23 2025-10-17 | Week 08 | Lecture 23

> ***The nexus question of this lecture:*** *Why is matrix multiplication defined the way it is?*

This lecture is based on section 5.2

23.1 Combining linear transformations: the geometric perspective

A ***linear transformation*** is a function $T : V \to W$ given by with the property that

$$T(av + bv') = aT(v) + bT(v')$$

for any $v, v' \in V$ and $a, b \in \mathbb{R}$.

If $V = \mathbb{R}^2$ then linear transformations include things like

- rotations about the origin
- reflections across any line of the form $y = mx$
- projections of the plane onto a line of the form $y = mx$
- uniform stretching of space
- any shearing of the plane (i.e., which slants the plane)
- any combinations of these

(The story is similar for higher dimensions, but we get some weirder transformations: for example in $\mathbb{R}^4$, one can rotate space about two axes at once.)

The last bullet—that we can combine linear transformations—is significant. In math-speak, let

$$S, R : \mathbb{R}^2 \to \mathbb{R}^2$$

be linear transformations. If we do S and then do R, is the resulting transformation of space is the function composition $R \circ S$.

Example 107. For example, the following matrix rotates space clockwise about the origin by 30° (i.e., $\pi/6$ radians):

$$R = \begin{bmatrix} \cos(\pi/6) & \sin(\pi/6) \\ -\sin(\pi/6) & \cos(\pi/6) \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$$

and the following matrix stretches space in the x-direction by a factor of 2:

$$S = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

Then rotating and then stretching is achieved by doing S and then R. That is, by $R \circ S$. I claim this is given by the matrix product

$$RS = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & \frac{1}{2} \\ -1 & \frac{\sqrt{3}}{2} \end{bmatrix}$$

The core idea of this lecture is that matrix multiplication is defined in such a way that this always happens: i.e., that the product RS corresponds to doing transformation S and then doing transformation T.

In other words, the idea of composing linear transformations comes first. The reason why people defined matrix multiplication the way it is, is because they wanted it to represent composition of linear functions. We make this idea more precise in the next section.

End of Example 107. □

23.2 Matrix multiplication is composition of linear transformations

Recall that if $f : X \to Y$ and $g : Y \to Z$ are functions, we define the copmosition $g \circ f$ as

$$g \circ f(x) := g(f(x)),\ x \in X.$$

Function composition can be thought of a form of "multiplication" (sort of).

Theorem 108 (Composition of linear transformations)**.** *If $T : V \to W$ and $S : W \to U$ are linear transformations, then the composite function $ST = S \circ T$ is a linear transformation $ST : V \to U$ is a linear transformation.*

Claim 1: Every linear transformation can be represented by a matrix.

Proof of Claim 1. This is the subject of the next lecture. $\square_{\text{Claim}}$

Theorem 109. *When we compose linear transformations T and S, this corresponds to multiplying their matrices.*

We will prove the case of 2×2 matrices. For a proof of the general case (which holds for any matrices), see Theorem 5.10 in the textbook (p 258).

Proof. Suppose $T, S : \mathbb{R}^2 \to \mathbb{R}^2$ are linear transformations. By Claim 1 both T and S can be represented by 2×2 matrices. That is, there exist $a, b, c, d, A, B, C, D \in \mathbb{R}$ such that

$$T \begin{bmatrix} x \\ y \end{bmatrix} := \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix} \quad \text{and} \quad S \begin{bmatrix} x \\ y \end{bmatrix} := \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} Ax + By \\ Cx + Dy \end{bmatrix}.$$

The matrix product is

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \underbrace{\begin{bmatrix} Aa + Bc & Ab + Bd \\ Ca + Dc & Cb + Dd \end{bmatrix}}_{\text{call this } M}$$

To show that $S \circ T$ corresponds exactly to matrix multiplication, we need to show that

$$S \circ T \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = M \begin{bmatrix} x \\ y \end{bmatrix} \tag{22}$$

for all $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$. This is because left-hand side of the aboe equation is the composition $S \circ T$. The right hand side is the product of the matrices for S and T. By showing that the equality holds, we will show that they are the same thing.

Indeed,

$$\begin{aligned} S \circ T \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) &= S \left(T \begin{bmatrix} x \\ y \end{bmatrix} \right) \\ &= S \left(\begin{bmatrix} ax + by \\ cx + dy \end{bmatrix} \right) \\ &= \begin{bmatrix} A(ax + by) + B(cx + dy) \\ C(ax + by) + D(cx + dy) \end{bmatrix} \\ &= \begin{bmatrix} (Aa + Bc)x + (Ab + Bd)y \\ (Ca + Dc)x + (Cb + Dd)y \end{bmatrix} \\ &= M \begin{bmatrix} x \\ y \end{bmatrix}. \end{aligned}$$

We have demonstrated that Eq. (22) holds. We are done. :) $\square$