Chapter 5

Linear Transformations and Isomorphism

5.1 Linear Transformations

Linear transformations allow us to compare vector spaces over the same field: They are those functions between vector spaces that are compatible with the vector space structures. Formally,

Definition 5.1. A *linear transformation* is a function

$$T : V \longrightarrow W$$

such that for all $\mathbf{x}, \mathbf{y} \in V$ and $\lambda, \mu \in \mathbb{F}$

$$T(\lambda\mathbf{x} + \mu\mathbf{y}) \;=\; \lambda T(\mathbf{x}) + \mu T(\mathbf{y}).$$

Lemma 5.2. *Let U, V and W be vector spaces over the field $\mathbb{F}$.*

(a) The zero map

$$\mathbf{0} : V \longrightarrow W, \quad \mathbf{v} \longmapsto \mathbf{0}_W$$

is a linear transformation.

(b) The identity map

$$id_V : V \longrightarrow V, \quad \mathbf{v} \longmapsto \mathbf{v}$$

is a linear transformation.

(c) If $S : U \longrightarrow V$ and $T : V \longrightarrow W$ are linear transformations, so is their composition

$$T \circ S \colon U \longrightarrow W, \quad \mathbf{u} \longmapsto T(S(\mathbf{u}))$$

Proof. (a) is an immediate consequence of the fact that $\lambda\mathbf{0}_W + \mu\mathbf{0}_W = \mathbf{0}_W$ for all $\lambda, \mu \in \mathbb{F}$.

(b) is immediate from definition.

(c) Take $\mathbf{u}, \mathbf{v} \in U$ and $\lambda, \mu \in \mathbb{F}$. Then

$$\begin{aligned}
(T \circ S)(\lambda\mathbf{u} + \mu\mathbf{v}) &:= T\Big(S(\lambda\mathbf{u} + \mu\mathbf{v})\Big) \\
&= T\Big(\lambda S(\mathbf{u}) + \mu S(\mathbf{v})\Big) && \text{as } S \text{ is linear} \\
&= \lambda T\big(S(\mathbf{u})\big) + \mu T\big(S(\mathbf{v})\big) && \text{as } T \text{ is linear} \\
&=: \lambda(T \circ S)(\mathbf{u}) + \mu(T \circ S)(\mathbf{v})
\end{aligned}$$

□

Our first examples of vector spaces included F^n. We determine all linear transformations between vector spaces of this form.

Theorem 5.3. *The function $T\colon \mathbb{F}^n \longrightarrow \mathbb{F}^m$ is a linear transformation if and only if there are $a_{ij} \in \mathbb{F}$ $(i = 1, \dots, m,\ j = 1, \dots, n)$ such that for all $(x_1, \dots, x_n) \in \mathbb{F}^n$,*

$$T(x_1, \dots, x_n) := (\sum_{j=1}^{n} a_{1j}x_j, \dots, \sum_{j=1}^{n} a_{mj}x_j).$$

In particular $T : \mathbb{F} \longrightarrow \mathbb{F}$ is linear if and only if there is an $a \in \mathbb{F}$ such that $T(x) = ax$ for all $x \in \mathbb{F}$.

Proof. Suppose that for some $a_{ij} \in \mathbb{F}$ $(i = 1, \dots, m,\ j = 1, \dots, n)$

$$T(x_1, \dots, x_n) = (\sum_{j=1}^{n} a_{1j}x_j, \dots, \sum_{j=1}^{n} a_{mj}x_j).$$

Take $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{F}^n$ and $\lambda, \mu \in \mathbb{F}$. Then

$$\begin{aligned}
&T\big(\lambda(x_1, \dots, x_n) + \mu(y_1, \dots, y_n)\big)\\
&\qquad = T\big((\lambda x_1 + \mu y_1), \dots, (\lambda x_n + \mu y_n)\big)\\
&\qquad = \big(\sum_{j=1}^{n} a_{1j}(\lambda x_j + \mu y_j), \dots, \sum_{j=1}^{n} a_{mj}(\lambda x_j + \mu y_j)\big)\\
&\qquad = \big(\lambda \sum_{j=1}^{n} a_{1j}x_j + \mu \sum_{j=1}^{n} y_j), \dots, \lambda \sum_{j=1}^{n} a_{mj}x_j + \mu \sum_{j=1}^{n} y_j)\big)\\
&\qquad = \lambda(\sum_{j=1}^{n} a_{1j}x_j, \dots, \sum_{j=1}^{n} a_{mj}x_j) + \mu(\sum_{j=1}^{n} a_{mj}y_j, \dots,, \sum_{j=1}^{n} a_{mj}y_j)\\
&\qquad = \lambda T(x_1, \dots, x_n) + \mu T(y_1, \dots, y_n)
\end{aligned}$$

Conversely, suppose $T\colon \mathbb{F}^m \longrightarrow \mathbb{F}^n$ is a linear transformation. Take $(x_1, \dots, x_n) \in \mathbb{F}^n$. Then

$$(x_1, \dots, x_n) = x_1(1, 0, \dots, 0) + \cdots + x_n(0, \dots, 0, 1),$$

so that

$$T(x_1, \dots, x_n) = x_1T(1, 0, \dots, 0) + \cdots + x_nT(0, \dots, 0, 1).$$

There are $a_{11}, \dots, a_{m1}, \dots, a_{1n}, \dots, a_{mn} \in \mathbb{F}$ such that

$$T(1, 0, \dots, 0) =: (a_{11}, \dots, a_{m1}),\ \cdots\cdots,\ T(0, 0, \dots, 1) =: (a_{1n}, \dots, a_{mn}).$$

Then

$$\begin{aligned}
T(x_1, \dots, x_n) &= x_1(a_{11}, \cdots, a_{m1}) + \cdots + x_n(a_{1n}, \dots, a_{mn})\\
&= (\sum_{j=1}^{n} a_{1j}x_j, \dots, \sum_{j=1}^{n} a_{mj}x_j)
\end{aligned}$$

□

Definition 5.4. The *kernel* of the linear transformation $T: V \to W$ T, $\ker(T)$, is defined by

$$\ker(T) := \{\mathbf{x} \in V \mid \; T(\mathbf{x}) = \mathbf{0}_W\}$$

Definition 5.5. The linear transformation $T : V \longrightarrow W$ is

(a) *1–1, injective* or a *monomorphism* if and only if $\mathbf{u} = \mathbf{v}$ whenever $T(\mathbf{u}) = T(\mathbf{v})$;

(b) *onto, surjective* or an *epimorphism* if and only if for each $\mathbf{w} \in W$ there is a $\mathbf{v} \in V$ with $\mathbf{w} = T(\mathbf{v})$;

(c) *1–1 and onto* or *bijective* if and only it is both 1–1 and onto;

(d) an *endomorphism* if and only if $W = V$.

Lemma 5.6. *Let $T : V \longrightarrow W$ be a linear transformation. Then*

(i) T is injective if and only if $\ker(T) = \{\mathbf{0}_V\}$;

(ii) T is surjective if and only if $\operatorname{im}(T) = W$.

Proof. (i). $T(\mathbf{u}) = T(\mathbf{v})$ if and only if $T(\mathbf{u}-\mathbf{v}) = \mathbf{0}_W$, which is the case if and only if $\mathbf{u}-\mathbf{v} \in \ker(T)$.
(ii) This is immediate from the definition. □

We now consider further examples.

Example 5.7. Take $V = \mathbb{R}^3$, $W = \mathbb{R}^2$ and

$$T : V \longrightarrow W, \quad (x, y, z) \longmapsto (2x + y + z, x - y)$$

Then T is easily shown to be a linear transformation.

Example 5.8. Put

$$\mathcal{D}(\mathbb{R}) := \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid f \text{ is differentiable}\}\,.$$

It is an elementary result from calculus that $\mathcal{D}(\mathbb{R})$ is a real vector space and that the derivative defines a linear transformation

$$D : \mathcal{D}(\mathbb{R}) \longrightarrow \mathcal{F}(\mathbb{R}), \quad f \longmapsto f'$$

where $\mathcal{F}(\mathbb{R})$ is as defined Example 3.16, and we have written f' for the derivative of f.

That D is a linear transformation is simply a restatement of the familiar rule from calculus that given $\alpha, \beta \in \mathbb{R}$ and $f, g \in \mathcal{D}(\mathbb{R})$,

$$\frac{d}{dx}(\alpha f + \beta g) = \alpha\frac{df}{dx} + \beta\frac{dg}{dx}$$

5.2 Isomorphism

We consider two vector spaces over $\mathbb{F}$ to be essentially the same when the only differences between them are purely notational. We formulate rigorously using the language of linear transformations, including the notion of *isomorphism*.

Definition 5.9. The linear transformation $T : V \longrightarrow W$ is an *isomorphism* if and only if there is a linear transformation $S : W \to V$ such that

$$S \circ T = id_V \qquad \text{and} \qquad T \circ S = id_W.$$

An endomorphism $T : V \longrightarrow V$ which is an isomorphism is called an *automorphism*.

The vector spaces V and W over the field $\mathbb{F}$ are *isomorphic* if and only if there is an isomorphism $T : V \longrightarrow W$.

We write $V \cong W$ when V is isomorphic to W.

To decide whether a given linear transformation $T : V \to W$ is an isomorphism requires, *prima facie*, trying all possible linear transformations $S : W \to V$ to see which, if any, satisfy the conditions of the last definition. This is, at best, an unwieldy task. It is therefore desirable to find intrinsic criteria to decide whether $T : V \to W$ is an isomorphism. Fortunately, this is possible.

Theorem 5.10. *The linear transformation* $T\colon V \to W$ *is an isomorphism if and only if it is both an epimorphism and a monomorphism.*

Proof. Since a function has an inverse if and only if it is bijective (Theorem 1.33), it is sufficient to show that if the linear transformation $T : V \to W$ has an inverse *function*, $S : W \to V$, then S must also be a linear transformation.

Let $S\colon W \to V$ be the inverse function to the linear transformation $T\colon V \to W$. Take $\mathbf{u}, \mathbf{w} \in W$ and $\lambda, \mu \in \mathbb{F}$. Then

$$\begin{aligned}
S(\lambda\mathbf{u} + \mu\mathbf{w}) &= S\big(\lambda(T \circ S)(\mathbf{u}) + \mu(T \circ S)(\mathbf{w})\big) && \text{as } T \circ S = id_W \\
&= S\big(\lambda T(S(\mathbf{u})) + \mu T(S(\mathbf{w}))\big) \\
&= S\big(T(\lambda S(\mathbf{u}) + \mu S(\mathbf{w}))\big) && \text{as } T \text{ is linear} \\
&= (S \circ T)(\lambda S(\mathbf{u}) + \mu S(\mathbf{w})) \\
&= \lambda S(\mathbf{u}) + \mu S(\mathbf{w}) && \text{as } S \circ T = id_V
\end{aligned}$$

□

While this theorem provides a satisfactory intrinsic criterion for deciding whether a given linear transformation is an isomorphism, it does not do the same for deciding whether two given vector spaces are isomorphic. After all, we still face the unwieldy and, in principle, infinite task of finding an isomorphism between the vector spaces in question. We shall see (Theorem 8.1) that the question can be decided by means of a single numerical invariant, the *dimension* of a vector space.

5.3 Exercise

Exercise 5.1. Show $f : \mathbb{R}^2 \to \mathbb{R}^3,\ \ (x, y) \mapsto (2x + y, 4x + 17y, 3x - 4y)$ defines a linear transformation of vector spaces over $\mathbb{R}$.

Exercise 5.2. Let $V := \mathcal{C}^\infty(\mathbb{R})$ be the real vector space of all infinitely differentiable functions $f : \mathbb{R} \to \mathbb{R}$, with the vector space operations defined point-wise. Prove that $D : V \to V,\ f \mapsto f'$ is a linear transformation, where f' denotes the derivative of f.

Exercise 5.3. Prove that if the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation of real vector spaces, then there are uniquely determined real numbers a, b, c, d such that for all $(x, y) \in \mathbb{R}^2$, $f(x, y) = (ax + by, cx + dy)$.