 since V is finite-dimensional and so is spanned by a finite set of vectors. $\square$

Theorem 1.4.4. *Let V be a finite-dimensional vector space over $\mathbb{F}$ spanned by $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$. The following hold.*

1. *V has a finite basis and $\dim(V) \leq n$.*
2. *Every independent set of vectors in V can be enlarged to a basis of V by adding vectors from any fixed basis of V.*
3. *If W is a subspace of V, then*
   (a) *W is a finite dimensional subspace and $\dim(W) \leq \dim(V)$.*
   (b) *If $\dim(W) = \dim(V)$, then $W = V$.*

Proof. To see that (1) holds, if $V = \{\vec{0}_V\}$, then V has an empty basis, so $\dim(V) = 0 \leq n$. Otherwise, if $\vec{v} \neq \vec{0}_V$, then $\{\vec{v}\}$ is linearly independent, so (1) follows from Lemma 1.4.3 with $W = V$.

To see that (2) holds, let $\mathcal{B}$ be a basis of V and let A be a linearly independent subset of V. If $\operatorname{span}(A) = V$, then A is a basis for V. Otherwise, $\mathcal{B}$ is not contained in A since $\mathcal{B}$ spans V, so choose $\vec{w}_1 \in \mathcal{B} \setminus \operatorname{span}(A)$ so that

$A \cup \{\vec{w}_1\}$ is linearly independent by Lemma 1.4.1. If $\operatorname{span}(A \cup \{\vec{w}_1\}) = V$, then we are done. Otherwise, a similar argument shows that $A \cup \{\vec{w}_1, \vec{w}_2\}$ is linearly independent for some $\vec{w}_2 \in \mathcal{B}$. Continuing this process, as in the proof of Lemma 1.4.3, a basis of V will be reached eventually.

To see that (3) part (a) holds, note that if $W = \{\vec{0}_V\}$, then this is easy. Otherwise, let $\vec{w} \neq \vec{0}_V \in W$. Then $\{\vec{w}\}$ can be enlarged to a finite basis $\mathcal{B}$ of W by Lemma 1.4.3, so W is finite-dimensional. But $\mathcal{B}$ is also linearly independent, so $\dim(W) \leq \dim(V)$ by Theorem 1.3.10. To see that part (b) holds, if $W = \{\vec{0}_V\}$, then this is trivial since V has a basis. Otherwise, it follows from (2). $\square$

Example 1.4.5. Let $V = \mathcal{M}_2(\mathbb{R})$ and consider the set

$$A = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \right\}$$

Right now, this is not a basis for $\mathcal{M}_2(\mathbb{R})$ just yet since $\dim(\mathcal{M}_2(\mathbb{R})) = 4$, we are missing one element. Recall that the standard basis for $\mathcal{M}_2(\mathbb{R})$ is given by

$$\mathcal{SB}_{\mathcal{M}_2(\mathbb{R})} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

So including one of these in A will produce a basis by Theorem 1.4.4. In fact, *any* one of these matrices in A produces a linearly independent set, and hence, a basis, as we will mention later below.

Example 1.4.6. Let $V = \mathcal{M}_2(\mathbb{C})$ and consider the set

$$A = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & i \\ 0 & 0 \end{bmatrix} \right\}$$

Right now, this is not a basis for $\mathcal{M}_2(\mathbb{C})$ just yet, since $\dim(\mathcal{M}_2(\mathbb{C})) = 4$, we are missing two elements. To make it into a basis, we can simply add two standard basis elements to A to make it into a basis, for example,

$$\mathcal{B} = A \cup \left\{ \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

$\mathcal{B}$ is still a linearly independent set, and moreover, it is now a basis for $\mathcal{M}_2(\mathbb{C})$.

Example 1.4.7. Let $V = \mathcal{P}_3(\mathbb{R})$ and consider the linearly independent set $A = \{1 + x, 1 + x^2\}$. Suppose we want to find a basis containing A. Recall that the standard basis for $\mathcal{P}_3(\mathbb{R})$ is given by

$$\mathcal{SB}_{\mathcal{P}_3(\mathbb{R})} = \{1, x, x^2, x^3\}$$

so including two of these vectors will do. If we use 1 and x^3, then we have $\mathcal{B} = \{1, 1 + x, 1 + x^2, x^3\}$, and this is independent because the polynomials have distinct degrees, and so $\mathcal{B}$ is a basis by Theorem 1.4.4. However, note that if we add $\{1, x\}$ or $\{1, x^2\}$ instead, this would not work!

Example 1.4.8. Let $\mathcal{P}(\mathbb{F}) = \bigcup_{n=1}^{\infty} \mathcal{P}_n(\mathbb{F})$ be the set of all polynomials. It is elementary to verify that $\mathcal{P}(\mathbb{F})$ is a vector space over $\mathbb{F}$. We claim that $\mathcal{P}(\mathbb{F})$ is infinite-dimensional. To see this, note that for each $n \in \mathbb{N}$, $\mathcal{P}(\mathbb{F})$ has a subspace $\mathcal{P}_n(\mathbb{F})$ of dimension $n + 1$.

Suppose for a contradiction that $\mathcal{P}(\mathbb{F})$ is finite dimensional, say $\dim(\mathcal{P}(\mathbb{F})) = m$. Then $\dim(\mathcal{P}_n(\mathbb{F})) \leq \dim(\mathcal{P}(\mathbb{F}))$ by Theorem 1.4.4, so $n + 1 \leq m$. This is absurd since n is arbitrary, so $\mathcal{P}(\mathbb{F})$ must be infinite dimensional.

Proposition 1.4.9. *Let V be a finite-dimensional vector space over $\mathbb{F}$, and let U and W be subspaces of V.*

1. *If $U \subseteq W$, then $\dim(U) \leq \dim(W)$.*
2. *If $U \subseteq W$ and $\dim(U) = \dim(W)$, then $U = W$.*

Proof. Since W is finite-dimensional, (1) follows by taking $V = W$ in part (3) of Theorem 1.4.4. Now, assume that $\dim(U) = \dim(W) = n$, and let $\mathcal{B}$ be a basis for U. Then $\mathcal{B}$ is a linearly independent set in W. If $U \neq W$, then $\operatorname{span}(\mathcal{B}) \neq W$, so $\mathcal{B}$ can be extended to a linearly independent set of $n + 1$ vectors in W, by Lemma 1.4.1. This contradicts Theorem 1.3.10, since W is spanned by $\dim(W) = n$ vectors. Hence, $U = W$, proving (2). □

Example 1.4.10. Let $V = \mathcal{P}_n(\mathbb{R})$, and let

$$W = \{p(x) \in \mathcal{P}_n(\mathbb{R}) : p(1) = 0\}$$

We claim that $\mathcal{B} = \{(x-1), (x-1)^2, ..., (x-1)^n\}$ is a basis for W. To see this, note that $(x - 1), (x - 1)^2, ..., (x - 1)^n$ are all members of W, and that they are linearly independent because they have distinct degrees. Observe then that $\operatorname{span}(\mathcal{B}) \subseteq W \subseteq \mathcal{P}_n(\mathbb{R})$ and $\dim(\operatorname{span}(\mathcal{B})) = n$, and $\dim(\mathcal{P}_n(\mathbb{R})) = n+1$, so $n \leq \dim(W) \leq n + 1$ by Proposition 1.4.9. Since $\dim(W)$ is an integer, we must have $\dim(W) = n$ or $\dim(W) = n + 1$, but then $W = \operatorname{span}(\mathcal{B})$ or $W = \mathcal{P}_n(\mathbb{R})$ by Proposition 1.4.9. Since $W \neq \mathcal{P}_n(\mathbb{R})$, it follows that $W = \operatorname{span}(\mathcal{B})$, as required.

Lemma 1.4.11. *Let V be a vector space over $\mathbb{F}$ and let $A = \{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\} \subseteq V$. Then A is linearly dependent if and only if there exists some vector in A that is a linear combination of the others.*

Proof. Suppose $\vec{v}_2$ is a linear combination of $\vec{v}_1, \vec{v}_3, ..., \vec{v}_n$, so for some $\alpha_1, \alpha_3, ..., \alpha_n \in \mathbb{F}$,

$$\vec{v}_2 = \alpha_1\vec{v}_1 + \alpha_3\vec{v}_3 + \cdots + \alpha_n\vec{v}_N$$

Then

$$\alpha_1\vec{v}_1 + (-1)\vec{v}_2 + \cdots + \alpha_n\vec{v}_n = \vec{0}_V$$

is a nontrivial linear combination that vanishes, so A is linearly dependent.

Conversely, if A is linearly dependent, let

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n = \vec{0}_V$$

where some coefficients is nonzero. If, say $\alpha_2 \neq 0$, then

$$\vec{v}_2 = -\frac{\alpha_1}{\alpha_2}\vec{v}_1 - \frac{\alpha_3}{\alpha_2}\vec{v}_3 - \cdots - \frac{\alpha_n}{\alpha_2}\vec{v}_n$$

is a linear combination of the others. □

Lemma 1.4.1 gives us a way to enlarge linearly independent sets to a basis, Theorem 1.4.12 shows that spanning sets can be cut down to a basis.

Theorem 1.4.12. *Let V be a finite-dimensional vector space over $\mathbb{F}$. Any spanning set for V can be cut down to a basis of V.*

Proof. Since V is finite-dimensional, it has a finite spanning set S. Among all spanning sets contained in S, choose S_0 containing the smallest number of vectors. It suffices to show that S_0 is linearly independent.

Assume for a contradiction that S_0 is linearly dependent. Then by Lemma 1.4.11, there exists a vector $\vec{u} \in S_0$ that is a linear combination of the set $S_1 = S_0 \setminus \{\vec{u}\}$ of vectors in S_0 other than $\vec{u}$. It follows that $\mathrm{span}(S_0) = \mathrm{span}(S_1)$, so $V = \mathrm{span}(S_1)$, but S_1 has fewer elements than S_0, so this contradicts the choice of S_0, so S_0 is linearly independent, as claimed. □

Example 1.4.13. Let $V = \mathcal{P}_3(\mathbb{R})$, and consider

$$S = \{1, x + x^2, 2x - 3x^2, 1 + 3x - 2x^2, x^3\}$$

Right now, we have 5 elements in this set, and we know that $\dim(\mathcal{P}_3(\mathbb{R})) = 4$, so we need to eliminate an element from S. However, we note that we cannot remove x^3, since the span of the rest of S is contained in $\mathcal{P}_2(\mathbb{R})$, but by eliminating $1 + 3x - 2x^2$, we can get a basis, since $1 + 3x - 2x^2$ is the sum of the first three polynomials in S.

Theorem 1.4.4 and 1.4.12 have the following consequence.

Corollary 1.4.14. *Let V be a finite-dimensional vector space over $\mathbb{F}$ with $\dim(V) = n$ and let S be a set of exactly n vectors in V. Then S is linearly independent if and only if S spans V.*

Proof. Assume that S is linearly independent. Then by Theorem 1.4.4, S is contained in a basis $\mathcal{B}$ of V. Hence, $|S| = n = |\mathcal{B}|$ so, as $S \subseteq \mathcal{B}$, it follows that $S = B$, so S spans V.

Conversely, if S spans V, then S contains a basis $\mathcal{B}$ by Theorem 1.4.12. Thus, as $|S| = n = |\mathcal{B}|$, then $S \supseteq \mathcal{B}$, so $S = \mathcal{B}$. Therefore, S is linearly independent. □

Example 1.4.15. Let $V = \mathcal{S}_2(\mathbb{R})$ denote the set of all 2×2 symmetric matrices with real entries and let $GL_2(\mathbb{R})$ denote the subspace of $\mathcal{S}_2(\mathbb{R})$ consisting of invertible matrices with real entries. Suppose we want to find a basis for $GL_2(\mathbb{R})$. From Example 1.3.21, we showed that $\dim(V) = 3$, so what is needed is a set of three invertible matrices that (using Corollary 1.4.14) is either independent or spans V. The set

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}$$

is indeed a linearly independent set, and spans $GL_2(\mathbb{R})$. We call $GL_n(\mathbb{R})$ the space of all general linear matrices with real entries, or all $n \times n$ matrices that are invertible.

Let V be a vector space over $\mathbb{F}$ and let U and W be subspaces of V. There are two subspaces that are of interest, their *sum* $U + W$, and their *intersection* $U \cap W$, defined by

$$U + W = \{\vec{u} + \vec{w} : \vec{u} \in U, \vec{w} \in W\}$$

and

$$U \cap W = \{\vec{v} \in V : \vec{v} \in U \text{ and } \vec{v} \in W\}$$

It is routine to verify that these are indeed subspaces of V, that $U \cap W$ is contained in both U and W, and that $U + W$ contains both U and W.

Theorem 1.4.16. *Let V be a vector space over $\mathbb{F}$, and let U and W be finite dimensional subspaces of V. Then $U + W$ is finite dimensional and*

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$$

Proof. Exercise. □

Remark 1.4.17. Although we can have intersections of vector subspaces, it may not be the case that we have unions of vector subspaces. For example, suppose we have $V = \mathbb{R}^2$, and we let

$$U = \{(x, 0) : x \in \mathbb{R}\} \quad \text{and} \quad W = \{(0, y) : y \in \mathbb{R}\}$$

i.e. U is the set of all points on the x-axis, and W is the set of all points on the y-axis. Their union $U \cup W$ is the set of all points on the x-axis or all points on the y-axis.

To see that $U \cup W$ is not a subspace of $\mathbb{R}^2$, we claim that it is not closed under vector addition. Consider the points $(3, 0), (0, 4) \in U \cup W$. But then the point

$$(3, 4) = (3, 0) + (0, 4) \notin U \cup W$$

Therefore, as $(3, 4) \notin U \cup W$, $U \cup W$ is not closed under vector addition, as claimed.

Chapter 2

Linear Transformations

In the previous chapter, we studied vector spaces, which provide the foundational setting for linear algebra. But vector spaces alone are only part of the story. Equally important is the study of functions between vector spaces that preserve the algebraic structure—these are called linear transformations.

A linear transformation is a rule that maps vectors from one vector space to another while preserving the two basic operations of vector addition and scalar multiplication. That is, the image of a sum is the sum of the images, and the image of a scalar multiple is the scalar multiple of the image.

Understanding linear transformations is essential because they allow us to translate abstract problems about vectors into concrete problems about matrices. Indeed, every linear transformation between finite-dimensional vector spaces can be represented by a matrix, and conversely, every matrix defines a linear transformation.

This chapter builds on the theory of vector spaces to develop a precise and rich understanding of linear transformations. We will define linear transformations and study their basic properties, explore the kernel (null space) and image (range) of a linear transformation, understand the conditions under which a transformation is injective (one-to-one), surjective (onto), or bijective, learn how linear transformations between finite-dimensional spaces correspond to matrix multiplication, introduce change of basis and the idea of similarity between matrices, and analyze the role of dimension through the powerful Dimension Theorem (or Rank-Nullity Theorem).

The study of linear transformations not only deepens our understanding of vector spaces, but it also sets the stage for key ideas in many advanced areas—such as eigenvalues and eigenvectors, diagonalization, and inner product spaces.

Throughout this chapter, we will move fluidly between the abstract world of transformations and the concrete world of matrices, building a bridge between structure and computation.

2.1 Linear Transformations

The most natural functions to consider between vector spaces are those that preserve the algebraic operations of vector addition and scalar multiplication. These functions are called *linear transformations*. They play a central role in linear algebra, as they encode structure-preserving maps between vector spaces.

Definition 2.1.1. Let V and W be vector spaces over $\mathbb{F}$. A *linear transformation* is a function $T : V \to W$ such that

1. (Preserves Vector Addition) For all $\vec{v}_1, \vec{v}_2 \in V$,
$$T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2)$$
2. (Preserves Scalar Multiplication) For all $\vec{v} \in V$ and $\alpha \in \mathbb{F}$,
$$T(\alpha\vec{v}) = \alpha T(\vec{v})$$

A linear transformation $T : V \to V$ is called a *linear operator* on V.

Remark 2.1.2. The two properties of a linear transformation implies that $T(\vec{0}_V) = \vec{0}_W$ and that T preserves linear combinations, that is, if $\vec{v} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$, where $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n \in V$, then

$$T(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n) = \alpha_1 T(\vec{v}_1) + \alpha_2 T(\vec{v}_2) + \cdots + \alpha_n T(\vec{v}_n)$$

Therefore, if one were to show that T is a linear transformation, it is sometimes quicker to check that for any $\vec{v}_1, \vec{v}_2 \in V$ and $\alpha_1, \alpha_2 \in \mathbb{F}$, that

$$T(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2) = \alpha_1 T(\vec{v}_1) + \alpha_2 T(\vec{v}_2)$$

Example 2.1.3. Let $V = W = \mathbb{R}^2$, and let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the map defined by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x + y \\ x - y \end{bmatrix}$$

We claim that T is a linear transformation. To see this, we need to verify the properties of Definition 2.1.1.

To see that T preserves vector addition, let $\vec{v}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ are in $\mathbb{R}^2$, then

$$\begin{aligned}
T(\vec{v}_1 + \vec{v}_2) &= T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) \\
&= T\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right) \\
&= \begin{bmatrix} 2(x_1 + x_2) + (y_1 + y_2) \\ (x_1 + x_2) - (y_1 + y_2) \end{bmatrix} \\
&= \begin{bmatrix} 2x_1 + 2x_2 + y_1 + y_2 \\ x_1 + x_2 - y_1 - y_2 \end{bmatrix} \\
&= \begin{bmatrix} 2x_1 + y_1 \\ x_1 - y_1 \end{bmatrix} + \begin{bmatrix} 2x_2 + y_2 \\ x_2 - y_2 \end{bmatrix} \\
&= T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) \\
&= T(\vec{v}_1) + T(\vec{v}_2)
\end{aligned}$$

So T preserves vector addition.

To see that T preserves scalar multiplication, let $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $\alpha \in \mathbb{R}$. Then

$$\begin{aligned}
T(\alpha\vec{v}) &= T\left(\alpha \begin{bmatrix} x \\ y \end{bmatrix}\right) \\
&= T\left(\begin{bmatrix} \alpha x \\ \alpha y \end{bmatrix}\right) \\
&= \begin{bmatrix} 2(\alpha x) + \alpha y \\ \alpha x - \alpha y \end{bmatrix} \\
&= \begin{bmatrix} \alpha(2x + y) \\ \alpha(x - y) \end{bmatrix} \\
&= \alpha \begin{bmatrix} 2x + y \\ x - y \end{bmatrix} \\
&= \alpha T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) \\
&= \alpha T(\vec{v})
\end{aligned}$$

So T preserves scalar multiplication.

Therefore, by Definition 2.1.1, we have shown that T is a linear transformation.

Example 2.1.4. Let $V = \mathbb{C}^2$, $W = \mathbb{C}$, and let $T : \mathbb{C}^2 \to \mathbb{C}$ be the map defined by

$$T\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) = z_1 - iz_2$$

We claim that T is a linear transformation. To see this, we need to verify the properties of Definition 2.1.1.

To see that T preserves vector addition, let $\vec{v}_1 = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \in \mathbb{C}^2$. Then

$$\begin{aligned} T(\vec{v}_1 + \vec{v}_2) &= T\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}\right) \\ &= T\left(\begin{bmatrix} z_1 + w_1 \\ z_2 + w_2 \end{bmatrix}\right) \\ &= (z_1 + w_1) - i(z_2 + w_2) \\ &= (z_1 - iz_2) + (w_1 - iw_2) \\ &= T\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) + T\left(\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}\right) \\ &= T(\vec{v}_1) + T(\vec{v}_2) \end{aligned}$$

So T preserves vector addition.

To see that T preserves scalar multiplication, let $\vec{v} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$ and let $\alpha \in \mathbb{C}$. Then

$$\begin{aligned} T(\alpha\vec{v}) &= T\left(\alpha \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) \\ &= T\left(\begin{bmatrix} \alpha z_1 \\ \alpha z_2 \end{bmatrix}\right) \\ &= (\alpha z_1) - i(\alpha z_2) \\ &= \alpha(z_1 - iz_2) \\ &= \alpha T\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) \\ &= \alpha T(\vec{v}) \end{aligned}$$

So T preserves scalar multiplication.

Therefore, by Definition 2.1.1, we have shown that T is a linear transformation.

Example 2.1.5. Let $T : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ be defined by $T(p(x)) = p(1)$. We claim that T is a linear transformation. To see this, we need to verify the properties of Definition 2.1.1.

To see that T preserves vector addition, let $p(x), q(x) \in \mathcal{P}_2(\mathbb{R})$ be arbitrary. Then

$$\begin{aligned} T(p(x) + q(x)) &= T((p+q)(x)) \\ &= (p+q)(1) \\ &= p(1) + q(1) \\ &= T(p(x)) + T(q(x)) \end{aligned}$$

So T preserves vector addition.

To see that T preserves scalar multiplication, let $p(x) \in \mathcal{P}_2(\mathbb{R})$ and $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} T(\alpha p(x)) &= T((\alpha p)(x)) \\ &= (\alpha p)(1) \\ &= \alpha p(1) \\ &= \alpha T(p(x)) \end{aligned}$$

So T preserves scalar multiplication.

Therefore, by Definition 2.1.1, we have shown that T is a linear transformation.

Linear transformations arise in countless mathematical settings, from the finite-dimensional spaces of vectors and matrices to the infinite-dimensional worlds of functions and operators. Beyond familiar examples like matrix multiplication and geometric transformations, we encounter linear operators in calculus (e.g., the derivative and integral), in probability (e.g., the expectation operator), and in analysis (e.g., the Laplace and Fourier transforms). These transformations all share the essential properties of linearity, preserving addition and scalar multiplication. The abundance and variety of linear transformations underscore their fundamental role in connecting different areas of mathematics, unifying diverse problems under a common framework. We leave the verification of the following examples to the reader.

Example 2.1.6. Let V be a vector space over $\mathbb{F}$. The *identity operator on V* $\mathrm{id}_V : V \to V$ defined by $\mathrm{id}_V(\vec{v}) = \vec{v}$ is a linear transformation.

Example 2.1.7. Let $V = \mathbb{F}^n$ and define the *left-* and *right-shift operators* by $\mathfrak{s}_l, \mathfrak{s}_r : \mathbb{F}^n \to \mathbb{F}^n$ by

$$\mathfrak{s}_l(x_1, x_2, ..., x_n) = (x_2, x_3, ..., x_n, 0)$$
$$\mathfrak{s}_r(x_1, x_2, ..., x_n) = (0, x_1, x_2, ..., x_{n-1})$$

for all $(x_1, x_2, ..., x_n) \in \mathbb{F}^n$. It is elementary to verify that both $\mathfrak{s}_l$ and $\mathfrak{s}_r$ are linear transformations.

Example 2.1.8. Let $\mathcal{D}[0,1]$ be the space of all differentiable functions on $[0,1]$ (see Example 1.2.7). Define the *differential operator* $D : \mathcal{D}[0,1] \to \mathcal{C}[0,1]$ by

$$D(f) = f'$$

for all $f \in \mathcal{D}[0,1]$. It is elementary to verify that D is a linear transformation.

Example 2.1.9. Let $V = \mathcal{C}[0,1]$. Define the *Volterra operator* $V : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ by

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

for all $f \in \mathcal{C}[0,1]$. Indeed, this linear transformation is well-defined by the Fundamental Theorem of Calculus. It is elementary to verify that V is a linear transformation.

Example 2.1.10. Let $V = \mathcal{C}[0,1]$ and define the *expectation operator* $\mathbb{E} : \mathcal{C}[0,1] \to \mathbb{R}$ by

$$\mathbb{E}(f) = \int_0^1 xf(x)\ dx$$

for all $f \in \mathcal{C}[0,1]$. It is elementary to verify that $\mathbb{E}$ is a linear transformation.

Example 2.1.11. Let $A \in \mathcal{M}_{mn}(\mathbb{R})$ and define the map $T_A : \mathbb{R}^n \to \mathbb{R}^m$ by

$$T_A(\vec{x}) = A\vec{x}$$

for all $\vec{x} \in \mathbb{R}^n$. It is elementary to verify that T_A is a linear transformation.

Example 2.1.12. Let $V = \mathbb{R}^n$, and define the kth *coordinate projection* $\pi_k : \mathbb{R}^n \to \mathbb{R}$ by

$$\pi_k(x_1, x_2, ..., x_n) = x_k$$

for all $(x_1, x_2, ..., x_n) \in \mathbb{R}^n$. It elementary to verify that π_k is a linear transformation.

Example 2.1.13. Let $V = \mathcal{M}_n(\mathbb{R})$. Recall that for $A \in \mathcal{M}_n(\mathbb{R})$ given by

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

The *trace of* A is defined to by

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdot + a_{nn}$$

That is, you take the sum of the diagonal entries of the matrix A. Define the *trace operator* $T : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ by

$$T(A) = \operatorname{tr}(A)$$

for all $A \in \mathcal{M}_n(\mathbb{R})$. It is elementary to verify that T is a linear transformation.

Example 2.1.14. Let $V = \mathcal{C}_p[0, \infty)$ be the space of all piecewise continuous functions on the interval $[0, \infty)$, and define the *Laplace transform* $\mathcal{L} : \mathcal{C}[0, \infty) \to \mathcal{F}[0, \infty)$ by

$$\mathcal{L}(f)(s) = \int_0^\infty e^{-st} f(t)\ dt$$

for all $f \in \mathcal{C}_p[0, \infty)$. It is elementary to verify that $\mathcal{L}$ is a linear transformation.

The following theorem collects three useful properties of all linear transformations. They can be described by saying that, in addition to preserving addition and scalar multiplication, linear transformations preserve the zero vector, negatives, and all linear combinations. Two of which are already mentioned in Remark 2.1.2, but we state as a general statement.

Proposition 2.1.15. *Let V and W be vector spaces over $\mathbb{F}$, and let $T : V \to W$ be a linear transformation. Then*

1. $T(\vec{0}_V) = \vec{0}_W$.
2. *For all* $\vec{v} \in V$, $T(-\vec{v}) = -T(\vec{v})$.
3. *For all* $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n \in V$ *and* $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}$,

$$T(\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_n \vec{v}_n) = \alpha_1 T(\vec{v}_1) + \alpha_2 T(\vec{v}_2) + \cdots + \alpha_n T(\vec{v}_n)$$

As for functions in general, two linear transformations $T, S : V \to W$ are said to be equal, denoted by $T = S$, if they have the same action; that is, $T(\vec{v}) = S(\vec{v})$ for all $\vec{v} \in V$.

Theorem 2.1.16. *Let V, W be vector spaces over $\mathbb{F}$, and let $T, S : V \to W$ be linear transformations. If $V = \operatorname{span}\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ and $T(\vec{v}_k) = S(\vec{v}_k)$ for all k, then $T = S$.*

Proof. Let $\vec{v} \in V$. Then $\vec{v}$ can be written as a linear combination of the vectors $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$, say

$$\vec{v} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$$

Then applying T to both sides, and using the fact that $T(\vec{v}_k) = S(\vec{v}_k)$ for all k, we obtain

$$\begin{aligned} T(\vec{v}) &= T(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n) \\ &= \alpha_1 T(\vec{v}_1) + \alpha_2 T(\vec{v}_2) + \cdots + \alpha_n T(\vec{v}_n) \\ &= \alpha_1 S(\vec{v}_1) + \alpha_2 T(\vec{v}_2) + \cdots + \alpha_n S(\vec{v}_n) \\ &= S(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n) \\ &= S(\vec{v}) \end{aligned}$$

Therefore, as $\vec{v} \in V$ was arbitrary, we have $T = S$, as desired. □

Theorem 2.1.16 can be expressed as follows: If we know what a linear transformation $T : V \to W$ does to each vector in a spanning set for V, then we know what T does to every vector in V. If the spanning set is a basis, we can say much more.

Theorem 2.1.17. *Let V, W be vector spaces over $\mathbb{F}$, and let $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ be a basis for V. Given any vectors $\vec{w}_1, \vec{w}_2, ..., \vec{w}_n \in W$, there exists a unique linear transformation $T : V \to W$ such that $T(\vec{v}_k) = \vec{w}_k$ for each $k = 1, 2, ..., n$. In fact, the action of T is as follows: For $\vec{v} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$ with $\alpha_k \in \mathbb{F}$, then*

$$T(\vec{v}) = T(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n) = \alpha_1\vec{w}_1 + \alpha_2\vec{w}_2 + \cdots + \alpha_n\vec{w}_n$$

Proof. If a linear transformation T does exist with $T(\vec{v}_k) = \vec{w}_k$ for each k, and if S is any other such transformation, then $T(\vec{v}_k) = \vec{w}_k = S(\vec{v}_k)$ holds for each k, so $S = T$. Hence, T is linear if it exists.

We need to show that T is a linear transformation. For any $\vec{v} \in V$, we must specify $T(\vec{v}) \in W$. Since $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ is a basis of V, we have

$$\vec{v} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$$

where $\alpha_1, \alpha_2, ..., \alpha_n$ are uniquely determined by $\vec{v}$. Hence, we may define $T : V \to w$ by

$$T(\vec{v}) = T(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n) = \alpha_1\vec{w}_1 + \alpha_2\vec{w}_2 + \cdots + \alpha_n\vec{w}_n$$

for all $\vec{v} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$. This satisfies $T(\vec{v}_k) = \vec{w}_k$ for each k. The verification that T is linear is left to the reader. $\square$

This theorem shows that a linear transformation can be defined almost at will. Simply specifying where the basis vectors go, and the rest of the action is dictated by the linearity. Theorem 2.1.16 shows that deciding whether two linear transformations are equal comes down to determining whether they have the same effect on the basis vectors.

Example 2.1.18. Suppose we want to find a $T : \mathcal{P}_2(\mathbb{R}) \to \mathcal{M}_2(\mathbb{R})$ be defined such that

$$T(1+x) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, T(x+x^2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, T(1+x^2) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Note that $\{1+x, x+x^2, 1+x^2\}$ is a basis of $\mathcal{P}_2(\mathbb{R})$, so any polynomial $p(x) = a + bx + cx^2$ in $\mathcal{P}_2(\mathbb{R})$ is a linear combination of these vectors. Indeed,

$$p(x) = \frac{1}{2}(a+b-c)(1+x) + \frac{1}{2}(-a+b+c)(x+x^2) + \frac{1}{2}(a-b+c)(1+x^2)$$

Hence, Theorem 2.1.16 gives

$$\begin{aligned} T(p(x)) &= \frac{1}{2}(a+b-c)\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{2}(-a+b+c)\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \frac{1}{2}(a-b+c)\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} a+b-c & -a+b+c \\ -a+b+c & a-b+c \end{bmatrix} \end{aligned}$$

To conclude this section, we will introduce the concept of composition and composite linear operators.

Definition 2.1.19. Given linear transformations $V \xrightarrow{T} W \xrightarrow{S} U$, the *composite* $ST : V \to U$ of T and S is defined by

$$ST(\vec{v}) = S(T(\vec{v}))$$

for all $\vec{v} \in V$. The operation of forming the new function ST is called *composition*.

Sometimes we may see composition denoted as $S \circ T$, but ST will be used more frequently and for simplicity.

The action of ST can be described as follows: ST means T first, then S.

Remark 2.1.20. Not all pairs of linear transformations can be composed. For example, if $T : V \to W$ and $S : W \to U$ are linear transformations, then $ST : V \to U$ is defined, but TS cannot be formed, unless $U = V$ because $S : V \to U$ and $T : V \to W$ do not link in that order.

Moreover, if ST and TS can be formed, they may not be equal. In fact, if $S : \mathbb{R}^m \to \mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ are induced by matrices A and B, respectively, then ST and TS can be formed (they are induced by AB and BA respectively), but the matrix products AB and BA may not be equal (they may not even be the same size).

Example 2.1.21. Let $S : \mathcal{M}_2(\mathbb{R}) \to \mathcal{M}_2(\mathbb{R})$ and $T : \mathcal{M}_2(\mathbb{R}) \to \mathcal{M}_2(\mathbb{R})$ be defined by

$$S\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} c & d \\ a & b \end{bmatrix}, \quad T(A) = A^T$$

for $A \in \mathcal{M}_2(\mathbb{R})$. Here, we describe ST and TS as follows:

$$ST\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = S\left(\begin{bmatrix} a & c \\ b & d \end{bmatrix}\right) = \begin{bmatrix} b & d \\ a & c \end{bmatrix}$$

On the other hand,

$$TS\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = T\left(\begin{bmatrix} c & d \\ a & b \end{bmatrix}\right) = \begin{bmatrix} c & a \\ d & b \end{bmatrix}$$

It is easy to see that TS need not equal to ST, so $TS \neq ST$.

Example 2.1.22. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(\vec{x}) = A\vec{x}$ where $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ (rotation by 90°), and let $S : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $S(\vec{x}) = B\vec{x}$ where $B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ (scaling in y-axis). Then ST and TS are both linear but yield different geometric effects and different matrices.

The following proposition collects some basic properties of the composition operation.

Proposition 2.1.23. *Let $V \xrightarrow{T} W \xrightarrow{S} U \xrightarrow{R} Z$ be linear transformations.*

1. *The composite $ST : V \to U$ is a linear transformation.*

2. $T\,\mathrm{id}_V = T$ *and* $\mathrm{id}_W\, T = T$.

3. $(RS)T = R(ST)$.

Proof. The proofs of (1) and (2) are left for the reader. To see that (3) holds, note that for every $\vec{v} \in V$,

$$[(RS)T](\vec{v}) = (RS)(T(\vec{v})) = R(S(T(\vec{v}))) = R((ST)(\vec{v})) = [R(ST)](\vec{v})$$

which shows that $(RS)T = R(ST)$. $\square$

2.2 The Kernel and Range

When studying a linear transformation, it's not enough to simply understand how it maps inputs to outputs—we also want to analyze what gets sent to zero, and what the image of the transformation looks like. This leads to two fundamental subspaces associated with any linear transformation: the kernel and the range.

The kernel (also called the null space) of a linear transformation captures all vectors that are mapped to the zero vector. In many applications, the kernel tells us about the "degeneracy" or redundancy in the transformation—it tells us which vectors are "lost" or "collapsed" under the mapping.

On the other hand, the range (also called the image) consists of all possible outputs of the transformation. It tells us how much of the codomain is actually "reached" by the linear map. Together, the kernel and range give a precise way to understand the structure and behavior of a transformation.

In this section, we define the kernel and range formally, explore their properties, and compute them through concrete examples in both real and complex vector spaces. These concepts will be essential later on when we study the rank-nullity theorem, invertibility, and the classification of linear maps.

Definition 2.2.1. Let V and W be vector spaces over $\mathbb{F}$, and let $T : V \to W$ be a linear transformation.

1. The *kernel of* T is the set

$$\ker(T) = \{\vec{v} \in V : T(\vec{v}) = \vec{0}_W\}$$

It consists of all vectors in V that are mapped to the zero vector in W. The kernel of T is a subspace of V.

2. The *range of* T is the set

$$\operatorname{ran}(T) = \{T(\vec{v}) : \vec{v} \in V\}$$

It consists of all vectors in W that are images of vectors in V under T. The range of T is a subspace of W.

Example 2.2.2. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y \\ y+z \end{bmatrix}$$

To find the kernel of T, we need to solve

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or

$$\begin{cases} x+y=0 \\ y+z=0 \end{cases}$$

Solving the system gives $x = -y$ and $z = -y$, so

$$\begin{aligned} \ker(T) &= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in V : T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\ &= \left\{ \begin{bmatrix} -y \\ 1 \\ -y \end{bmatrix} : y \in \mathbb{R} \right\} \\ &= \operatorname{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \right\} \end{aligned}$$

To find the range of T, note that

$$\operatorname{ran}(T) = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

So $\ker(T)$ is 1-dimensional, and $\operatorname{ran}(T) = \mathbb{R}^2$.

Example 2.2.3. Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be defined by

$$T\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) = \begin{bmatrix} z_1 + iz_2 \\ z_2 \end{bmatrix}$$

To find the kernel, we need to solve

$$T\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or

$$\begin{cases} z_1 + iz_2 = 0 \\ z_2 = 0 \end{cases}$$

In this case, $z_2 = 0$, so $z_1 = 0$. Therefore, $\ker(T) = \{\vec{0}\}$.

For the range, since both components are independent linear expressions, $\operatorname{ran}(T) = \mathbb{C}^2$.

Example 2.2.4. Let $D : \mathcal{D}[0,1] \to \mathcal{C}[0,1]$ be the differential operator defined by $D(f) = f'$, where $\mathcal{D}[0,1]$ is the space of differentiable functions on $[0,1]$ and $\mathcal{C}[0,1]$ is the space of continuous functions on $[0,1]$.

To determine the kernel, we must find all functions $f \in \mathcal{D}[0,1]$ such that $D(f) = f' = 0$. From calculus, if $f'(x) = 0$ for all $x \in [0,1]$, then f must be a constant function. Thus, for some $c \in \mathbb{R}$, we have $f(x) = c$ for all $x \in [0,1]$. Therefore,

$$\ker(D) = \{f \in \mathcal{D}[0,1] : f(x) = c \text{ for some } c \in \mathbb{R}\} = \operatorname{span}\{1\}$$

To determine the range, observe that for any $g \in \mathcal{C}[0,1]$, there exists a function $f \in \mathcal{D}[0,1]$ such that $f'(x) = g(x)$ for all $x \in [0,1]$. Indeed, by the Fundamental Theorem of Calculus, the function

$$f(x) = \int_0^x g(t)\,dt$$

satisfies $f' = g$. Therefore, every continuous function is the derivative of some differentiable function. Hence,

$$\operatorname{ran}(D) = \mathcal{C}[0,1]$$

Definition 2.2.5. Let V and W be vector spaces over $\mathbb{F}$ and let $T : V \to W$ be a linear transformation.

1. The *nullity of* T is the dimension of the kernel of T, so $\dim(\ker(T))$.

2. The *rank of* T is the dimension of the range of T, so $\dim(\operatorname{ran}(T))$.

Often, a useful way to study a subspace of a vector space is to exhibit it as the kernel or range of a linear transformation. Here is an example.

Example 2.2.6. Let $T : \mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R})$ be the linear transformation defined by $T(A) = A - A^T$. Indeed, it is elementary to verify that T is linear. To find the kernel, note that a matrix A lies in $\ker(T)$ when

$$0 = T(A) = A - A^T$$

so $A = A^T$. In other words, the kernel of T is the set of all symmetric matrices, so

$$\ker(T) = \mathcal{S}_n(\mathbb{R})$$

To find the range, note that $\operatorname{ran}(T)$ consists of all matrices $T(A)$ for $A \in \mathcal{M}_n(\mathbb{R})$. Every such matrix is of the form of a skew-symmetric matrix because

$$T(A)^T = (A - A^T)^T = A^T - A = -T(A)$$

On the other hand, if S is skew-symmetric, so $S^T = -S$, then S lies in $\operatorname{ran}(T)$, so

$$T\left(\frac{1}{2}S\right) = \frac{1}{2}S - \left(\frac{1}{2}S\right)^T = \frac{1}{2}(S - S^T) = \frac{1}{2}(S + S) = S$$

In this case,

$$\operatorname{ran}(T) = \mathcal{S}_n^s(\mathbb{R})$$

where $\mathcal{S}_n^s(\mathbb{R})$ denotes the space of all skew-symmetric matrices.

In the context of linear transformations, it is natural to ask when a transformation behaves like a "function with an inverse." To answer this, we study two fundamental properties of linear transformations: being *one-to-one* (injective) and being *onto* (surjective).

Definition 2.2.7. Let V and W be vector spaces over $\mathbb{F}$ and let $T : V \to W$ be a linear transformation.

1. T is said to be *one-to-one* (*injective*) if $T(\vec{v}_1) = T(\vec{v}_2)$ implies $\vec{v}_1 = \vec{v}_2$.

2. T is said to be *onto* (*surjective*) if $\operatorname{ran}(T) = W$; that is, for any $\vec{w} \in W$, there exists a $\vec{v} \in V$ such that $T(\vec{v}) = \vec{w}$.

3. T is said to be a *bijection* (*bijective*) if T is both injective and surjective.

Clearly, the onto transformations T are those for which $\operatorname{ran}(T) = W$ is a large subspace of W as possible. By contrast, the following shows that one-to-one transformations are the ones with the kernel as small a subspace of V as possible.

Theorem 2.2.8. *Let V and W be vector spaces over $\mathbb{F}$, and let $T : V \to W$ be a linear transformation. Then T is one-to-one if and only if* $\ker(T) = \{\vec{0}_V\}$.

Proof. If T is one-to-one, let $\vec{v}$ be any vector in $\ker(T)$. Then $T(\vec{v}) = \vec{0}_W$, so $T(\vec{v}) = T(\vec{0}_V)$. Hence, $\vec{v} = \vec{0}_V$, because T is one-to-one, os $\ker(T) = \{\vec{0}_V\}$.

Conversely, if $\ker(T) = \{\vec{0}_V\}$ and $T(\vec{v}_1) = T(\vec{v}_2)$ where $\vec{v}_1, \vec{v}_2 \in V$, then

$$T(\vec{v}_1 - \vec{v}_2) = T(\vec{v}_1) - T(\vec{v}_2) = \vec{0}_W$$

$\vec{v}_1 - \vec{v}_2 \in \ker(T) = \{\vec{0}_V\}$, so $\vec{v}_1 - \vec{v}_2 = \vec{0}_V$, and thus, $\vec{v}_1 = \vec{v}_2$, showing that T is one-to-one, as desired. □

Example 2.2.9. For any vector space V, the identity operator $\mathrm{id}_V : V \to V$ is a bijection.

Example 2.2.10. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformations given by

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x-y \end{bmatrix}$$

We claim that T is surjective, but not injective.

To see this, note that the vector $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ lies in $\ker(T)$, so T cannot be injective. On the other hand, for any $\begin{bmatrix} s \\ t \end{bmatrix} \in \mathbb{R}^2$ it lies in $\operatorname{ran}(T)$ because

$$\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} x+y \\ x-y \end{bmatrix}$$

whenever $x = \frac{1}{2}(s+t)$ and $y = \frac{1}{2}(s-t)$ and $z = 0$. Therefore, z is onto.

Example 2.2.11. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation defined by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x-y \\ x \end{bmatrix}$$

We claim that T is injective, but not surjective.

To see this, note that T is injective because

$$\ker(T) = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x + y = x - y = x = 0 \right\} = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

However, note that T is not onto, because the vector $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ does not lie in $\operatorname{ran}(T)$ because if

$$\begin{bmatrix} x + y \\ x - y \\ x \end{bmatrix}$$

for some $x, y \in \mathbb{R}$, then $x + y = 0 = x - y$ and $x = 1$, which is absurd.

Example 2.2.12. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $T(\vec{v}) = A\vec{v}$, where

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

Since A is an invertible matrix (for example, check that $\det(A) \neq 0$), then T is both one-to-one and onto. Therefore, T is a bijection.

Example 2.2.13. Let $D : \mathcal{D}[0,1] \to \mathcal{C}[0,1]$ be the differential operator by $D(f) = f'$. As we have discussed earlier in Example 2.2.4, $\ker(D) = \operatorname{span}\{1\}$, so D cannot be injective. However since $\operatorname{ran}(D) = \mathcal{C}[0,1]$, it follows that D is onto.

Example 2.2.14. Let $V : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ be the Volterra operator

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

Here, $\ker(T) = \{0\}$, since if $V(f(x)) = 0$ for all x, then $f = 0$, so V is injective. On the other hand, V is not nto, because not all continuous functions can be written in the above form.

2.3 The Dimension Theorem

One of the most elegant and powerful results in linear algebra is the *Dimension Theorem*, also known as the *Rank-Nullity Theorem*. This theorem describes a fundamental relationship between three key components of any linear transformation: the dimension of the domain, the dimension of the kernel (null space), and the dimension of the range (image).

Recall that the kernel of a linear transformation measures the extent to which the transformation *fails to be injective*, while the range tells us how much of the codomain is *actually reached* by the transformation. The Dimension Theorem tells us that if we know the dimension of the domain and the size of the kernel, then the size of the image is completely determined—and vice versa.

In essence, this result formalizes the intuitive idea that every vector in the domain contributes either to the kernel or to the image, but not both. This has profound implications across mathematics, including solutions of systems of equations, properties of matrices, and the study of invertibility and isomorphisms.

In this section, we present a formal statement and proof of the theorem, and work through illustrative examples to reinforce its importance.

Theorem 2.3.1 (The Dimension Theorem). *Let V and W be vector spaces over $\mathbb{F}$, and let $T : V \to W$ be a linear transformation. If* $\ker(T)$ *and* $\operatorname{ran}(T)$ *are finite-dimensional, then V is also finite-dimensional, and*

$$\dim(V) = \dim(\ker(T)) + \dim(\operatorname{ran}(T))$$

Proof. First note that every vector in $\operatorname{ran}(T)$ has the form $T(\vec{v})$ for some $\vec{v} \in V$. Hence, let $\{T(\vec{e}_1), T(\vec{e}_2), ..., T(\vec{e}_n)\}$ be a basis for $\operatorname{ran}(T)$, where the $\vec{e}_k$ lie in V. Let $\{\vec{f}_1, \vec{f}_2, ..., \vec{f}_m\}$ be any basis for $\ker(T)$. Then $\dim(\operatorname{ran}(T)) = n$ and $\dim(\ker(T)) = m$, so it suffices to show that

$$\mathcal{B} = \{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n, \vec{f}_1, \vec{f}_2, ..., \vec{f}_m\}$$

is a basis of V.

To see that $\mathcal{B}$ spans V, first note that if $\vec{v} \in V$, then $T(\vec{v}) \in \operatorname{ran}(T)$, so

$$T(\vec{v}) = \alpha_1 T(\vec{e}_1) + \alpha_2 T(\vec{e}_2) + \cdots + \alpha_n T(\vec{e}_n)$$

where $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}$. Then this implies that

$$\vec{v} - \alpha_1\vec{e}_1 - \alpha_2\vec{e}_2 - \cdots - \alpha_n\vec{e}_n$$

lies in $\ker(T)$ and so it is a linear combination of $\vec{f}_1, \vec{f}_2, ..., \vec{f}_m$. Therefore, $\vec{v}$ is a linear combination of the vectors in $\mathcal{B}$.

To see that $\mathcal{B}$ is linearly independent, assume that $\alpha_1, \alpha_2, ..., \alpha_n, \beta_1, \beta_2, ..., \beta_m \in \mathbb{F}$ are such that

$$\alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \cdots + \alpha_n\vec{e}_n + \beta_1\vec{f}_1 + \beta_2\vec{f}_2 + \cdots + \beta_m\vec{f}_m = \vec{0}_V$$

Applying T gives

$$\alpha_1 T(\vec{e}_1) + \alpha_2 T(\vec{e}_2) + \cdots + \alpha_n T(\vec{e}_n) = \vec{0}_W$$

Hence, the linear independence of $\{T(\vec{e}_1), T(\vec{e}_2), ..., T(\vec{e}_n)\}$ gives $\alpha_1 = 0$, $\alpha_2 = 0$,..., $\alpha_n = 0$. But then

$$\beta_1\vec{f}_1 + \beta_2\vec{f}_2 + \cdots + \beta_m\vec{f}_m = \vec{0}_V$$

But by linear independence of $\{\vec{f}_1, \vec{f}_2, ..., \vec{f}_m\}$, we have $\beta_1 = 0$, $\beta_2 = 0$, ..., $\beta_m = 0$. Therefore, we have shown that $\mathcal{B}$ is linearly independent. $\square$

Remark 2.3.2. Note that the vector space V is not assumed to be finite-dimensional in Theorem 2.3.1. In fact, verifying that $\ker(T)$ and $\operatorname{ran}(T)$ are both finite-dimensional is often an important way to prove that V is finite dimensional.

Corollary 2.3.3. *Let V and W be vector spaces over $\mathbb{F}$, and let $T : V \to W$ be a linear transformation. Let $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_r, \vec{e}_{r+1}, ..., \vec{e}_n\}$ be a basis for V such that $\{\vec{e}_{r+1}, \vec{e}_{r+2}, ..., \vec{e}_n\}$ is a basis for $\ker(T)$. Then $\{T(\vec{e}_1), T(\vec{e}_2), ..., T(\vec{e}_r)\}$ is a basis for $\operatorname{ran}(T)$, and hence, $r = \operatorname{rank}(T)$.*

Example 2.3.4. Let $A \in \mathcal{M}_{mn}(\mathbb{R})$ with $\operatorname{rank}(A) = r$. We claim that the space $\operatorname{null}(A)$ of all solutions of the system $A\vec{v} = \vec{0}$ of m homogeneous equations in n variables has dimension $n - r$. Indeed, note that the space in question is simply $\ker(T_A)$, where $T_A(\vec{v}) = A\vec{v}$ for all columns $\vec{v} \in \mathbb{R}^n$. But $\dim(\operatorname{ran}(T_A)) = \operatorname{rank}(T_A) = \operatorname{rank}(A) = r$, so $\dim(\ker(T_A)) = n - r$ by the Dimension Theorem.

For example, let $T : \mathbb{R}^4 \to \mathbb{R}^3$ be defined by the matrix

$$A = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Then $T(\vec{x}) = A\vec{x}$ defines a linear transformation. To apply the Dimension Theorem, we compute the rank and nullity of A.

From the matrix above, we see that the rows are already in row-echelon form, with two leading ones. Therefore, the rank of A is

$$\dim(\operatorname{ran}(T)) = \operatorname{rank}(A) = 2$$

Since the domain of T is $\mathbb{R}^4$, the Dimension Theorem tells us:

$$\dim(\ker(T)) = 4 - 2 = 2$$

So the nullity of T is 2. This means that the kernel of T is a 2-dimensional subspace of $\mathbb{R}^4$, and the image of T is a 2-dimensional subspace of $\mathbb{R}^3$.

Example 2.3.5. Consider the differential operator $D : \mathcal{P}_3(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$ given by

$$D(p(x)) = p'(x)$$

We know that $\mathcal{P}_3(\mathbb{R})$ has dimension 4 and $\mathcal{P}_2(\mathbb{R})$ has dimension 3. The kernel of D consists of all polynomials $p(x)$ such that $p'(x) = 0$, which are precisely the constant polynomials, so $\ker(D) = \operatorname{span}\{1\}$, and thus, $\dim(\ker(D)) = 1$. Then by the Dimension Theorem:

$$\dim(\operatorname{ran}(D)) = \dim(\mathcal{P}_3(\mathbb{R})) - \dim(\ker(D)) = 4 - 1 = 3$$

So the image of D is all of $\mathcal{P}_2(\mathbb{R})$, i.e., D is onto.

Example 2.3.6. Let $T : \mathcal{M}_2(\mathbb{R}) \to \mathcal{M}_2(\mathbb{R})$ be defined by

$$T(A) = A - A^T$$

From Example 2.2.6, we know that $\ker(T) = \mathcal{S}_2(\mathbb{R})$ and $\operatorname{ran}(T) = \mathcal{S}_2^s(\mathbb{R})$. Both $\ker(T)$ and $\operatorname{ran}(T)$ are 3- and 1-dimensional, respectively. So, $\dim(\ker(T)) = 3$, $\dim(\operatorname{ran}(T)) = 1$, and $\dim(\mathcal{M}_2(\mathbb{R})) = 4$. The Dimension Theorem tells us that:

$$\dim(\ker(T)) + \dim(\operatorname{ran}(T)) = 3 + 1 = 4 = \dim(\mathcal{M}_2(\mathbb{R}))$$

Example 2.3.7. Let $V = \mathcal{P}_n(\mathbb{R})$ and $W = \mathbb{R}$, and consider the *evaluation map* $\varphi_a : \mathcal{P}_n(\mathbb{R}) \to \mathbb{R}$ by $\varphi_a(p(x)) = p(a)$. It is easy to note that φ_a is a linear operator and that φ_a is onto. In this case, $\dim(\operatorname{ran}(\varphi_a)) = \dim(\mathbb{R}) = 1$, so $\dim(\ker(\varphi_a)) = (n+1) - 1 = n$ by the Dimension Theorem. Now, each of the n polynomials $(x-a)$, $(x-a)^2$,..., $(x-a)^n$ is in $\ker(\varphi_a)$, and they are linearly independent, so they are a basis because $\dim(\ker(\varphi_a)) = n$.

2.4 Isomorphisms of Linear Transformations

In linear algebra, an *isomorphism* is a special type of linear transformation that perfectly preserves the structure of vector spaces. Two vector spaces are said to be *isomorphic* if there exists a bijective linear transformation (i.e., one-to-one and onto) between them. In essence, isomorphic vector spaces are "the same" for all practical linear algebraic purposes—they may appear different, but their structure, dimension, and behavior under linear operations are identical.

Studying isomorphisms allows us to classify vector spaces according to their structure. For example, any n-dimensional vector space over a field $\mathbb{F}$ is isomorphic to $\mathbb{F}^n$. This means that every finite-dimensional vector space behaves, in some sense, like $\mathbb{R}^n$ or $\mathbb{C}^n$, depending on the field.

Understanding isomorphisms not only simplifies our perspective on vector spaces, but also plays a key role in applications across mathematics, physics, and computer science. In this section, we will define isomorphisms formally, discuss when two vector spaces are isomorphic, and explore how to construct and verify such transformations.

Definition 2.4.1. Let V and W be vector spaces over $\mathbb{F}$. A linear transformation $T : V \to W$ is said to be an *isomorphism* if T is both one-to-one and onto, that is, T is a bijection. If there exists an isomorphism $T : V \to W$, then we say that V and W are *isomorphic* vector spaces, and write $V \simeq W$.

Example 2.4.2. For any vector space V, the identity transformation $\mathrm{id}_V : V \to V$ given by $\mathrm{id}_V(\vec{v}) = \vec{v}$ is an isomorphism. In other words, $V \simeq V$.

Example 2.4.3. Let $T : \mathcal{M}_{nm}(\mathbb{R}) \to \mathcal{M}_{mn}(\mathbb{R})$ be the linear transformation defined by $T(A) = A^T$. Then T is an isomorphism, and thus, $\mathcal{M}_{nm} \simeq \mathcal{M}_{mn}$. To see this, we first show that T is injective. Indeed, for any $A \in \mathcal{M}_{nm}(\mathbb{R})$ such that $T(A) = \vec{0}$, we have $A^T = \vec{0}$, but then $A = (A^T)^T = \vec{0}^T = \vec{0}$, which shows that $\ker(T) = \{\vec{0}\}$.

Of course, it is also easy to determine that T is onto, since for any matrix $B \in \mathcal{M}_{mn}(\mathbb{R})$, there is another matrix $B^T \in \mathcal{M}_{nm}(\mathbb{R})$ such that $T(B^T) = B$, so T is also onto. Therefore, since T is both one-to-one and onto, T is an isomorphism, and thus, $\mathcal{M}_{nm}(\mathbb{R}) \simeq \mathcal{M}_{mn}(\mathbb{R})$.

Example 2.4.4. Let $V = \mathbb{R}^3$ and $W = \mathcal{P}_2(\mathbb{R})$, the vector space of real polynomials of degree at most 2. Define the map $T : V \to W$ by

$$T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = a + bx + cx^2$$

It is easy to check that T is linear and that T is one-to-one and onto, so that T is an isomorphism, so $\mathbb{R}^3 \simeq \mathcal{P}_2(\mathbb{R})$.

The word *isomorphism* comes from two Greek roots: *iso* meaning "same", and *morphos*, meaning "form". An isomorphism $T : V \to W$ induces a pairing $\vec{v} \leftrightarrow T(\vec{v})$ between vectors $\vec{v} \in V$, and vectors in $T(\vec{v})$ that preserves vector addition and scalar multiplication. Hence, as far as their vector space properties are concerned, the spaces V and W are identical except for notation.

The following theorem gives a pretty useful characterization of isomorphisms.

Theorem 2.4.5. *Let V and W be finite-dimensional vector spaces over $\mathbb{F}$ and let $T : V \to W$ be a linear transformation. The following are equivalent.*

1. *T is an isomorphism.*
2. *If $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ is any basis of V, then $\{T(\vec{e}_1), T(\vec{e}_2), ..., T(\vec{e}_n)\}$ is a basis of W.*
3. *There exists a basis $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ of V such that $\{T(\vec{e}_1), T(\vec{e}_2), ..., T(\vec{e}_n)\}$ is a basis of W.*

Proof. (1) $\Rightarrow$ (2): Let $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ be a basis of V. If

$$\alpha_1 T(\vec{e}_1) + \alpha_2 T(\vec{e}_2) + \cdots + \alpha_n T(\vec{e}_n) = \vec{0}_W$$

with $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}$, then

$$T(\alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2 + \cdots + \alpha_n \vec{e}_n) = \vec{0}_V$$

(since $\ker(T) = \{\vec{0}_V\}$). But then each $\alpha_i = 0$ by the independence of $\vec{e}_i$, so $\{T(\vec{e}_1), T(\vec{e}_2), ..., T(\vec{e}_n)\}$ is linearly independent. To see that it spans W, let $\vec{w} \in W$. Since T is onto, $\vec{w} = T(\vec{v})$ for some $\vec{v} \in V$, so write

$$\vec{v} = \alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2 + \cdots + \alpha_n \vec{e}_n$$

Hence we obtain

$$\vec{w} = T(\vec{v}) = \alpha_1 T(\vec{e}_1) + \alpha_2 T(\vec{e}_2) + \cdots + \alpha_n T(\vec{e}_n)$$

proving that $\{T(\vec{e}_1), T(\vec{e}_2), ..., T(\vec{e}_n)\}$ spans W, and thus, $\{T(\vec{e}_1) T(\vec{e}_2), ..., T(\vec{e}_n)\}$ is a basis of W.

(2) $\Rightarrow$ (3): This holds because V is a basis.

(3) $\Rightarrow$ (1): If $T(\vec{v}) = \vec{0}_W$, then write

$$\vec{v} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_n \vec{v}_n$$

where $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}$. Then

$$\vec{0}_W = T(\vec{v}) = \alpha_1 T(\vec{e}_1) + \alpha_2 T(\vec{v}_2) + \cdots + \alpha_n T(\vec{e}_n)$$

so $\alpha_1 = 0$, $\alpha_2 = 0$, ..., $\alpha_n = 0$, so $\ker(T) = \{\vec{0}_V\}$, and T is one-to-one. To show that T is onto, let $\vec{w} \in W$. By assumption, there exists $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}$ such that

$$\vec{w} = \alpha_1 T(\vec{e}_1) + \alpha_2 T(\vec{e}_2) + \cdots + \alpha_n T(\vec{e}_n) = T(\alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2 + \cdots + \alpha_n \vec{e}_n)$$

Therefore, T is onto. □

Remark 2.4.6. Theorem 2.4.5 dovetails nicely with Theorem 2.1.17 as follows. Let V and W be vector spaces over $\mathbb{F}$ of dimension n, and suppose that $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ and $\{\vec{f}_1, \vec{f}_2, ..., \vec{f}_n\}$ are bases of V and W, respectively. Theorem 2.1.17 asserts that there exists a linear transformation $T : V \to W$ such that

$$T(\vec{e}_i) = \vec{f}_i$$

for all $i = 1, 2, ..., n$. Then $\{T(\vec{e}_1), T(\vec{e}_2), ..., T(\vec{e}_n)\}$ is a basis of W, so T is an isomorphism by Theorem 2.4.5. Furthermore, the action of T is prescribed by

$$T(\alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2 + \cdots + \alpha_n \vec{e}_n) = \alpha_1 \vec{f}_1 + \alpha_2 \vec{f}_2 + \cdots + \alpha_n \vec{f}_n$$

so isomorphisms between spaces of equal dimension can easily be defined as soon as bases are known. In particular, this shows that if two vector spaces V and W have the same dimension, then they are isomorphic, that is, $V \simeq W$.

Theorem 2.4.7. *If V and W are finite-dimensional vector spaces over $\mathbb{F}$, then $V \simeq W$ if and only if $\dim(V) = \dim(W)$.*

Proof. It is easy to note that if $\dim(V) = \dim(W)$, then $V \simeq W$ by Remark 2.4.6, so it suffices to show that if $V \simeq W$, then $\dim(V) = \dim(W)$. In this case, there exists an isomorphism $T : V \to W$. Since V is finite-dimensional, let $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ is a basis of V, then $\{T(\vec{e}_1), T(\vec{e}_2), ..., T(\vec{e}_n)\}$ is a basis of W by Theorem 2.4.5 so $\dim(W) = n = \dim(V)$. □

As it turns out, the notion of composition of linear operators and isomorphisms have a well-established connection among each other.

Theorem 2.4.8. *Let V and W be finite-dimensional vector spaces over $\mathbb{F}$, and let $T : V \to W$ be a linear transformation. The following are equivalent.*

1. *T is an isomorphism.*
2. *There exists a linear transformation $S : W \to V$ such that $ST = \mathrm{id}_V$ and $TS = \mathrm{id}_W$.*

Moreover, in this case, S is also an isomorphism and is uniquely determined by T: If $\vec{w} \in W$ is written as $\vec{w} = T(\vec{v})$, then $S(\vec{w}) = \vec{v}$.

Proof. To see that (1) $\Rightarrow$ (2), let $\mathcal{B}_1 = \{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ be a basis of V. Then $\mathcal{B}_2 = \{T(\vec{e}_1), T(\vec{e}_2), ..., T(\vec{e}_n)\}$ is a basis of W by Theorem **??**. Hence, by Theorem 2.1.17, define $S : W \to V$ by

$$S(T(\vec{e}_i)) = \vec{e}_i$$

for each i.

Since $\vec{e}_i = \mathrm{id}_V(\vec{e}_i)$, then $ST = \mathrm{id}_V$ by Theorem 2.1.16. But applying T gives

$$T(S(T(\vec{e}_i))) = T(\vec{e}_i)$$

for each i, so $TS = \mathrm{id}_W$.

To see that (2) $\Rightarrow$ (1), if $T(\vec{v}_1) = T(\vec{v}_2)$, then $S(T(\vec{v}_1)) = S(T(\vec{v}_2))$. Since $ST = \mathrm{id}_V$ by assumption, this implies $\vec{v}_1 = \vec{v}_2$, so T is one-to-one. Given $\vec{w} \in W$, the fact that $TS = \mathrm{id}_W$ means that $\vec{w} = T(S(\vec{w}))$, so T is onto.

Finally, S is uniquely determined by the condition $ST = \mathrm{id}_V$ since this condition implies $S(T(\vec{e}_i)) = \vec{e}_i$ for each i. S is an isomorphism since it carries the basis $\mathcal{B}_2$ to $\mathcal{B}_1$. As to the last assertion, for any $\vec{w} \in W$, write

$$\vec{w} = \alpha_1 T(\vec{e}_1) + \alpha_2 T(\vec{e}_2) + \cdots + \alpha_n T(\vec{e}_n)$$

Then $\vec{w} = T(\vec{v})$, where

$$\vec{v} = \alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2 + \cdots + \alpha_n \vec{e}_n$$

Then $S(\vec{w}) = \vec{v}$. ☐

Definition 2.4.9. Given an isomorphism $T : V \to W$, the unique isomorphism $S : W \to V$ satisfying (2) of Theorem 2.4.8 is called the *inverse* of T, denoted by T^{-1}. Hence, $T : V \to W$ and $T^{-1} : W \to V$ are related by fundamental identities

$$T^{-1}T(\vec{v}) = \vec{v} \text{ for all } \vec{v} \in V$$

and

$$TT^{-1}(\vec{w}) = \vec{w} \text{ for all } \vec{w} \in W$$

In other words, each of T and T^{-1} reverses the action of the other. In particular, the equation in the proof of Theorem 2.4.8 $S(T(\vec{e}_i)) = \vec{e}_i$ shows how to define T^{-1} using the mage of a basis under the isomorphism T.

Corollary 2.4.10. *"$\simeq$" forms an equivalence relation on the class of all isomorphic vector spaces That is, if U, V, and W are vector spaces over $\mathbb{F}$, then*

1. *$V \simeq V$ for every vector space V.*
2. *If $V \simeq W$, then $W \simeq V$.*
3. *If $U \simeq V$ and $V \simeq W$, then $U \simeq W$.*

Proof. To see that "$\simeq$" forms an equivalence relation, we need to verify reflexivity, symmetry, and transitivity. For (1), note that by Example 2.4.2, the identity operator is an isomorphism from V to V, so $V \simeq V$ for any vector space V. For (2), if $V \simeq W$ then there exists an isomorphism $T : V \to W$. But because T is a bijection, there exists an inverse $T^{-1} : W \to V$ such that T^{-1} is an isomorphism, so $W \simeq V$. Finally, if $U \simeq V$ and $V \simeq W$, then there exists isomorphisms $T : U \to V$ and $S : V \to W$. Then taking the composition $ST : U \to W$ is also an isomorphism, so $U \simeq W$. $\square$

Corollary 2.4.11. *If V is a finite-dimensional vector space over $\mathbb{F}$ with $\dim(V) = n$, then $V \simeq \mathbb{R}^n$.*

If V is a vector space of dimension n, note that there are important explicit isomorphisms $V \to \mathbb{R}^n$. Fix a basis $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, ..., \vec{b}_n\}$ of V, and write $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ for the standard basis of $\mathbb{R}^n$. By Theorem 2.4.5, there is a unique linear combination $C_{\mathcal{B}} : V \to \mathbb{R}^n$ given by

$$C_{\mathcal{B}}(\alpha_1\vec{b}_1 + \alpha_2\vec{b}_2 + \cdots + \alpha_n\vec{b}_n) = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \cdots + \alpha_n\vec{e}_n = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

where each $\alpha_i \in \mathbb{R}$. Moreover, $C_{\mathcal{B}}(\vec{b}_i) = \vec{e}_i$ for each i, so $C_{\mathcal{B}}$ is an isomorphism by Theorem 2.4.5, called the *coordinate isomorphism* corresponding to the basis $\mathcal{B}$. Such isomorphisms will be talked more about in the next chapter.

Example 2.4.12. Let $V = \mathcal{S}_2(\mathbb{R})$. Suppose we want to find an isomorphism $T : \mathcal{P}_2(\mathbb{R}) \to \mathcal{S}_2(\mathbb{R})$ such that $T(1) = I_2$, where I_2 is the 2×2 identity matrix.

Note that $\{1, x, x^2\}$ is a basis of $\mathcal{P}_2(\mathbb{R})$ and we want a basis of $\mathcal{S}_2(\mathbb{R})$ containing I_2. Indeed, the set

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

is linearly independent of V, so it is a basis because $\dim(V) = 3$ (see Example 1.3.21). Hence, define $T : \mathcal{P}_2(\mathbb{R}) \to \mathcal{S}_2$ by taking $T(1) = I_2$, $T(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, and $T(x^2) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, and extending linearly as in Theorem 2.1.17. T is an isomorphism by Theorem 2.1.17 and its action is given by

$$T(a + bx + cx^2) = aT(1) + bT(x) + cT(x^2) = \begin{bmatrix} a & b \\ b & a + c \end{bmatrix}$$

The Dimension Theorem (Theorem 2.3.1) gives the following useful fact about isomorphisms.

Theorem 2.4.13. *If V and W are finite-dimensional vector spaces over $\mathbb{F}$ with $\dim(V) = \dim(W) = n$, a linear transformation $T : V \to W$ is an isomorphism if it is either one-to-one or onto.*

Proof. The Dimension Theorem asserts that $\dim(\ker(T)) + \dim(\operatorname{ran}(T)) = n$, so $\dim(\ker(T)) = 0$ if and only if $\dim(\operatorname{ran}(T)) = n$, so T is one-to-one if and only if T is onto, and thus, the result follows. $\square$

Example 2.4.14. Let $T : \mathcal{P}_1(\mathbb{R}) \to \mathcal{P}_1(\mathbb{R})$ be defined by $T(a + bx) = (a - b) + ax$. Then T is a linear transformation. Note that since $T(1) = 1 + x$ and $T(x) = -1$, T carries the basis $\mathcal{B}_1 = \{1, X\}$ to the basis $\mathcal{B}_2 = \{1 + x, -1\}$. Hence, T is an isomorphism, and T^{-1} carries $\mathcal{B}_2$ back to $\mathcal{B}_1$, that is,

$$T^{-1}(1 + x) = 1 \text{ and } T^{-1}(-1) = x$$

Since $a + bx = b(1 + x) + (b - a)(-1)$, we obtain

$$T^{-1}(a + bx) = bT^{-1}(1 + x) + (b - a)T^{-1}(-1) = b + (b - a)x$$

Example 2.4.15. If $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, ..., \vec{b}_n\}$ is a basis of a vector space V, the coordinate transformation $C_{\mathcal{B}} : V \to \mathbb{R}^n$ is an isomorphism defined by

$$C_{\mathcal{B}}(\alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \cdots + \alpha_n \vec{b}_n) = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

The way to reverse the action of $C_{\mathcal{B}}$ is clear: $C_{\mathcal{B}}^{-1} : \mathbb{R}^n \to V$ is given by

$$C_{\mathcal{B}}^{-1}\left(\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}\right) = \alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \cdots + \alpha_n \vec{b}_n$$

for all $\alpha_i \in V$.

Example 2.4.16. Define $T : \mathbb{R}^3 \to \mathbb{R}^3$ to be the linear transformation defined by $T(x, y, z) = (z, x, y)$. Note that

$$T^2(x, y, z) = T(T(x, y, z)) = T(z, x, y) = (y, z, x)$$

and then

$$T^3(x, y, z) = T(T^2(x, y, z)) = T(y, z, x) = (x, y, z)$$

Since this holds for all $(x, y, z) \in \mathbb{R}^3$, $T^3 = \mathrm{id}_{\mathbb{R}^3}$, so $T(T^2) = (T^2)T$, so $T^{-1} = T^2$.

Index