LINEAR ALGEBRA: LECTURE 11

LEO GOLDMAKHER

Recall that given any function $f : A \to B$, the *image* of f is the set of all outputs of f:

$$\text{im}\,(f) := \{f(a) : a \in A\}.$$

Example 1. The image of a rotation map is the entire plane:

$$\text{im}\,(R_{\pi/2}) = \mathbb{R}^2.$$

Example 2. The image of the squashing map is the horizontal axis:

$$\text{im}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} = \{(x, 0) : x \in \mathbb{R}\}.$$

The following result shows that the typical linear map has full image.

Theorem 1. im $\begin{pmatrix}a & b\\ c & d\end{pmatrix} = \mathbb{R}^2$ *if and only if* $ad - bc \neq 0$.

Proof. As usual with iff statements, we split the proof into two.

($\Rightarrow$) This follows immediately from problem **3.6(a)**.

($\Leftarrow$) Given $\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ such that $ad - bc \neq 0$, and suppose $\begin{pmatrix}p\\ q\end{pmatrix} \in \mathbb{R}^2$ is an arbitrary point. Let

$$x_0 := \frac{pd - qb}{ad - bc} \qquad \text{and} \qquad y_0 := \frac{qa - pc}{ad - bc}$$

A quick computation shows that

$$\begin{pmatrix}p\\ q\end{pmatrix} = \begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}x_0\\ y_0\end{pmatrix} \in \text{im}\begin{pmatrix}a & b\\ c & d\end{pmatrix}$$

Since $\begin{pmatrix}p\\ q\end{pmatrix}$ was arbitrary, this shows that every point in $\mathbb{R}^2$ belongs to the image of $\begin{pmatrix}a & b\\ c & d\end{pmatrix}$. □

Remark. Note that I'm *not* explaining how I came up with the crazy expressions for x_0 and y_0 – this is intentional. The way in which one came up with the proof is not part of the proof itself, just as the way in which an author came up with the idea of a short story is not described in the story itself. (In class we saw how to come up with these choices of x_0 and y_0; I urge you to make sure you can reconstruct our work.)

Combining the above theorem with problem **3.6**, we see that the image of any linear map is either all of $\mathbb{R}^2$ or is entirely contained in some line – there's no middle ground. Thus, for example, there are no linear maps whose image is a parabola, or some region of the plane.

Given four random real numbers a, b, c, d, we don't expect to have $ad - bc = 0$. Thus, our theorem shows that a linear map $f : \mathbb{R}^2 \to \mathbb{R}^2$ selected at random will almost always have im $(f) = \mathbb{R}^2$. This explains the choice of terminology in the following definition:

Definition. A linear map $f : \mathbb{R}^2 \to \mathbb{R}^2$ is said to be *singular* iff im $(f) \neq \mathbb{R}^2$. Otherwise, f is called *nonsingular*.

Our next goal is to prove the following.

Theorem 2. *If $f : \mathbb{R}^2 \to \mathbb{R}^2$ is a nonsingular linear map, then f is invertible, and $f^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ is linear.*

Date: March 2, 2016.

Before we can prove this theorem, we need to define one of the main characters:

Definition. Given a function $f : A \to B$ and any element $b \in B$, we define the *pre-image of* b *under* f to be

$$f^{-1}(b) := \{a \in A : f(a) = b\}.$$

Example 3. Consider the familiar function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto x^2 \end{aligned}$$

(The bottom arrow is a convenient notation for expressing what the function does to a given element x.) What can we say about preimages of different numbers under f? Here are a few examples.

- $f^{-1}(4) = \{2, -2\}$
- $f^{-1}(0) = \{0\}$
- $f^{-1}(-1) = \emptyset$ (the symbol $\emptyset$ denotes the empty set)

Note that the output of f^{-1} is a *set*. Thus, generally speaking, f^{-1} isn't a function – given an input, there might be multiple possible outputs to choose from, or there might be no outputs to choose from. We explore this further with the following exercise:

Exercise 1. Given $f : A \to B$, what can you say about $f^{-1}\Big(f(x)\Big)$? How about $f\Big(f^{-1}(x)\Big)$?

Let's start with $f^{-1}\Big(f(x)\Big)$. By definition,

$$f^{-1}\Big(f(x)\Big) = \{a \in A : f(a) = f(x)\}.$$

Without more information about f, we can't say much about this set. However, there is one element of A which we know must belong to this set:

$$x \in f^{-1}\Big(f(x)\Big).$$

(Make sure you can explain why!)

Next we turn to $f\Big(f^{-1}(x)\Big)$. We are applying the function f to $f^{-1}(x)$, a set. How does one apply a function to a set? Just apply f to each element of the set! In other words, given a set S, we define

$$f(S) := \{f(s) : s \in S\}.$$

Thus,

$$f\Big(f^{-1}(x)\Big) = \{f(a) : a \in f^{-1}(x)\} = \begin{cases} \{x\} & \text{if } f^{-1}(x) \neq \emptyset \\ \emptyset & \text{otherwise.} \end{cases}$$

To make this look nicer, we will abuse notation and adopt the following convention:

We shall treat any set consisting of a single element as that element itself.

Thus, for example, we will not distinguish between the set $\{5\}$ and the number 5. This convention allows us to write

$$f\Big(f^{-1}(x)\Big) = x \qquad \forall x \in \text{im}\,(f).$$

It also allows us to define the concept of the inverse of a function.

Definition. Given $f : A \to B$. We say f is *invertible* if

$$\#f^{-1}(b) = 1 \quad \forall b \in B,$$

where $\#S$ denotes the number of distinct elements in the set S. If this is the case, we view f^{-1} as a function $f^{-1} : B \to A$ by using our convention above.

Next time we will discuss this more and prove Theorem 2.