LECTURE 10

Linear Transformations

Recall that rather than considering general subsets of a vector space V, our focus has thus far centered on the special subsets, the so-called *subspaces of* V, that were singled out precisely because of their intrinsic compatibility with our basic vector space operations of scalar multiplication and vector addition. In a similar manner, we for the most part forgo a discussion of general functions between vector spaces and instead concentrate on a special class of functions that are intrinsically compatible with scalar multiplication and vector addition.

DEFINITION 10.1. *Let $T : V - \rightarrow W$ be a function from a vector space V to a vector space W (with the same underlying field $\mathbb{F}$). T is called a* ***linear transformation*** *if*

(i) $T(\lambda v) = \lambda T(v)$ *for all* $\lambda \in \mathbb{F}$ *and all* $v \in V$.
(ii) $T(v_1) + T(v_2)$ *for all* $v_1, v_2 \in V$

The first thing I wish to point out about a linear transformation $T : V \rightarrow W$ is that it not only relates the individual vectors in V to individual vectors in V, it also relates subspaces of V to subspaces of W.

DEFINITION 10.2. *Let $T : V \rightarrow W$ be a linear transformation and let U be a subset of V. The* ***image*** *of U by T is the subset of W is denoted $T(U)$ and is defined by*

$$T(U) = \{w \in W \mid w = T(v) \text{ for some } v \in V\} \quad .$$

PROPOSITION 10.3. *Let $T : V \rightarrow W$ be a linear transformation and let U be a subspace of V. Then $T(U)$ is a subspace of W.*

Proof. Suppose $w_1, w_2 \in T(U)$. Then, by definition, there exists $u_1, u_2 \in U$ such that $w_1 = T(u_1)$ and $w_2 = T(u_2)$. We want to show that $T(U)$ is closed under linear combinations. So consider the general linear combination $\alpha w_1 + \beta w_2$

$$\begin{aligned} \alpha w_1 + \beta w_2 &= \alpha T(u_1) + \beta T(u_2) \\ &= T(\alpha u_1) + T(\beta u_2) && \text{by (i) above} \\ &= T(\alpha u_1 + \beta u_2) && \text{by (ii) above} \end{aligned}$$

By since U is a subspace, $u_1, u_2 \in U \quad \Rightarrow \quad \alpha u_1 + \beta u_2 \in U$. Thus, $\alpha w_1 + \beta w_2 \in T(U)$, since it is the image of the vector $\alpha u_1 + \beta u_2 \in U$. □

Thus, a linear transformation $T : V \rightarrow W$ allows us to map subspaces of the domain V to the subspaces of the codomain W. We can also go in the opposite direction.

DEFINITION 10.4. *Let $T : V \rightarrow W$ be a linear transformation and let U be a subset of the codomain W. The* ***inverse image of*** *U* ***by*** *T is the subset of V denoted by $T^{-1}(U)$ and defined by*

$$T^{-1}(U) = \{v \in V \mid T(v) \in U\} \quad .$$

PROPOSITION 10.5. *Let $T : V \rightarrow W$ be a linear transformation and let U be a subset of W. Then $T^{-1}(U)$ is a subspace of V.*

Proof. Let $v_1, v_2 \in T^{-1}(U)$. Then, by definition, there exist vectors u_1 and u_2 in $U \subset W$ such that $u_1 = T(v_1)$ and $u_2 = T(v_2)$. We want to show that $\alpha v_1 + \beta v_2 \in T^{-1}(U)$. But

$$\begin{aligned} T(\alpha v_1 + \beta v_2) &= T(\alpha v_1) + T(\beta v_2) \\ &= \alpha T(v_1) + \beta T(v_2) \\ &= \alpha u_1 + \beta u_2 \\ &\in U \qquad \text{(since } u_1, u_2 \in U \text{ and } U \text{ is a subspace)} \end{aligned}$$

Thus, the image of $\alpha v_1 + \beta v_1$ is in U and so $\alpha v_1 + \beta v_2 \in T^{-1}(U)$. □

There are two especially important special cases for the subspaces $T(U) \subset W$ and $T^{-1}(U) \in V$.

Definition 10.6. *The **range** of a linear transformation $T : V \to W$ is the subspace $T(V)$ of W :*

$$\textit{range}(T) = \{w \in W \mid w = T(v) \text{ for some } v \in V\}$$

*The **kernel** of a linear transformation $T : V \to W$ is the subspace $T^{-1}(\{0_W\})$ of V :*

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

Remark 10.7. We have a bit of a notation pitfall here. Once we have a linear transformation $T : V \to W$, we also have a mapping that sends subspaces of V to subspaces of W and this is also denoted by T. Moreover, we always have a mapping T^{-1} that sends subspaces of W to subspaces of V. However, it may very well happen that this is no inverse linear transformation that sends vectors in W to vectors in V. Figuring out when T^{-1} exists as a function between W and V (as opposed to a function between subspaces of W and subspaces of V) is our next topic.

1. Digression: Functions Between Sets

Although I think it's safe to assume that everybody here is familiar with the utility of functions, it may not be the case that everybody keeps in mind the generality of this concept. So let me take a few minutes to remind us of what a function is in set theoretical terms.

Given two sets A and B a function $f : A \to B$ is a rule that links each element of A to a corresponding element of B. This situation is often represented pictorially as

It is common to say that a function $f : A \to B$ is a map from A *to* B. We also write $a \longmapsto f(a)$ and say that an element a is mapped to the element $f(a)$ of B. If $f : A \to B$ is a function from A to B

- the set A is called the domain of f
- the set B is called the codomain of f
- the set $\textit{image}(f) := \{b \in B \mid b = f(a) \text{ for some } a \in A\}$ is called the image of f

- If $b \in B$, the set $f^{-1}(b) = \{a \in A \mid f(a) = b\}$ is called the pullback (or fiber) of b.

It is important to note that if $f : A \to B$ is a function from A to B, then for each element a of A there is exactly one element $f(a)$ of B. On the other hand,

- It is not necessarily true that to each element $b \in B$ there is an $a \in A$ such that $b = f(a)$. If it does happen that for each $b \in B$ there is an $a \in A$ such that $b = f(a)$, then the function $f : A \to B$ is said to be **surjective** or **onto**. Equivalently, $f : A \to B$ is surjective if *image* $(f) = B$.
- It is not necessarily true that to each element $b \in B$ rhere is unique element a such that $b = f(a)$. If
$$\text{whenever } f(a) = f(a') \quad \text{we have} \quad a = a'$$
then the function $f : A \to B$ is called **injective** or **into** (or *one-to-one*, however, this latter terminology sometimes causes confusion with the notion of a one-to-one correspondence).
- A function that both injective and surjective is said to be **bijective**.

THEOREM 10.8. *If $f : A \to B$ is a function that is both surjective and injective, then there exists a function $f^{-1} : B \to A$ with the properties that*
$$f^{-1} \circ f(a) = a \qquad , \quad \forall\, a \in A$$
and
$$f \circ f^{-1}(b) = b \quad , \quad \forall\, b \in B\ .$$

REMARK 10.9. Although the notation is the same, a pullback $f^{-1}(b)$ of a non-injective function from A to B is not the same thing as an inverse function; because a function from B to A would have to have a unique value for each point in its domain. But sometimes $f^{-1}(b) = \{\}$ (e.g, which can happen when f is not surjective) and sometimes $f^{-1}(b) = \{a, a', \dots\}$ which can happen when f is not injective. On the other hand, when $f : A \to B$ is both injective and surjective, then f^{-1} can be identified as the inverse function of A; since in this situation for each $b \in B$ there will be exactly one element in $f^{-1}(b)$
$$\text{the pullback of } b \text{ by } f = \{a\} \quad \Longleftrightarrow \quad \text{the value of the inverse function } f^{-1} \text{ at } b \text{ is } a$$

EXAMPLES 10.10. (1) $f : \mathbb{R} \to \mathbb{R} : x \longmapsto x^2$ is neither surjective nor injective. For
$$-1 \neq x^2 \quad \text{for any } x \in \mathbb{R}$$
$$f(2) = f(-2) = 4 \qquad \text{but } 2 \neq -2 \quad .$$

(2) By restricting the codomain of the function of f in Example 1 to be $\mathbb{R}_{\geq 0}$, we can obtain a surjective function $\widetilde{f} : \mathbb{R} \to \mathbb{R}_{\geq 0} : x \longmapsto x^2$.
(3) By restricting the domain of the function of f in Example 1 to be $\mathbb{R}_{\geq 0}$, we can obtain a injective function $\overline{f} : \mathbb{R}_{\geq 0} \to \mathbb{R} : x \longmapsto x^2$.
(4) By restricting both the domain and the codomain of the function of f in Example 1 to be $\mathbb{R}_{\geq 0}$, we can obtain a bijective function $\overline{\widetilde{f}} : \mathbb{R} \to \mathbb{R}_{\geq 0} : x \longmapsto x^2$.

2. Back to Linear Transformations

THEOREM 10.11. *Suppose $T : V \to W$ is a linear transformation. Then the inverse function $T^{-1} : W \to V$ exists if and only if the following two conditions hold:*

(i) *Range* $(T) = W$
(ii) $\ker(T) = \{0_V\}$

Proof. We shall show that these two conditions on T are equivalent to T being bijective as a map between the sets V and W. That (i) implies that T is surjective is just a matter of notation:
$$W = \mathit{Range}(T) \equiv \{w \in W \mid w = f(v) \text{ for some } v \in V\} \equiv \mathit{image}(T)$$

Let me now show that (ii) is equivalent to T being injective. Suppose

$$T(v_1) = T(v_2)$$

Then

$$T(v_1) - T(v_2) = 0_W$$

But then, by linearity

$$T(v_1 - v_2) = 0_W$$

which in turn implies $v_1 - v_2 \in \ker(T)$. But by hypothesis, $\ker(T)$ consists of a single vector 0_V. So

$$v_1 - v_2 = 0_V \quad \Rightarrow \quad v_1 = v_2$$

Thus,

$\ker(T) = \{0_V\} \quad \Rightarrow \quad$ whenever $T(v_1) = T(v_2)$ then $v_1 = v_2$ and so T is injective

Let me now show the converse is true as well. I first note that

$$T(0_V) = T(v_1 - v_1) = T(v_1) - T(v_1) = 0_W$$

and so

$$0_V \in \ker(T)$$

for any linear transformation T. Now suppose T is also injective and that $v \in \ker(T)$. Then, by definition $T(v) = 0_W$, and by injectivity

$$T(0_V) = 0_W = T(v) \quad \Rightarrow \quad v = 0_V$$

thus, $\ker(T) = \{0_V\}$. □

Proposition 10.12. *Suppose $T : V \to W$ is a bijective linear transformation. Then $T^{-1} : W \to V$ is also a linear transformation.*

Proof. Let $w_1, w_2 \in W$. Since T is surjective, there exists $v_1, v_2 \in V$ such that $w_1 = T(v_1)$ and $w_2 = T(v_2)$. Now consider

$$\begin{aligned}
T^{-1}(\alpha w_1 + \beta v_2) &= T^{-1}(\alpha T(v_1) + \beta T(v_2)) \\
&= T^{-1}(T(\alpha v_1) + T(\beta v_2)) \\
&= T^{-1}(T(\alpha v_1 + \beta v_2)) \\
&= \alpha v_1 + \beta v_2 \qquad \text{since } T^{-1} \circ T = Id_V \\
&= \alpha T^{-1}(w_1) + \beta T^{-1}(w_2) \\
&\Rightarrow \quad T^{-1} \text{ is a linear transformation}
\end{aligned}$$

□

Nomenclature 10.13. A linear transformation is also called a **vector space homomorphism**. A bijective linear transformation is called a **vector space isomorphism**.