LECTURES 17: LINEAR TRANSFORMATIONS

MA1111: LINEAR ALGEBRA I, MICHAELMAS 2016

In addition to vector spaces, the other main structures of linear algebra are certain functions mapping between different vector spaces. We don't care about generic functions, but only ones which play nicely with the basic structure of vector spaces. In general, throughout your mathematical career you will find that many subjects can be studied in terms of 1). sets with "extra structure" together with 2). maps between these sets respecting this structure. Although it is tempting at the beginning to simply focus on the "objects" in 1) (which could be vector spaces, groups, fields, etc.), you will quickly find that in fact maps between them are just as fundamental as the objects themselves. As a rough idea for why this might be true, if you have an object which you don't know much about, it is often fruitful to find a function which "notices" some of the basic features of that object and sends you to a much simpler object. For example, you may consider a function mapping you to sets of matrices, which as we have seen have many nice properties.

This brief aside having been said, we make the following definition.

Definition. If V and W are vector spaces over a field F, then a function $T\colon V \to W$ (that is, a procedure taking a vector $v \in V$ and spitting out a vector $w \in W$) is called a **linear transformation** if for all $x, y \in V$, $c \in F$, we have the usual linearity properties

$$T(x + y) = T(x) + T(y),$$

$$T(cx) = cT(x).$$

For brevity, we will often just call such a function linear. I will also use words like map, transformation, and function, interchangably.

Example. *For any linear transformation T, we have $T(0) = 0$. Indeed, $T(0) = T(0 \cdot 0) = 0 \cdot T(0) = 0$.*

Example. *The most important property of derivatives which you frequently used in your calculus class is that the derivative operator D is linear. For example, we have the linear function $T\colon \mathcal{P}_{\leq n} \to \mathcal{P}_{\leq n-1}$ (where $\mathcal{P}_{\leq n}$ is the space of real-valued polynomials of degree at most n) defined by $T(f) = f'$.*

Date: November 17, 2016.

Example. *Consider the vector space $\mathcal{C}(\mathbb{R})$ of continuous functions from $\mathbb{R}$ to $\mathbb{R}$. Then for any real numbers $a < b$, we have a linear map $T\colon \mathcal{C}(\mathbb{R}) \to \mathbb{R}$ which sends f to*

$$T(f) = \int_a^b f(x)dx.$$

The usual linearity properties of integrals show that this is indeed a linear map.

Example. *For any pair of vector spaces V, W, there are two obvious, but trivial linear transformations to consider. Firstly, we have the* **identity transformation** $I_V\colon V \to V$ *defined by*

$$I_V(v) = v$$

for all $v \in V$. We also have the **zero transformation** $I_0\colon V \to W$ *defined by*

$$I_0(v) = 0$$

for all $v \in V$. Of course, this is a rather boring function, but if W is zero-dimensional, this is the only possibility, so we do have to consider it sometimes. It is not a very nice function though, as it completely forgets everything about the input vectors; this is like crumpling a piece of paper up into just one point, a very messy operation which can't be undone.

Example. *The most important example of a linear map is one which is associated to any $m \times n$ matrix with real entries. Namely, given such a matrix A, we have a function $T_A\colon \mathbb{R}^n \to \mathbb{R}^m$ which sends*

$$T_A(v) = Av.$$

That is, this map is just left-multiplication by A. The basic rules of matrix arithmetic directly show that this is indeed a linear transformation. As we saw in the last lecture, if you have a finite dimensional vector space, via coordinate vectors one may think of the vector space as a real vector space $\mathbb{R}^n$ once a basis for the original vector space is chosen. If we have any linear map, then after choosing bases for both V and W, every example is an instance of such a simple matrix multiplication transformation.

Example. *A geometric example of the construction in the last example is the linear transformation given by rotation of vectors in $\mathbb{R}^2$ by an angle ϑ. Namely, for $0 \leq \vartheta < 2\pi$, we define $T_\vartheta = T_A$ where*

$$A = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}.$$

Explicitly, $T_\vartheta(x, y) = (x\cos\vartheta - y\sin\vartheta, x\sin\vartheta + y\cos\vartheta)$, and geometrically, this corresponds to rotating the vector (x, y) by an angle ϑ (counter-clockwise of course). For instance, multiplying a vector by

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

rotates it by 90 degrees.

Example. *The linear operator (the use of the word operator signifies that it maps from a space V to V itself) $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ with*

$$T(x, y) = (x, -y)$$

reflects vectors about the x-axis.

Example. *The linear operator $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ with*

$$T(x, y) = (x, 0)$$

projects *vectors onto the x-axis. That is, given a vector (x, y), the image under this map is found by dropping a perpendicular line down to the x-axis and finding the intersection point. Thus, although this operator is a map $T\colon \mathbb{R}^2 \to \mathbb{R}^2$, it is also a linear map $T\colon \mathbb{R}^2 \to W$ where $W \subset \mathbb{R}^2$ is the x-axis, a subspace of $\mathbb{R}^2$. This kind of behavior in some ways is undesirable. For example, note that infinitely many points get sent to the same points on the x-axis, so this projection operation forgets some of the original information about the vector.*

Given any linear transformation, there are two very important associated subspaces. As you can guess from the language we have chosen, these have something to do with the vector spaces arising from matrices which we have seen before.

Definition. The **kernel** (or null space) of $T\colon V \to W$, denoted $\ker(T)$, is the set of all vectors $v \in V$ with $T(v) = 0$. The **image** (or range), denoted $\operatorname{Im}(T)$, is the set of all images of V under T, that is,

$$\operatorname{Im}(T) = \{w \in W : w = T(v) \text{ for some } v \in V\}.$$

The **nullity**, denoted null(T), is the dimension of $\ker(T)$, and the **rank**, denoted $\operatorname{rk}(T)$ is the dimension of $\operatorname{Im}(T)$.

It is easy to show that $\ker(T)$ is a subspace of V, and that $\operatorname{Im}(T)$ is a subspace of W. Indeed, the proof for the kernel is almost identical to the proof we gave for matrix kernels, and to see that the image is a subspace of W, note that our example above stated that $T(0) = 0$, so that $0 \in \operatorname{Im}(T)$, and closure under addition and scalar multiplication are direct from the linearity properties of T.

Example. *If A is a matrix, then the kernel of A is clearly the same as the kernel of T_A. Moreover, the image of T_A is the same as the column space of A. Indeed, a vector w is a linear combination of the columns of A if and only if $Ax = w$ has a solution, which is the same as saying that $T_A(x) = w$ for some x, or w is in the image of T_A.*

Example. *For any vector spaces V, W, we have*

$$\ker(I_V) = \{0\}, \qquad \operatorname{Im}(I_V) = V,$$

$$\ker(T_0) = V, \qquad \operatorname{Im}(T_0) = \{0\}.$$

Generalizing the example above, once we pick a basis for V, we can easily find a spanning set for $\mathrm{Im}(T)$.

Theorem. *If $\{v_1, \ldots, v_n\}$ is a basis for V and $T \colon V \to W$ is linear, then*

$$\mathrm{Im}(T) = \mathrm{span}\{T(v_1), \ldots T(v_n)\}.$$

Proof. It is obvious that $T(v_j) \in \mathrm{Im}(T)$ for all j. We have already shown above that $\mathrm{Im}(T)$ is a subspace, and it hence

$$\mathrm{span}\{T(v_1), \ldots T(v_n)\} \subseteq \mathrm{Im}(T).$$

To show the reverse containment, suppose that $w \in \mathrm{Im}(T)$. By assumption, we can find a $v \in V$ for which $T(v) = w$. In the given basis for V, suppose that the coordinate vector of v is c, that is,

$$v = c_1 v_1 + \ldots + c_n v_n.$$

Using the linearity of T applied to both sides of this equation, we find

$$w = T(v) = c_1 T(v_1) + \ldots + c_n(v_n),$$

and hence

$$\mathrm{Im}(T) \subseteq \mathrm{span}\{T(v_1), \ldots T(v_n)\}.$$

As we have shown both containments, we have shown that

$$\mathrm{Im}(T) = \mathrm{span}\{T(v_1), \ldots T(v_n)\},$$

as desired. □

Example. *Consider the projection operator $T \colon \mathbb{R}^2 \to \mathbb{R}^2$ with*

$$T(x, y) = (x, 0),$$

described above. Clearly, the kernel is the y-axis, and the image is the x-axis.

Example. *We will later see that every linear transformation between finite-dimensional vector spaces can be represented by matrix multiplication with a fixed matrix. Using the rank-nullity theorem that we showed before using elementary row reduction, the following fundamental result for linear transformations falls out.*

Theorem (Rank-Nullity Theorem)**.** *If $T \colon V \to W$ is linear and V is finite-dimensional, then*

$$\mathrm{null}(T) + \mathrm{rk}(T) = \dim(V).$$

Example. *Continuing the example above of the linear operator $T \colon \mathcal{P}_{\leq n} \to \mathcal{P}_{\leq n-1}$ given by*

$$T(f) = f',$$

the kernel kernel consists of those polynomials with derivative zero, namely the constant polynomials. Thus, the kernel is the subspace of polynomials of degree 0. The image is all of $\mathcal{P}_{\leq n-1}$. For example, we can take any polynomial in the latter and choose any antiderivative of it, which will then be a polynomial with degree one larger.

Example. *Let's reconsider the example* $T\colon \mathcal{C}(\mathbb{R}) \to \mathbb{R}$ *sending* f *to*

$$T(f) = \int_a^b f(x)dx.$$

Clearly, the image is all of $\mathbb{R}$*, as for any constant* c*,* $\int_a^b cdx = (b-a)c$*, so any* $\alpha \in \mathbb{R}$ *is the image* $T(f)$ *where* f *is the constant function* $\alpha/(b-a)$*. The kernel is also quite large; it consists of all those functions whose averages on the interval* $[a, b]$ *are zero.*