

Notes for 18.02 Recitation 5

18.02 Recitation MW9

EVAN CHEN

18 September 2024

The most competent people, with weak processes, will screw up.

— Petey

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

Also, I'm going to start updating the online PDF's with **post-recitation notes** based on what we talk about during the recitation. So check the PDF later. (They're up already for R03 and R04.)

§1 How to find eigen-things of an $n \times n$ matrix A

This is covered in full detail in Poonen section 6. Here is a tl;dr, but please read Poonen. (Unlike Poonen, I prefer $\det(A - \lambda I)$ to $\det(\lambda I - A)$ due to fewer minus signs. This makes no difference.)

1. To find the **eigenvalues**, solve the equation $\det(A - \lambda I) = 0$ for λ .
 - For $n = 2$ it turns out that this equation is actually just

$$\lambda^2 - (\text{Tr } A)\lambda + \det A = 0$$

so you can take a shortcut specifically when $n = 2$.

- In real life, half the time this quadratic equation will have non-real roots. I hope this doesn't happen in 18.02, though, but I'll give you one example in a moment for your curiosity.
2. To find the **eigenvectors**, for *each* eigenvalue λ , manually solve the system of n equations

$$(A - \lambda I) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$$

for x_1, \dots, x_n . (For 18.02, usually $n \leq 3$.)

- By the last step, $A - \lambda I$ is *not* invertible, (i.e. $\det(A - \lambda I) = 0$) so this system of n equations is always degenerate. It will always have infinitely many solutions, i.e. at least one redundant equation. (To tie into last recitation: it has at least one obvious solution, $x_1 = \dots = x_n = 0$, and because the matrix of coefficients $A - \lambda I$ is not invertible, it thus has infinitely many.)

§1.1 Example with non-real eigenvalues (not for exam)

Let $M = \begin{pmatrix} \cos(60^\circ) & -\sin(60^\circ) \\ \sin(60^\circ) & \cos(60^\circ) \end{pmatrix}$ be the matrix corresponding to rotation by 60 degrees. (Feel free to replace 60 by a different number.) I claim that M has no real eigenvalues or eigenvectors.

Indeed, if $\mathbf{v} \in \mathbb{R}^2$ was an eigenvector, then $M\mathbf{v}$ needs to point in the same direction as \mathbf{v} , by definition. But that can never happen: M is rotation by 60° , so $M\mathbf{v}$ and \mathbf{v} necessarily point in different directions — 60 degrees apart.

If you want, you can compute the characteristic polynomial for M . Indeed, $\text{Tr}(M) = 2 \cos(60^\circ) = 1$ and $\det(M) = \cos^2(60^\circ) + \sin^2(60^\circ) = 1$. So the characteristic polynomial is $\lambda^2 - \lambda + 1 = 0$, which has two nonreal roots $\frac{1 \pm \sqrt{-3}}{2}$.

§2 An example showing a use for eigenvectors (not for exam)

Problem

Let M be the matrix $\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$. Calculate M^{100} .

This looks impossible at first glance without a computer. I'll show you how to do it with eigenvectors.

Solution

In R03 and R04, I repeatedly stress the idea if you have a linear operator T , and you know the outputs of T on *any* basis, that tells you all the outputs of T .

Idea: we'll find the eigenvectors for M , and pray they're independent (they usually are). Since $\det M = 5$ and $\text{Tr } M = 6$, the characteristic polynomial is $\lambda^2 - 5\lambda + 6 = 0$, so the eigenvalues are 2 and 3. If you grind out the corresponding system, you'll find these eigenvectors:

- The vector $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ is an eigenvector with eigenvalue 2, because $M \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.
- The vector $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is an eigenvector with eigenvalue 3, because $M \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Then it's really easy to apply M^{100} because it's just multiplication by a constant:

$$M^{100} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 2^{100} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}; \quad M^{100} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 3^{100} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

So now we know the outputs of M^{100} at two linearly independent vectors. We can now rewrite this as

$$M^{100} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2^{100} \\ 0 \\ 2^{100} \end{pmatrix}; \quad M^{100} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = M^{100} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - M^{100} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3^{100} - 2^{100} \\ 3^{100} \\ 3^{100} \end{pmatrix}.$$

Thus encoding M as in R03 gives the answer: $M^{100} = \begin{pmatrix} 2^{100} & 3^{100} - 2^{100} \\ 0 & 3^{100} \end{pmatrix}$.

Remark

A natural question might be whether every 2×2 matrix (or any $n \times n$ matrix) has such a basis consisting of only eigenvectors, i.e. whether you could do this trick with any matrix. It turns out (I won't prove this) that it's always possible for $n \times n$ matrices as long as the characteristic polynomial has no repeated roots, although the eigenvalues might end up being complex numbers.

§3 Recitation problems from official course

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & 0 & 1 \\ 3 & 4 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

- 1 Find the eigenvalues and eigenvectors of A and B .
- 2 Find the eigenvalues and eigenvectors of C .

- 3 For any real numbers a and b , find the eigenvalues and eigenvectors of $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ and $B = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$.
If $a = b$, how do these answers differ?

§4 Post-recitation notes: let's redo everything

Based on my pulse today, I want to redo the whole lecture from first principles. Therefore, the discussion in this section is supposed to stand by itself, even if you missed both the lecture and recitation and didn't read any other material. I will try to explain, from first principles, how to find eigenvectors and eigenvalues.

However, I *will* assume you've read r03.pdf and r04.pdf from my website, so if you haven't read those yet, now's a good time to do so. Math builds on itself, so you need to have those down to make further progress here.

§4.1 Terminology

First, the relevant definition:

Definition 4.1: Given a square matrix $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$, suppose that

$$A\mathbf{v} = \lambda\mathbf{v}$$

for some nonzero vector \mathbf{v} . Then we say that \mathbf{v} is an *eigenvector* for A , and λ is the *eigenvalue*.

The prefix “eigen-” means self or own. It means that A maps \mathbf{v} to a multiple of itself.

The basic problem is:

Problem 4.2: If we have a matrix A , can we find any and all the eigen-things for it?

§4.1.1 Type signatures

I forgot to harp about types as I did in R01.

- The type of “eigenvector” is a *vector*.
- The type of “eigenvalue” is a *scalar* (i.e. it's a real number).

§4.2 Story of how to derive the eigen-things

Let's pretend we didn't go to lecture or recitation, and see how to solve this problem. For this story, our protagonist will be the matrix

$$A = \begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix}.$$

Phrased another way, the problem of finding eigenvectors is, by definition, looking for λ, x, y such that

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \iff \begin{cases} 5x - 2y = \lambda x \\ 3x + 10y = \lambda y \end{cases}$$

Smart-alecks will say $x = y = 0$ always works for every λ . *Are there other solutions?*

§4.2.1 Why guessing the eigenvalues is ill-fated

As an example, let's see if there are any eigenvectors $\begin{pmatrix} x \\ y \end{pmatrix}$ with eigenvalue 100. In other words, let's solve

$$\begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 100 \begin{pmatrix} x \\ y \end{pmatrix}.$$

If we solve the system of equations, we get

$$\begin{cases} 5x - 2y = 100x \\ 3x + 10y = 100y \end{cases} \implies \begin{cases} -95x - 2y = 0 \\ 3x - 90y = 0 \end{cases} \implies x = y = 0.$$

Well, that's boring. In this system of equations, the only solution is $x = y = 0$.

We can try a different guess: maybe we use 1000 instead of 100. An eigenvector with eigenvalue 1000 ought to be correspond to

$$\begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1000 \begin{pmatrix} x \\ y \end{pmatrix}.$$

If we solve the system of equations, we get

$$\begin{cases} 5x - 2y = 1000x \\ 3x + 10y = 1000y \end{cases} \implies \begin{cases} -995x - 2y = 0 \\ 3x - 990y = 0 \end{cases} \implies x = y = 0$$

which... isn't any better. We still don't get any solutions besides $x = y = 0$.

At this point, you should be remembering something I told you in R04: a "random" system of equations and variables usually only has a unique solution. So if I keep picking numbers out of a hat like 100, 1000, etc., then I'm unlikely to find anything interesting. In order to get a system that doesn't just solve to $x = y = 0$, I'm going to need to cherry-pick my number λ .

§4.2.2 Cherry-picking λ

Let's try to figure out what value of λ would make the system more interesting. If we copy what we did above, we see that the general process is:

$$\begin{cases} 5x - 2y = \lambda x \\ 3x + 10y = \lambda y \end{cases} \implies \begin{cases} (5 - \lambda)x - 2y = 0 \\ 3x + (10 - \lambda)y = 0 \end{cases}$$

We need to cherry-pick λ to make sure that the system doesn't just solve to $x = y = 0$ like the examples we tried with 100 and 1000. But we learned how to do this in R04: in order to get a degenerate system you need to make sure that

$$0 = \det \begin{pmatrix} 5 - \lambda & -2 \\ 3 & 10 - \lambda \end{pmatrix}.$$

i Remark

At this point, you might notice that this is secretly an explanation of why $A - \lambda I$ keeps showing up on your formula sheet. Writing $A\mathbf{v} = \lambda\mathbf{v}$ is the same as $(A - \lambda I)\mathbf{v} = 0$, just more opaquely.

Expanding the determinant on the left-hand side gives

$$0 = \det \begin{pmatrix} 5 - \lambda & -2 \\ 3 & 10 - \lambda \end{pmatrix} = (5 - \lambda)(10 - \lambda) + 6 = \lambda^2 - 15\lambda + 56 = (\lambda - 7)(\lambda - 8).$$

Great! So we expect that if we choose either $\lambda = 7$ and $\lambda = 8$, then we will get a degenerate system, and we won't just get $x = y = 0$. Indeed, let's check this:

- When $\lambda = 7$, our system is

$$\begin{cases} 5x - 2y = 7x \\ 3x + 10y = 7y \end{cases} \implies \begin{cases} -2x - 2y = 0 \\ 3x + 3y = 0 \end{cases} \implies x = -y.$$

So for example, $\begin{pmatrix} -13 \\ 13 \end{pmatrix}$ and $\begin{pmatrix} 37 \\ -37 \end{pmatrix}$ will be eigenvectors with eigenvalue 7:

$$A \begin{pmatrix} -13 \\ 13 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix} \begin{pmatrix} -13 \\ 13 \end{pmatrix} = \begin{pmatrix} -91 \\ 91 \end{pmatrix} = 7 \begin{pmatrix} -13 \\ 13 \end{pmatrix}.$$

On exam, you probably answer “the eigenvectors with eigenvalue 7 are the multiples of $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ”. Or you can say “the eigenvectors with eigenvalue 7 are the multiples of $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ” if you want; these are the same thing. You could even say “the eigenvectors with eigenvalue 7 are the multiples of $\begin{pmatrix} 100 \\ -100 \end{pmatrix}$ ” and still get credit, but that's silly.

- When $\lambda = 8$, our system is

$$\begin{cases} 5x - 2y = 8x \\ 3x + 10y = 8y \end{cases} \implies \begin{cases} -3x - 2y = 0 \\ 3x + 2y = 0 \end{cases} \implies x = -\frac{2}{3}y.$$

So for example, $\begin{pmatrix} -20 \\ 30 \end{pmatrix}$ is an eigenvector with eigenvalue 8:

$$A \begin{pmatrix} -20 \\ 30 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix} \begin{pmatrix} -20 \\ 30 \end{pmatrix} = \begin{pmatrix} -160 \\ 240 \end{pmatrix} = 8 \begin{pmatrix} -20 \\ 30 \end{pmatrix}.$$

On exam, you should answer “the eigenvectors with eigenvalue 8 are the multiples of $\begin{pmatrix} -2 \\ 3 \end{pmatrix}$ ”. Or you can say “the eigenvectors with eigenvalue 7 are the multiples of $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$ ” if you want; these are the same thing. You could even say “the eigenvectors with eigenvalue 7 are the multiples of $\begin{pmatrix} 200 \\ -300 \end{pmatrix}$ ” and still get credit, but that's silly.

§4.3 Summary

To summarize the story above:

- We had the matrix $A = \begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix}$ and wanted to find λ 's for which the equation

$$\begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$

had solutions other than $x = y = 0$.

- We realized that guessing λ was never going to fly, so we went out of our way to cherry-pick λ to make sure the system was degenerate. The buzzwords for this are “find the roots of the characteristic polynomial”, but I wanted to show that it flows naturally from the end goal.
- For the two values of λ we cherry-picked, we know the system of equations is degenerate. So we solve the two degenerate systems and see what happens.

In lectures and notes, the last two bullets are separated as two different steps, to make it into a recipe. But don't lose sight of how they're connected! I would rather call it the following interlocked thing:

We cherry-pick λ to make sure the system doesn't just solve to $x = y = 0$.

To do the cherry-picking, ensure $\det(A - \lambda I) = 0$ so that our system is degenerate.

§4.3.1 Degenerate systems of two equations all look stupid

This is worth repeating: **degenerate systems of two equations all look stupid**. Earlier on, we saw the two systems

$$\begin{cases} -2x - 2y = 0 \\ 3x + 3y = 0 \end{cases} \quad \text{and} \quad \begin{cases} -3x - 2y = 0 \\ 3x + 2y = 0 \end{cases}.$$

Both look moronic to the eye, because in each equation, the two equations say the same thing. This is by design: when you're solving the eigenvector problem, *you're going out of your way to find degenerate systems* so that there will actually be solutions besides $x = y = 0$.

In particular: if you do all the steps right, **you should never wind up with $x = y = 0$ as your only solution**. That means you either didn't do the cherry-picking step correctly, or something went wrong when you were solving the system. If that happens, check your work!

§4.3.2 Degenerate systems of three equations may not look stupid, but they are

When you have three or more equations instead, they don't necessarily look as stupid. To reuse the example I mentioned from R04, we have

$$\begin{aligned} x + 10y - 9z &= 0 \\ 3x + y + 10z &= 0 \\ 4x + 11y + z &= 0 \end{aligned}$$

which doesn't look stupid. But again, if you check the determinant, you find out

$$\det \begin{pmatrix} 1 & 10 & -9 \\ 3 & 1 & 10 \\ 4 & 11 & 1 \end{pmatrix} = 0.$$

So you know *a priori* that there will be solutions besides $x = y = z = 0$.

I think 18.02 won't have too many situations where you need to solve a degenerate three-variable system of equations, because it's generally annoying to do by hand. But if it happens, you should fall back on your high school algebra and solve the system however you learned it in 9th or 10th grade. The good news is that one of the three equations is always redundant, so you can just throw one away and solve for the other two. For example, in this case we would solve

$$\begin{aligned} x + 10y &= 9z \\ 3x + y &= -10z \end{aligned}$$

for x and y , as a function of z . I think this particular example works out to $x = -\frac{109}{29}z$, $y = \frac{37}{29}z$. And it indeed fits the third equation too.

§4.4 Good news for people who don't like eigenthings

If you don't like eigen-things, I have some good news for you: according to what I'm hearing, they won't actually be used much later on this course.