Isomorphisms

The linear transformations studied so far are generally referred to as homomorphisms. They satisfy a linearity condition that ensures that their domain and range have a similar vector space structure: their elements behave similarly under vector addition and scalar-vector multiplication. However, they need not be in a one-to-one correspondence[1]. When this is the case, the domain can be 'identified' with the image. This type of mapping is called an isomorphism. We have already come across one such map: the coordinate mapping. Recall that given any $\mathbf{v} \in V$ expanded in a basis B_V as $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n$, we can define the mapping

$$\mathbf{v} \mapsto \mathbf{x} \in \mathbb{F}^n.$$

This linear map establishes a one-to-one correspondence between elements in a generic vector space V and $\mathbb{F}^n$. In this lecture, we outline some of the key results and properties associated with such maps. We start with invertibility.

13.1 Invertible maps

Let $f \colon V \to W$ be a linear map and let $\mathbf{v} \in V$. Recall that the **image** of $\mathbf{v}$ under f is a vector $\mathbf{w} \in W$ given by $\mathbf{w} = f(\mathbf{v})$. Indeed, we have defined the set of such images as the image set $f(V)$ (or im f):

$$f(V) := \{\mathbf{w} \in W : \mathbf{w} = f(\mathbf{v}), \mathbf{v} \in V\} \subseteq W.$$

Assume now that f is one-to-one (injective). Then **only one** vector in the domain of f is mapped to one vector in the image of f. Given this uniqueness, we can choose to define it as follows.

> **Definition 13.1 — Pre-image.** Let $f \in \mathcal{L}(V, W)$ be injective and let $\mathbf{v} \in V$ and $\mathbf{w} = f(\mathbf{v}) \in \operatorname{im} f$. The vector $\mathbf{v}$ is called the **pre-image** or **inverse image** of $\mathbf{w}$ under f and is denoted by $\mathbf{v} = f^{-1}(\mathbf{w})$.

Not all $\mathbf{w} \in W$ are guaranteed to have an inverse image under f, unless f is surjective. On the other hand, by definition of image set, all $\mathbf{w} \in \operatorname{im} f$ have a pre-image when f is injective. Hence, we can extend/apply the notation f^{-1} to all elements of im f. Thus, we can view f^{-1} as a map with domain im f and codomain V: $f^{-1} \colon \operatorname{im} f \to V$. Moreover, note that any pre-image $\mathbf{v}$ of $\mathbf{w} \in \operatorname{im} f = f(V)$, satisfies

$$\mathbf{v} = f^{-1}(\mathbf{w}) = f^{-1}(f(\mathbf{v})),$$

---

[1]One-to-one correspondence means bijection; one-to-one mapping means injection.

while any $\mathbf{w} \in \operatorname{im} f$ satisfies

$$\mathbf{w} = f(\mathbf{v}) = f(f^{-1}(\mathbf{w})).$$

In other words, $f \circ f^{-1} = id_W$ and $f^{-1} \circ f = id_V$. This discussion justifies the following definition.

> **Definition 13.2 — Invertible map.** Let $f \colon V \to W$ be a linear map. We say f is **invertible** if there exists a linear map $g \colon W \to V$ such that
>
> $$g \circ f = id_V \quad \text{and} \quad f \circ g = id_W.$$
>
> The map g is denoted by $g = f^{-1}$ and is called the **inverse** (map) of f.

> **Proposition 13.1** The inverse of a linear map $f \colon V \to W$ is unique, if it exists.

Proof. Let $g, h \colon W \to V$ be inverses of f. Then $h = h \circ id_W = h \circ (f \circ g) = (h \circ f) \circ g = id_V \circ g = g$. ∎

> **Proposition 13.2** Let $f \colon V \to W$ be a linear map. Then f is invertible if and only if it is bijective.

Proof. Let f be a linear map.
$\boxed{\Longrightarrow}$ Assume f is invertible. Let $f(\mathbf{v}_1) = f(\mathbf{v}_2)$. Then $\mathbf{v}_1 = f^{-1}(f(\mathbf{v}_1)) = f^{-1}(f(\mathbf{v}_2)) = \mathbf{v}_2$, so that f is injective. Let now $\mathbf{w} \in W$. Then there exists $\mathbf{v} = f^{-1}(\mathbf{w})$ such that $f(\mathbf{v}) = f(f^{-1}(\mathbf{w})) = \mathbf{w}$, so that f is surjective. Therefore f is bijective.
$\boxed{\Longleftarrow}$ Assume f is bijective. We show that (i) there exists a map $g \colon W \to V$ such that $g \circ f = id_V$ and $f \circ g = id_W$ and (ii) that g is linear. This is essentially the argument that motivated the definition of invertibility. Since f is surjective, every $\mathbf{w} \in W$ will have a corresponding pre-image $\mathbf{v} \in V$, which by injectivity is unique. Therefore, the map $g \colon W \to V$ given by $g(\mathbf{w}) = \mathbf{v}$ is well-defined. As a pre-image, $\mathbf{v}$ also satisfies $f(\mathbf{v}) = \mathbf{w}$. Hence,

$$\mathbf{w} = f(\mathbf{v}) = f(g(\mathbf{w})), \quad \mathbf{v} = g(\mathbf{w}) = g(f(\mathbf{v}))$$

so that

$$f \circ g = id_W, \qquad g \circ f = id_V.$$

Finally, let $\mathbf{v}_1 = g(\mathbf{w}_1), \mathbf{v}_2 = g(\mathbf{w}_2)$. Then g is a linear map since

$$f(a\mathbf{v}_1 + b\mathbf{v}_2) = f(ag(\mathbf{w}_1) + bg(\mathbf{w}_2)) = af(g(\mathbf{w}_1)) + bf(g(\mathbf{w}_2)) = a\mathbf{w}_1 + b\mathbf{w}_2,$$

by linearity of f, so that

$$g(a\mathbf{w}_1 + b\mathbf{w}_2) = g(f(a\mathbf{v}_1 + b\mathbf{v}_2)) = a\mathbf{v}_1 + b\mathbf{v}_2 = ag(\mathbf{w}_1) + ag(\mathbf{w}_2).$$

∎

> **Proposition 13.3** Let $f \colon V \to W$ be a linear map. Then f is invertible if and only if its matrix representation is invertible.

Proof. Let $f \colon V \to W$ be a linear map and let $n = \dim V$.
$\boxed{\Longrightarrow}$ Assume f is invertible. Then it is bijective and therefore $\ker f$ is trivial and $\operatorname{im} f = W$. By the rank-nullity formula, $\dim W = n$. Moreover, by definition of invertible maps, there exists a linear map such that

$$g \circ f = id_V \quad \text{and} \quad f \circ g = id_W.$$

By Proposition , the $n \times n$ matrix representations of f and g satisfy

$$A_g A_f = I_n \quad \text{and} \quad A_f A_g = I_n.$$

Hence, by the definition of matrix inverse, A_f has an inverse and is therefore invertible.

$\boxed{\Longleftarrow}$ Assume the matrix representation of f is a square matrix A which is invertible. Then there exists a matrix B such that

$$AB = I_n \quad \text{and} \quad BA = I_n.$$

By Proposition 11.10, the above relations correspond to compositions of linear maps f and g with matrix representations A and B, respectively in some bases for V and W. More precisely,

$$g \circ f = id_V \quad \text{and} \quad f \circ g = id_W.$$

By definition, f is an invertible map. $\blacksquare$

The corresponding commutative diagram is included below.

13.2   Isomorphic spaces

Definition 13.3 A linear bijection $f\colon V \to W$ is called an **isomorphism**. In this case, the spaces V and W are said to be **isomorphic**; we write $V \cong W$.

Proposition 13.4 Let $f\colon V \to W$ be an isomorphism. Then $f^{-1}\colon W \to V$ is an isomorphism.

Proof. Since f is bijective, it is invertible, so the inverse map f^{-1} is linear. By definition, f^{-1} is invertible, therefore bijective, so an isomorphism. $\blacksquare$

Proposition 13.5 Let $f\colon V \to W$ be a linear map and let B_V be a basis for V. Then f is an isomorphism if and only if $f(B_V)$ is a basis for W.

Proof. Let $f\colon V \to W$ be a linear map and let $B_V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for V.

$\boxed{\Longrightarrow}$ Assume f is an isomorphism. Since f is injective, its kernel is trivial. Since f is surjective, im $f = W$. By Proposition 9.4, since B_V is a linearly independent set, $f(B_V)$ is also linearly independent. Moreover, since f is linear, by Proposition 9.3, span$f(B_V) = f(\text{span}B_V) = f(V) = $ im $f = W$. Hence, $f(B_V)$ is a linearly independent spanning set for W, i.e., a basis for W.

$\boxed{\Longleftarrow}$ Assume that $f(B_V)$ is a basis for W. By Proposition 9.3, since f is linear, span$f(B_V) = f(\text{span}B_V)$, i.e., $W = $ im f, so that f is surjective. To show f is injective, we show that ker$f$ is trivial. Let $\mathbf{w}_i = f(\mathbf{v}_i)$ for $i = 1, 2, \ldots, n$ and note that $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ is a basis for W. Let $\mathbf{v} \in \ker f$. Then

$$\mathbf{0}_W = f(\mathbf{v}) = f\left(\sum_{i=1}^{n} x_i \mathbf{v}_i\right) = \sum_{i=1}^{n} x_i f(\mathbf{v}_i) = \sum_{i=1}^{n} x_i \mathbf{w}_i.$$

But this implies that $x_i = 0$ for all i, since $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ is a basis for W. Hence $\mathbf{v} = \mathbf{0}_V$ and the kernel is trivial. Thus, f is a bijection, which is also linear, and therefore an isomorphism. $\blacksquare$

Why are we interested in isomorphisms? It is because they allow for results or propertise derived for one space to also hold for the other space, as the following results shows.

Proposition 13.6 Let $f\colon V \to W$ be an isomorphism. Let $S \subseteq V$. Then
   i. S is a spanning set for V if and only if $f(S)$ is a spanning set for W.
   ii. S is linearly independent if and only if $f(S)$ is.
   iii. S is a basis for V if and only if $f(S)$ is a basis for W.
   iv. U is a subspace of V if and only if $f(U)$ is a subspace of W with same dimension.

Proof. Exercise. ∎

The last item in the above proposition is related in some sense to the following key result.

Theorem 13.7 Let V, W be finite-dimensional vector spaces. Then

$$V \cong W \quad \text{if and only if} \quad \dim V = \dim W.$$

Proof. Let $f\colon V \to W$ be a linear map.

$\boxed{\Longrightarrow}$ Let $V \cong W$, with f a bijection. Since f is injective, $\ker f = \{\mathbf{0}_V\}$. Since f is surjective, $\operatorname{im} f = W$. By the rank-nullity formula,

$$\dim V = \text{nullity } f + \operatorname{rank} f = 0 + \dim W.$$

$\boxed{\Longleftarrow}$ Let $\dim V = \dim W = n$. Let $B_V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}, B_W = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ be bases for V and W, respectively. Define the map $f\colon V \to W$ via

$$f(\mathbf{v}) = f(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n) = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \cdots + x_n\mathbf{w}_n =: \mathbf{w}.$$

Then
- f is surjective, as for any $\mathbf{w} \in W$ with coordinates x_i in the basis B_W, there exists $\mathbf{v}$ having the same coordinates in the basis B_V.
- f is injective since

$$f(\mathbf{v}) = f(\mathbf{u}) \iff \mathbf{w} = \mathbf{z} \iff \sum_{i=1}^n x_i\mathbf{w}_i = \sum_{i=1}^n z_i\mathbf{w}_i \iff \sum_{i=1}^n (x_i - z_i)\mathbf{w}_i = \mathbf{0} \iff x_i = z_i \iff \mathbf{v} = \mathbf{u}.$$

   Note that we used the linear independence of B_W to deduce that $x_i = z_i$ for all i.
- f is linear (show this!).

Hence f is a linear bijection and therefore $V \cong W$. ∎

The isomorphism f described in the above theorem has matrix representation $A = I_n$:

Theorem 13.7 allows us to establish quickly which pairs of spaces are not isomorphic, simply by examining the dimensions of the spaces.

An important consequence of Theorem 13.7 is that we can view all n-dimensional real vector spaces as being essentially the same. In particular, they are all isomorphic to $\mathbb{R}^n$.

Corollary 13.8 Let $V(\mathbb{R})$ be an n-dimensional vector space. Then $V \cong \mathbb{R}^n$.

We cannot stress enough the importance of this result: since every finite-dimensional space is isomorphic to $\mathbb{R}^n$, we can restrict our study to this specific vector space, as properties and results obtained for $\mathbb{R}^n$ will be paralleled in any other vector space of dimension n.

(R) The more general result is $V(\mathbb{F}) \cong \mathbb{F}^n$, for any n-dimensional V.

The above descriptions allow us to establish a criterion for identifying isomorphisms.

Proposition 13.9 — Isomorphism criterion. Let $f : V \to W$ be a linear map on a finite-dimensional vector space V. Then f is an isomorphism if any only $\ker f = \{\mathbf{0}_V\}$ and $\dim V = \dim W$.

Example 13.1 — Coordinate map. Let V be an n-dimensional vector space over $\mathbb{R}$ with basis set $B_V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, v_n\}$ and consider the coordinate map $\varphi_V : V \to \mathbb{R}^n$ given by

$$\varphi_V(\mathbf{v}) = \mathbf{x},$$

where $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n$. We can readily check the isomorphism criterion:
- $\ker \varphi_V = \{\mathbf{0}_V\}$, since any $\mathbf{v} \in \ker \varphi_V$ satisfies

$$\mathbf{0}_n = \varphi_V(\mathbf{v}) = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n \implies x_1 = x_2 = \cdots = x_n = 0,$$

  since the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is linearly independent. Hence, any $\mathbf{v} \in \ker \varphi_V$ satisfies $\mathbf{v} = \mathbf{0}_V$.
- $\dim \mathbb{R}^n = \dim V$.

Exercise 13.1 Let $V(\mathbb{F}), W(\mathbb{F})$ have bases B_V, B_W and dimensions n, m, respectively. Define the map $m : \mathcal{L}(V, W) \to \mathbb{F}^{m \times n}$ that associates with a linear map f its matrix representation relative to B_V, B_W:

$$m(f) = A_{VW}.$$

Use the isomorphism criterion to check that m is an isomorphism and therefore that $\mathcal{L}(V, W) \cong \mathbb{F}^{m \times n}$.

Example 13.2 — Counter-example. Let $V = \mathbb{C}, W = \mathbb{R}^2$. For any $z \in \mathbb{C}$ we can define a bijection

$$f(z) = f(a + ib) = \begin{bmatrix} a \\ b \end{bmatrix}.$$

However, V and W are not isomorphic since f does not satisfy the definition of linear map: recall that by definition V is a vector space $V(\mathbb{F})$ and W is a vector space $W(\mathbb{F})$. However, our vector spaces do not share the same field: $\mathbb{C}$ means $\mathbb{C}(\mathbb{C})$, while $\mathbb{R}^2$ means $\mathbb{R}^2(\mathbb{R})$. Hence, the linearity property is undefined in this case.

13.3   Matrix representations

Given the above isomorphism criterion (and also the discussion prior to it), it is evident that the matrix representation of an isomorphism f is square and invertible. This excludes the zero matrix, which means that while the set of isomorphisms is a subset of $\mathcal{L}(V, W)$, it cannot be a subspace. The set of square and invertible matrices forms a (non-Abelian) group under matrix multiplication; however, this property does not extend to the set of isomorphisms equipped with the operation of composition, since if fails to satisfy the closure property. Our next topic, endomorphisms, i.e., maps in $\mathcal{L}(V, V)$, will allow us to establish a group structure to the subset of bijections in $\mathcal{L}(V, V)$.