Notes for 18.02 Recitation 3

18.02 Recitation MW9

Evan Chen

11 September 2024

Every human has regrets, has things they'd like to go back and change. But I don't! 'Cause I'm a bear!

— *Monokuma in* Danganropa

This handout (and any other DLC's I write) are posted at `https://web.evanchen.cc/1802.html`.

§1 Housekeeping

- **Office hours both Monday and Wednesday 10am**; rooms posted on Evan's site.
- **You can now check recitation answers on Evan's site.** After the disaster that was the end of R02, I have pre-uploaded a quick answer key to R03 this time.[1] So, now you can check your answer on your phone/tablet/laptop/etc. But still feel free to check answers with me the old fashioned-way too. (Figured it's better to give people more options.)

I'll try to also yap less today and give you time to actually talk to people more 1:1 as I move around, since that seemed more useful. (Though Wed recitation is nicer than Mon, half as much stuff to cover.)

§2 Follow-ups from R02

- Errata from R02: Last week in R02 I mistakenly claimed that $\overrightarrow{AB} \times \overrightarrow{AC}$ is the area of the triangle ABC, but it's actually the parallelogram's area. Divide by 2 to get the area of the triangle instead.
- Maulik says any drawing of x, y, z axes satisfying right-hand rule is OK. I'll follow Maulik.
- Remember that every vector has both a **direction** and a **magnitude**, which both have geometric meaning (or insignificance); see table below.

Object	Direction	Magnitude
Normal vector to plane	Perpendicular to plane	*Irrelevant!*
Vector projection of $\mathbf{v}$ to $\mathbf{w}$	Same as $\mathbf{w}$	Scalar projection
Cross product $\mathbf{a} \times \mathbf{b}$	Perpendicular to both $\mathbf{a}$ and $\mathbf{b}$	Area of parallelogram

- Remember: **the $\mathbb{R}^2$ concept of slope is overridden by the concept of normal vector for** $\mathbb{R}^3$. This is an important lesson from last time that I didn't call enough attention to, so I want to re-emphasize it. Anything you used to rely on slopes for, you should try to retrain yourself to think of in terms of normal vector. (I'll say this yet again when $\nabla f(x, y)$ is introduced.)

[1] This time I just used ChatGPT. For future recitations, I might just type answers only, continue using ChatGPT more, or something else; undecided. Anyway, ChatGPT can make mistakes, so if you think something's wrong, flag me.

§3 Some abstraction on linear maps

Matrices are super confusing because they're actually the "wrong" way to think about things. Again, take 18.700 or 18.701 or read my Napkin book, or come to office hours. For now, a summary:

- A *linear transform* $T : \mathbb{R}^n \to \mathbb{R}^m$ is *any* map obeying $T(c\mathbf{v}) = cT(\mathbf{v})$ and $T(\mathbf{v}+\mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$. It's a chonky boy: for every $v \in \mathbb{R}^n$, there's an output value $T(\mathbf{v}) \in \mathbb{R}^m$.
- A *matrix* is a way of *encoding* the *outputs* of T using as few numbers as possible.
- Pop quiz: if $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transform and it's given that

$$T\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right) = \begin{pmatrix} \pi \\ 9 \end{pmatrix} \text{ and } T\left(\begin{pmatrix} 100 \\ 100 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 12 \end{pmatrix}$$

  what are the vectors for $T\left(\begin{pmatrix} 103 \\ 104 \end{pmatrix}\right)$ and $T\left(\begin{pmatrix} 203 \\ 204 \end{pmatrix}\right)$?
- More generally, if you know the outputs $T\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$ and $T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$, then that gives you enough information to figure out all other outputs of T, because

$$T\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = aT\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + bT\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right). \tag{1}$$

So the really, really important principle to understand is:

Idea

A matrix **encodes all outputs** of a linear transformation T by **writing the outputs** of $T(\mathbf{e}_1)$, …, $T(\mathbf{e}_n)$ as a list of **column vectors**.

There's a philosophical axe I'm grinding. In 18.02, a linear transform is defined as "act by a matrix". I've never liked this definition for ten years. I would rather define it as in the first bullet above (as in 18.701); then say the matrix is a way to succinctly *encode* or *identify* the transformation. (Think how that every human has a name that identifies them.)

If that made sense, you can explain the following two sentences (the last page of Maulik's L3 notes):

- Multiplication of matrix by vector is defined so that $M\mathbf{v} = T(\mathbf{v})$. (This is Equation 1.)
- Multiplication of two matrices defined so that AB corresponds to *function composition*.

§4 Recitation problems from official course

1 Consider the 2×2 matrices $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} -3 & 1 \\ 2 & 0 \end{pmatrix}$.

1a Calculate AB and BA.

1b Calculate the matrix A^{-1}.

1c Check directly that $AA^{-1} = A^{-1}A = I$.

1d Consider the linear system of equations: $2x + y = 7$ and $x + 2y = 11$. Write this system of equations as a single matrix equation and solve it using A^{-1} from the previous question. Check that your solution satisfies the original equations.

2a Viewed as a linear transformation, how does the matrix $\begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$ transform the unit square with vertices $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$?

2b What is the 2×2 matrix A associated with reflection across the line $y = -x$?

2c Find the 3×3 matrix corresponding to the linear transformation f: $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 3x-2z \\ x+y+z \\ 4y+z \end{pmatrix}$.

2d Consider the linear transformation $f : \mathbb{R}^2 \to \mathbb{R}^2$ where $f(\mathbf{v})$ is the vector projection of $\mathbf{v}$ in the direction $\mathbf{w} = \binom{1}{2}$. What is the 2×2 matrix A associated with this linear transformation?

§5 Post-recitation notes

If you ~~overslept~~ couldn't make it to recitation, here's an elaboration of Section 3 as I presented it today.

§5.1 One matrix

Here's the answer to the pop quiz in Section 3 (I'll restate the question first):

Problem 5.1: If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transform and it's given that

$$T\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right) = \begin{pmatrix} \pi \\ 9 \end{pmatrix} \text{ and } T\left(\begin{pmatrix} 100 \\ 100 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 12 \end{pmatrix}$$

what are the vectors for $T\left(\binom{103}{104}\right)$ and $T\left(\binom{203}{204}\right)$?

Solution

$$\begin{aligned} T\left(\begin{pmatrix} 103 \\ 104 \end{pmatrix}\right) &= \begin{pmatrix} \pi \\ 9 \end{pmatrix} + \begin{pmatrix} 0 \\ 12 \end{pmatrix} = \begin{pmatrix} \pi \\ 21 \end{pmatrix} \\ T\left(\begin{pmatrix} 203 \\ 204 \end{pmatrix}\right) &= \begin{pmatrix} \pi \\ 9 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 12 \end{pmatrix} = \begin{pmatrix} \pi \\ 33 \end{pmatrix}. \end{aligned}$$

□

Now more generally, here's the example with the made-up numbers (I forget which random numbers the audience gave me for the actual recitation, sorry).

Problem 5.2: If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transform and it's given that

$$T\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \text{ and } T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

what is $T\left(\binom{50}{70}\right)$?

Solution

$$T\left(\begin{pmatrix} 50 \\ 70 \end{pmatrix}\right) = 50\begin{pmatrix} 1 \\ 3 \end{pmatrix} + 70\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 190 \\ 430 \end{pmatrix}.$$

More generally, the second question shows that if you know $T\left(\binom{1}{0}\right)$ and $T\left(\binom{0}{1}\right)$ you ought to be able to *calculate* the output of T at any other vector like $\binom{50}{70}$. More generally, if $T : \mathbb{R}^n \to \mathbb{R}^m$, then telling you the output of $T(\mathbf{e}_1)$, ..., $T(\mathbf{e}_n)$ is the smallest amount of information I can give you that lets you reconstruct any other output.

Now, I told you a linear transformation T can be *encoded* as a matrix. This is really easy: glue the outputs of T at the basis vectors and format it as an array:

$$T\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}1\\3\end{pmatrix} \text{ and } T\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}2\\4\end{pmatrix} \iff T \quad \text{encoded as} \quad \begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}$$

The matrix multiplication rule is then rigged to correspond to evaluation:

$$T\left(\begin{pmatrix}50\\70\end{pmatrix}\right) = \begin{pmatrix}190\\430\end{pmatrix} \iff \begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}\begin{pmatrix}50\\70\end{pmatrix} = \begin{pmatrix}190\\430\end{pmatrix}.$$

And indeed, you can now verify that if you calculate $\begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}\begin{pmatrix}50\\70\end{pmatrix}$ as you were told to do in high school, you get the same answer (and do the same calculations) as we did for Problem 5.2. (In recitation, I actually circled where $1 \cdot 50$ and $2 \cdot 70$ showed up in the solution to Problem 5.2 on the board.)

Example: the identity matrix deserves its name

This also gives a "better" reason why the identity matrix is the one with 1's on the diagonal than the "well try multiplying by it".

Let $I : \mathbb{R}^3 \to \mathbb{R}^3$ denote the 3d identity function. To encode it, we look at its values at $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$:

$$I(\mathbf{e}_1) = \mathbf{e}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad I(\mathbf{e}_2) = \mathbf{e}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad I(\mathbf{e}_3) = \mathbf{e}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}.$$

We encode it as a matrix by writing the columns side by side, getting what you expect:

$$I \quad \text{encoded as} \quad \begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{pmatrix}.$$

§5.2 Two matrices

Any time we have functions in math, we can *compose* them.[2] So let's play the same game with a pair of functions S and T, and think about their composition $S \circ T$.

Problem 5.3: Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transform such that

$$T\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}1\\3\end{pmatrix} \text{ and } T\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}2\\4\end{pmatrix}.$$

Then let $S : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transform such that

$$S\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}5\\7\end{pmatrix} \text{ and } S\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}6\\8\end{pmatrix}.$$

Evaluate $S\left(T\left(\begin{pmatrix}1\\0\end{pmatrix}\right)\right)$ and $S\left(T\left(\begin{pmatrix}0\\1\end{pmatrix}\right)\right)$.

[2] The $\circ$ symbol means the function where you apply T first then S first. So for example, if $f(x) = x^2$ and $g(x) = x + 5$, then $(f \circ g)(x) = f(g(x)) = (x+5)^2$. We mostly use that circle symbol if we want to refer to $f \circ g$ itself without the x, since it would look really bad if you wrote "$f($g" or something.

Solution

$$S\left(T\left(\begin{pmatrix}1\\0\end{pmatrix}\right)\right) = S\left(\begin{pmatrix}1\\3\end{pmatrix}\right) = 1\begin{pmatrix}5\\7\end{pmatrix} + 3\begin{pmatrix}6\\8\end{pmatrix} = \begin{pmatrix}23\\31\end{pmatrix}$$
$$S\left(T\left(\begin{pmatrix}0\\1\end{pmatrix}\right)\right) = S\left(\begin{pmatrix}2\\4\end{pmatrix}\right) = 2\begin{pmatrix}5\\7\end{pmatrix} + 4\begin{pmatrix}6\\8\end{pmatrix} = \begin{pmatrix}34\\46\end{pmatrix}.$$

Now, $S \circ T$ is *itself* a function, so it makes sense to encode $S \circ T$ as a matrix too, using the answer to Problem 5.3:

$$S\left(T\left(\begin{pmatrix}1\\0\end{pmatrix}\right)\right) = \begin{pmatrix}23\\31\end{pmatrix} \text{ and } S\left(T\left(\begin{pmatrix}0\\1\end{pmatrix}\right)\right) = \begin{pmatrix}34\\46\end{pmatrix} \iff S \circ T \text{ encoded as } \begin{pmatrix}23 & 34\\31 & 46\end{pmatrix}.$$

The matrix multiplication rule is then rigged to give the same answer through the same calculation again:

$$\begin{pmatrix}5 & 6\\7 & 8\end{pmatrix}\begin{pmatrix}1 & 2\\3 & 4\end{pmatrix} = \begin{pmatrix}23 & 34\\31 & 46\end{pmatrix}.$$

This shows why the 18.700/18.701 definitions are better than the 18.02 ones. In 18.02, the recipe for matrix multiplication is a *definition*: "here is this contrived rule about taking products of columns and rows, trust me bro". But in 18.700/18.701, the matrix multiplication recipe is a *theorem*; it's what happens if you generalize Problem 5.3 to eight variables (or $n^2 + n^2 = 2n^2$ variables for $n \times n$ matrices).

As an aside, this should explain why matrix multiplication is associative but not commutative:

- Because function composition is associative, so is matrix multiplication.
- Because function composition is *not* commutative in general, matrix multiplication isn't either.

§5.3 Remark on recitation question 2

I want to point out that 2b/2c/2d are all doing the same thing: they take the general shape

> *Here is a particular linear transformation described in words; please encode it as a matrix.*

And you do all three problems by calculating the value of the transformation at $\mathbf{e}_i$, and then encoding it by just gluing them together. Conceptually, you can fit those into the following table. It's important to realize **all the work of the problem is the "values at basis"**; and that **only uses up to R02 material**. The only new step introduced in R03 is "to encode as matrix, glue your answers together".

Q	Transf.	Values at basis	Encoded matrix
2b	$T =$ "Reflect around $y = -x$"	$T(\mathbf{e}_1) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ $T(\mathbf{e}_2) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$
2c	$T =$ "ugly 3d equation above"	$T(\mathbf{e}_1) = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$ $T(\mathbf{e}_2) = \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}$ $T(\mathbf{e}_3) = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 3 & 0 & -2 \\ 1 & 1 & 1 \\ 0 & 4 & 1 \end{pmatrix}$
2d	$T =$ "vector projection as in R02"	$T(\mathbf{e}_1) = \begin{pmatrix} \frac{1}{5} \\ \frac{2}{5} \end{pmatrix}$ $T(\mathbf{e}_2) = \begin{pmatrix} \frac{2}{5} \\ \frac{4}{5} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{pmatrix}$

§6 TODO

Note to self: write up the following things.

1. Rotation matrix example (this seems to give a lot of people trouble).
2. Analogous table with "is zero" and "is not zero" for numbers (corr to vectors above).