LECTURE 8

Linear transformations (1)

The material studied so far, namely vector spaces, is motivated by the need to understand the behaviour of (and work with) functions defined on vector spaces. We will restrict our attention to functions that satisfy a linearity property, also known as linear transformations.

8.1 Definitions, notation, properties.

Definition 8.1 — Linear transformation. Let $V(\mathbb{F}), W(\mathbb{F})$ be vector spaces. A linear transformation is a map $f: V \to W$ satisfying the linearity property

$$f(a\mathbf{u} + b\mathbf{v}) = af(\mathbf{u}) + bf(\mathbf{v}),$$

for all vectors $\mathbf{u}, \mathbf{v} \in V$ and all scalars $a, b \in \mathbb{F}$.

The set of linear transformations is denoted by $\mathcal{L}(V, W)$. We write $f \in \mathcal{L}(V, W)$ to indicate a transformation $f: V \to W$ which is linear.

Note that the terms 'transformation' and 'map' will be used interchangeably throughout this course.

R It is important to note that the vector space operations are specific to each of the vector spaces V, W. Using the notation from Lecture 1, given $V(\mathbb{F}) = (V, +, \bullet, \mathbb{F})$ and $W(\mathbb{F}) = (W, +, \bullet, \mathbb{F})$, the linearity property reads

$$f(a \bullet \mathbf{u} + b \bullet \mathbf{v}) = a \bullet f(\mathbf{u}) + b \bullet f(\mathbf{v}).$$

This is clearly awkward to use: we will indeed continue using the vector space operations in the same way, assuming that distinctions are clear from the context.

We will encounter and study (some of) the following types of linear maps:

- **homomorphisms**: these are just general linear transformations $f: V \to W$;
- **isomorphisms**: these are homomorphisms $f: V \to W$ that are invertible;
- **endomorphisms**: these are linear transformations $f: V \to V$;
- **automorphisms**: these are endomorphisms $f: V \to V$ that are invertible.

In this lecture, we consider the general case of homomorphisms, with later lectures dedicated to the study of the other three types of maps.

Using Definition 8.1, we immediately derive the following properties of linear maps:

$$f(\mathbf{u} \pm \mathbf{v}) = f(\mathbf{u}) \pm f(\mathbf{v}), \qquad f(a\mathbf{v}) = af(\mathbf{v}),$$

and by setting $a = 0$ and $a = -1$, respectively, we obtain

$$f(\mathbf{0}_{\mathrm{V}}) = \mathbf{0}_{\mathrm{W}}, \quad f(-\mathbf{v}) = -f(\mathbf{v}),$$

where $\mathbf{0}_{\mathrm{V}}, \mathbf{0}_{\mathrm{W}}$ are the zero vectors in V and W, respectively. Note also that the second relation is rigourously written as $f(\mathbf{v}^-) = f(\mathbf{v})^-$.

We will use the following terminology and notation in relation to linear maps $f : V \to W$:

- V is the **domain** of f;
- W is the **codomain** of f;
- $\mathbf{w} = f(\mathbf{v})$ is the **image** of $\mathbf{v} \in V$;
- $\mathbf{v}$ is a **pre-image** of $\mathbf{w} \in W$, if $\mathbf{w} = f(\mathbf{v})$;
- $\mathrm{im}\ f := f(V) := \{\mathbf{w} \in W : \mathbf{w} = f(\mathbf{v}), \mathbf{v} \in V\} \subseteq W$ is the **image** or **range** of f;
- $\mathrm{ker} f = \{\mathbf{v} \in V : f(\mathbf{v}) = \mathbf{0}_{\mathrm{W}}\} \subseteq V$ is the **kernel** or **nullspace** of f.

We also recall the following properties of general maps.

Definition 8.2 — Injective map. A map $f : V \to W$ is said to be injective (or one-to-one) if for any $\mathbf{v}_1, \mathbf{v}_2 \in V$, $f(\mathbf{v}_1) = f(\mathbf{v}_2)$ implies $\mathbf{v}_1 = \mathbf{v}_2$.

Definition 8.3 — Surjective map. A map $f : V \to W$ is said to be surjective (or onto) if for all $\mathbf{w} \in W$ there exists at least one $\mathbf{v} \in V$ such that $f(\mathbf{v}) = \mathbf{w}$.

Definition 8.4 — Bijective map. A map $f : V \to W$ is said to be bijective if it is injective and surjective.

Definition 8.5 — Trivial kernel. The kernel of a linear map $f : V \to W$ is said to be trivial if $\mathrm{ker} f = \{\mathbf{0}\}$.

In addition to the notation $f(V)$, we will also employ the notation $f(S)$ to indicate the image of a finite set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$:

$$f(S) := \{f(\mathbf{v}_1), f(\mathbf{v}_2), \dots, f(\mathbf{v}_k)\}\,.$$

The following results provide additional descriptions of linear maps.

Proposition 8.1 Let $V(\mathbb{F}), W(\mathbb{F})$ be vector spaces. The set of linear maps $\mathcal{L}(V, W)$ is a vector space over $\mathbb{F}$.

Proof. Exercise. ■

Proposition 8.2 Let $f : V \to W$ be a linear map. Then

i. f is injective if and only if $\mathrm{ker} f = \{\mathbf{0}_{\mathrm{V}}\}$.

ii. f is surjective if and only if $\mathrm{im}\ f = W$.

Proof. i. First, note that since f is a linear map, we have $f(\mathbf{0}_{\mathrm{V}}) = \mathbf{0}_{\mathrm{W}}$.
$\boxed{\Longrightarrow}$ Let f be injective. Assume, by contradiction, that $\mathrm{ker} f$ is not trivial, i.e., there exists $\mathbf{v} \neq \mathbf{0}_{\mathrm{V}}$ such that $f(\mathbf{v}) = \mathbf{0}_{\mathrm{W}}$. Let $\mathbf{v}' \in V$. Then, using the linearity and also the injectivity of f, we find

$$\mathbf{0}_{\mathrm{W}} = f(\mathbf{v}) = f(\mathbf{v}-\mathbf{v}'+\mathbf{v}') = f(\mathbf{v}-\mathbf{v}')+f(\mathbf{v}') \iff f(\mathbf{v}-\mathbf{v}') = -f(\mathbf{v}') = f(-\mathbf{v}') \Longrightarrow \mathbf{v}-\mathbf{v}' = -\mathbf{v}' \iff \mathbf{v} = \mathbf{0}_{\mathrm{V}},$$

which is a contradiction. Hence, we must have $\ker f = \{\mathbf{0}_V\}$.
$\boxed{\Longleftarrow}$ Let $\ker f = \{\mathbf{0}_V\}$. Assume, by contradiction that f is not injective, i.e., there are two distinct vectors in V, such that $f(\mathbf{v}_1) = f(\mathbf{v}_2)$, with $\mathbf{v}_1 \neq \mathbf{v}_2$. Let $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2 \neq \mathbf{0}_V$. Then

$$f(\mathbf{v}) = f(\mathbf{v}_1 - \mathbf{v}_2) = f(\mathbf{v}_1) - f(\mathbf{v}_2) = \mathbf{0}_W \Longrightarrow \mathbf{v} \in \ker f,$$

which is a contradiction. Hence, f is injective.

ii. First, note that by definition of image set, im $f \subseteq W$.
$\boxed{\Longrightarrow}$ Let f be surjective. By definition, any $\mathbf{w} \in W$ satisfies $\mathbf{w} = f(\mathbf{v}) \in$ im f, so that $W \subseteq$ im f. Since im $f \subseteq W$, there holds $W =$ im f.
$\boxed{\Longleftarrow}$ Let im $f = W$. Let $\mathbf{w} \in W$. Then $\mathbf{w} \in$ im f, so that $\mathbf{w} = f(\mathbf{v})$ for some $\mathbf{v} \in V$. Since $\mathbf{w}$ is arbitrary, by definition, f is surjective. ■

Let us consider some standard examples of linear maps.

8.2 Examples

First, note two special linear maps that will arise in our later discussion.

Example 8.1 — Zero map. The **zero map** $o : V \to W$ is a linear map (check this) given by $o(\mathbf{v}) = \mathbf{0}_W$ for all $\mathbf{v} \in V$. This means that $\ker(o) = V$ and im $(o) = \{\mathbf{0}_W\}$. The map is injective if and only if $V = \{\mathbf{0}_V\}$ and surjective if and only $W = \{\mathbf{0}_W\}$.

Example 8.2 — Identity map. The **identity map** $id : V \to V$ is a linear map (check this) given by $id(\mathbf{v}) = \mathbf{v}$, for all $\mathbf{v} \in V$. Note that the map is an endomorphism, as $W = V$. The kernel of id is trivial, while im $id = V$. Hence, id is both injective and surjective.

Example 8.3 — Differentiation map. The **differentiation map** $D : \mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_{n-1}(\mathbb{R})$ is a linear map given by $(Dp)(x) = p'(x)$. It is linear since, using the properties of derivatives,

$$\left(D(ap + bq)\right)(x) = \left(ap(x) + bq(x)\right)' = ap'(x) + bq'(x) = aDp(x) + bDq(x).$$

Its kernel is non-trivial, since the constant polynomial $p(x) = a$, for $a \in \mathbb{R}$, yields $Dp(x) = p'(x) = 0$. The codomain $\mathcal{P}_{n-1}(\mathbb{R})$ is chosen to ensure surjectivity.

Example 8.4 — Integration map. The **definite integration map** $I : C^0([-1,1]) \to \mathbb{R}$ is a linear map given by $I(f) = \int_{-1}^{1} f(x)\mathrm{d}x$. It is linear since, using the properties of integrals,

$$I(af + bg) = \int_{-1}^{1} \left[af(x) + bg(x)\right] \mathrm{d}x = a\int_{-1}^{1} f(x)\mathrm{d}x + b\int_{-1}^{1} g(x)\mathrm{d}x = aI(f) + bI(g).$$

Its kernel is non-trivial, since $I(f) = 0$ for any even function f.

Example 8.5 — Matrix multiplication map. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be given by $f(\mathbf{x}) = A\mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^n$. Then f is a linear map since, using the properties of matrix-vector multiplication,

$$f(a\mathbf{x} + b\mathbf{y}) = A(a\mathbf{x} + b\mathbf{y}) = aA\mathbf{x} + bA\mathbf{y} = af(\mathbf{x}) + bf(\mathbf{y}).$$

This transformation is fundamental and will be studied in a later lecture.

Example 8.6 — Coordinate map. Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ denote a basis for V. Let

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n.$$

Define the **coordinate map** $\varphi : V \mapsto \mathbb{F}^n$ via

$$\varphi(\mathbf{v}) = \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Then φ is a linear map. To see, this, let $\mathbf{v}, \mathbf{w} \in V$ have their representations in the basis B given below

$$\mathbf{v} = \sum_{j=1}^{n} a_j\mathbf{v}_j, \quad \mathbf{w} = \sum_{j=1}^{n} b_j\mathbf{v}_j.$$

Then

$$a\mathbf{v} + b\mathbf{w} = a\sum_{j=1}^{n} a_j\mathbf{v}_j + b\sum_{j=1}^{n} b_j\mathbf{v}_j = \sum_{j=1}^{n}(aa_j + bb_j)\mathbf{v}_j =: \sum_{j=1}^{n} c_j\mathbf{v}_j.$$

Hence,

$$\varphi(a\mathbf{v} + b\mathbf{w}) = \mathbf{c} = a\mathbf{x} + b\mathbf{y} = a\varphi(\mathbf{v}) + b\varphi(\mathbf{w}).$$

This map is bijective, so by Proposition 8.2, its kernel is trivial and im $\varphi = \mathbb{R}^n$.