LINEAR ALGEBRA: LECTURE 16

LEO GOLDMAKHER

We started by reviewing the types of objects we've seen so far:

- $\mathbb{R}$, the set of real numbers, measure displacement;
- $\mathbb{R}^2$, the set of points in the plane, indicate location in the plane; and
- V^2, the set of all 2-dimensional vectors, indicate directions of motion.

Last time we showed that any linear map $f : \mathbb{R}^2 \to \mathbb{R}^2$ is also a well-defined when viewed as a function mapping $V^2 \to V^2$. More explicitly, given any vector $\alpha \vec{e}_1 + \beta \vec{e}_2 \in V^2$ we define $f(\alpha \vec{e}_1 + \beta \vec{e}_2)$ as follows:

$$f(\alpha \vec{e}_1 + \beta \vec{e}_2) := p\vec{e}_1 + q\vec{e}_2 \qquad \Longleftrightarrow \qquad f(\alpha, \beta) = (p, q).$$

Note that we could attempt to make such a definition for any map $f : \mathbb{R}^2 \to \mathbb{R}^2$, but for non-linear maps it might not end up being well-defined. For example, last time we defined $h : \mathbb{R}^2 \to \mathbb{R}^2$ by $h(x, y) := (x, |y|)$. Attempting to carry out the above transition to viewing $h : V^2 \to V^2$ led to a contradiction: $h(\vec{e}_1 + \vec{e}_2)$ sometimes was equal to $\vec{e}_1 + \vec{e}_2$, and other times to $\vec{e}_1 - \vec{e}_2$. This discrepancy isn't surprising – vectors and points are entirely different types of objects, so we shouldn't expect functions of points to also be functions of vectors. The surprise is that linear maps f can be viewed in either universe. Not only is this an unexpected fact, it's also useful because it gives us a way to do linear algebra *without coordinates*. Begin coordinate-free means that linear algebra is universal – you can describe results over the phone to someone and not worry whether they've drawn their axes in the same place as you.

Incidentally, linear maps, when viewed as functions from $V^2 \to V^2$, also obey additivity and scale. In other words,

$$f(\vec{v} + \vec{w}) = f(\vec{v}) + f(\vec{w}) \quad \forall \vec{v}, \vec{w} \in V^2 \qquad \text{and} \qquad f(\alpha \vec{v}) = \alpha f(\vec{v}) \quad \forall \alpha \in \mathbb{R}, \vec{v} \in V^2.$$

Here $\vec{v} + \vec{w}$ means you first follow the set of instructions given by $\vec{v}$, then those given by $\vec{w}$. For example, if $\vec{v} = 2\vec{e}_1 - 3\vec{e}_2$ and $\vec{w} = -\vec{e}_1 + 5\vec{e}_2$, then $\vec{v} + \vec{w}$ is the instruction "move 2 units to the right, then 3 units down, then 1 unit left, then 5 units up"; this is more succinctly expressed as "move 1 unit right and 2 units up", i.e., $\vec{v} + \vec{w} = \vec{e}_1 + 2\vec{e}_2$. Similarly, $\alpha \vec{v}$ simply rescales all the units by a factor of α.

Having discussed all this, we turned to our main subject of interest: change of basis. Recall that the goal of this is as follows. Suppose a traveler from the distant land of Shibboleth is visiting Williamstown. While we give directions in terms of the basic motions 1 unit east and 1 unit north (i.e., $\vec{e}_1$ and $\vec{e}_2$), the Shibbolese use two other basic motions to express directions: $\vec{v}_1$ and $\vec{v}_2$. Here's a picture of the situation (note that since we're dealing with vectors, we don't have coordinate axes!):
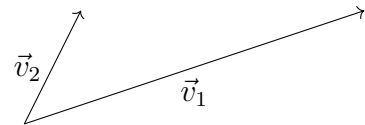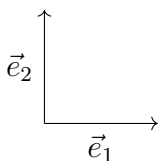
In the left plane, vectors are described in terms of $\vec{e}_1, \vec{e}_2$; in the right plane, in terms of $\vec{v}_1, \vec{v}_2$

We wish to be able to give directions to the Shibbolese, but to do so, we need to be able to express any directions in terms of their basic $\vec{v}_1$ and $\vec{v}_2$. Can this be done? And how?

Our approach (inspired by the example from last time) is to try to express $\vec{e}_1$ and $\vec{e}_2$ in terms of $\vec{v}_1$ and $\vec{v}_2$. What do we know about $\vec{v}_1$ and $\vec{v}_2$? Well, not much. All we can say is that since they're vectors, they can be

---

written in terms of the basic directions, say,

$$\vec{v}_1 = a\vec{e}_1 + b\vec{e}_2$$
$$\vec{v}_2 = c\vec{e}_1 + d\vec{e}_2.$$

Next we try to manipulate these relations to express $\vec{e}_1, \vec{e}_2$ in terms of $\vec{v}_1$ and $\vec{v}_2$. We observed that the above two imply

$$-c\vec{v}_1 + a\vec{v}_2 = (ad - bc)\vec{e}_2.$$

Similarly, we have

$$d\vec{v}_1 - b\vec{v}_2 = (ad - bc)\vec{e}_1.$$

Simplifying both of these gives

$$\vec{e}_1 = \frac{d}{ad - bc}\vec{v}_1 + \frac{-b}{ad - bc}\vec{v}_2$$
$$\vec{e}_2 = \frac{-c}{ad - bc}\vec{v}_1 + \frac{a}{ad - bc}\vec{v}_2. \tag{1}$$

Having done this, it's now a simple matter to express any vector $\vec{v}$ as a linear combination of $\vec{v}_1$ and $\vec{v}_2$: just write $\vec{v}$ as a linear combination of $\vec{e}_1$ and $\vec{e}_2$, then plug into the relations (1) and simplify.

All of the above looks strikingly similar to work we've done with linear maps. In particular, the coefficients in the equations (1) are exactly the entries of the inverse of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. This is a bit weird: the original question was just about writing vectors in terms of other vectors, and had nothing to do with matrices. So what's the connection?

Let's try to evaluate what we've done in terms of linear functions. Let

$$f := \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

We have $f(1,0) = (a,b)$ and $f(0,1) = (c,d)$. If instead of viewing $f : \mathbb{R}^2 \to \mathbb{R}^2$ we view f as the map from $V^2 \to V^2$, this would read

$$f(\vec{e}_1) = \vec{v}_1$$
$$f(\vec{e}_2) = \vec{v}_2$$

(Make sure you understand where this came from.)

Now, we want to express a given vector $\vec{v}$ in terms of $\vec{v}_1$ and $\vec{v}_2$, so we think of $\vec{v}$ as living in the right plane. There's a natural vector associated to $\vec{v}$ in the left plane: $f^{-1}(\vec{v})$.

English   *Shibbolese*
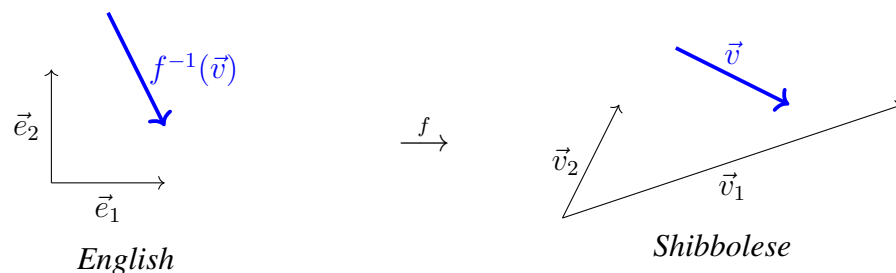
I like to think of the situation in terms of translations between languages. Imagine that the left plane represents the English language, while the right plane represents Shibbolese (the national language of Shibboleth). The linear map f is an English-Shibbolese dictionary. We're given some word $\vec{v}$ in Shibbolese, and we want to explain it in Shibbolese. The problem is, we don't speak Shibbolese! Here's what we can do. First, we translate $\vec{v}$ into English. Now we can explain it in English, since we speak the language. Finally, we translate our English explanation, word-for-word, back into Shibbolese. Let's interpret all this mathematically.

We start with $\vec{v}$, which we wish to express as a linear combination of $\vec{v}_1, \vec{v}_2$. We translate this into English by applying f^{-1}. Now we have a vector $f^{-1}(\vec{v})$ in the familiar world in which everything is described in terms of simple directions $\vec{e}_1, \vec{e}_2$; say

$$f^{-1}(\vec{v}) = \alpha\vec{e}_1 + \beta\vec{e}_2.$$

Now we translate back into Shibbolese: applying f, we find

$$\vec{v} = f(\alpha\vec{e}_1 + \beta\vec{e}_2).$$

Since f is linear, we can simplify this expression:

$$\vec{v} = \alpha f(\vec{e}_1) + \beta f(\vec{e}_2)$$
$$= \alpha\vec{v}_1 + \beta\vec{v}_2.$$

We have therefore found a linear combination of $\vec{v}_1, \vec{v}_2$ which forms $\vec{v}$! We've proved:

Proposition 1. *Given $\vec{v}_1, \vec{v}_2 \in V^2$, let $f : V^2 \to V^2$ be the linear map such that $f(\vec{e}_1) = \vec{v}_1$ and $f(\vec{e}_2) = \vec{v}_2$. If f is nonsingular, then every vector is a linear combination of $\vec{v}_1, \vec{v}_2$. In other words, given any $\vec{v} \in V^2$, there exist $\alpha, \beta \in \mathbb{R}$ such that*

$$\vec{v} = \alpha\vec{v}_1 + \beta\vec{v}_2.$$

Thus, so long as the English-Shibbolese dictionary $f : V^2 \to V^2$ is nonsingular, we can express every vector in Shibbolese. What if f were singular? We take this up next class.