Eigenvector puzzle solution

April 6, 2018

In the puzzle we were given the following matrix,

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix},$$

and we were asked to compute several powers of this matrix. Taking the time to work out the first few, here's what we get.

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

$$A^3 = AA^2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

$$A^4 = AA^3 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$$

$$A^5 = AA^4 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 5 & 8 \end{bmatrix}$$

$$A^6 = AA^5 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 5 & 8 \end{bmatrix} = \begin{bmatrix} 5 & 8 \\ 8 & 13 \end{bmatrix}$$

Two things stand out as we do this. First, this process is tedious! Finding matrix powers is no fun, and the prospect of finding something like A^{100} feels truly horrifying. And even if you have a computer to evaluate this for you, finding matrix powers simply by repeating matrix multiplication stands to be very computationally inefficient, especially for large matrices. Switching to an eigenbasis, as you are about to see, can greatly speed up the process, both for doodling humans and number-crunching machines.

The second thing that stands out is that the terms in side our matrix appear to be following the Fibonacci sequence! $1, 1, 2, 3, 5, 8, 13, \ldots$ There's actually a good reason for this. Notice how this matrix acts on an arbitrary vector:

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} b \\ a+b \end{bmatrix}$$

So the second component of the input, b, vector gets shifted to be the first component of the output, and the second component of the output becomes the sum of a and b. For example, think of applying that process repeatedly to a given vector:

$$\begin{bmatrix} a \\ b \end{bmatrix} \to \begin{bmatrix} b \\ a+b \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \to \begin{bmatrix} 1 \\ 2 \end{bmatrix} \to \begin{bmatrix} 2 \\ 3 \end{bmatrix} \to \begin{bmatrix} 3 \\ 5 \end{bmatrix} \to \begin{bmatrix} 5 \\ 8 \end{bmatrix} \to \begin{bmatrix} 8 \\ 13 \end{bmatrix} \to \cdots$$

This iterative process perfectly capture the one defining the Fibonacci sequence: Each term is the sum of the previous two terms.

So, since multiplying the matrix A^n by a vector corresponds to going through this process n times, figuring out how to represent A^n can lead us to an explicit formula for the nth Finbonacci term. I don't know about you, but I find that awesome! Well, maybe it's not obvious that a clean formula for A^n is achievable, but this is where eigenvectors come in.

We can find the eigenvalues of this matrix by subtracting a variable λ off the diagonal, and computing when its determinant is 0:

$$\det([A - \lambda I) = \det\left(\begin{bmatrix} 0-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix}\right) = \lambda^2 - \lambda - 1 = 0$$

This quadratic has two solutions:

$$\begin{aligned} \lambda_1 &= \frac{1+\sqrt{5}}{2} \\ \lambda_2 &= \frac{1-\sqrt{5}}{2} \end{aligned}$$

Some of you may recognize the first of these as the golden ratio $\varphi = 1.618\ldots$, and the second as $-\frac{1}{\varphi} = -0.618\ldots$.

To find the eigenvectors, you find a non-zero solution to the equation

$$\begin{bmatrix} 0-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Doing so, you'll find that the eigenvectors corresponding to λ_1 and λ_2 are

$$\mathbf{v}_1 = \begin{bmatrix} \frac{1}{2}(-1+\sqrt{5}) \\ 1 \end{bmatrix}$$

$$\mathbf{v}_2 = \begin{bmatrix} \frac{1}{2}(-1-\sqrt{5}) \\ 1 \end{bmatrix}$$

Why are we doing this? Well, think about what it means for $\mathbf{v}_1$ to be an eigenvector of A. It means multiplying A by $\mathbf{v}_1$ just has the effect of scaling it by a constant, namely the eigenvalue λ_1. So repeatedly multiplying A by this vector, say 100 times, just has the effect of scaling this vector by λ_1^{100}. So, if we switch to a coordinate system where the two eigenvalues are our basis, the action A^n will have a very simple form, since its effect on the basis vectors would simply be multiplying by λ_1^n and λ_2^n.

So...how do you switch to a different basis? If you'll recall from the video on that topic, a matrix whose columns are the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ can be thought of as translating a coordinates describing a certain vector in the language of that basis system into a description of that same vector in our basis. That is, define the matrix

$$S = \begin{bmatrix} \frac{1}{2}(-1+\sqrt{5}) & \frac{1}{2}(-1-\sqrt{5}) \\ 1 & 1 \end{bmatrix}.$$

Then S is the change of basis matrix which has the effect of taking in a vector written in the coordinates of the eigenbasis, and spitting out the coordinates for that same vector in our standard basis. So to take the action of the matrix A, and write it in the language of the eigenbasis, we'd write an expression like this:

$$D = S^{-1}AS$$

You might read the right side here as saying first translate from the language of the eigenbasis to our basis, then apply A, then translate back. The resulting matrix D will represent the same transformation as A, just in the language of the eigenbasis. The whole point of doing this is to make it so that the action of A just looks like stretching two basis vectors. Let's see how that plays out!

The inverse of S is

$$S^{-1} = \begin{bmatrix} \frac{1}{2}(-1+\sqrt{5}) & \frac{1}{2}(-1-\sqrt{5}) \\ 1 & 1 \end{bmatrix}^{-1} = \frac{1}{\det(S)} \begin{bmatrix} 1 & -\frac{1}{2}(-1-\sqrt{5}) \\ -1 & \frac{1}{2}(-1+\sqrt{5}) \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -\frac{1}{2}(-1-\sqrt{5}) \\ -1 & \frac{1}{2}(-1+\sqrt{5}) \end{bmatrix}$$

Crunching through it all, we then get

$$D = S^{-1}AS = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -\frac{1}{2}(-1-\sqrt{5}) \\ -1 & \frac{1}{2}(-1+\sqrt{5}) \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{2}(-1+\sqrt{5}) & \frac{1}{2}(-1-\sqrt{5}) \\ 1 & 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -\frac{1}{2}(-1-\sqrt{5}) \\ -1 & \frac{1}{2}(-1+\sqrt{5}) \end{bmatrix}\begin{bmatrix} 1 & 1 \\ \frac{1}{2}(1+\sqrt{5}) & \frac{1}{2}(1-\sqrt{5}) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2}(1+\sqrt{5}) & 0 \\ 0 & \frac{1}{2}(1-\sqrt{5}) \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

We actually could have concluded this ahead of time, since the action of A on the first eigenvector is to scale it by $\lambda_1 = \frac{1}{2}(1+\sqrt{5})$, and its action on the second basis vector is to scale it by $\lambda_1 = \frac{1}{2}(1-\sqrt{5})$. Nevertheless, it's nice to get a little computational confirmation now and then.

Okay, final stretch! This means in the eignebases, the action of our transformation n times is simply

$$D^n = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^n = \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix}$$

Compare that to how it felt to compute the first few powers of A^n, much easier! Of course, this is all in coordinate system of the eigenbasis, so to translate back to our coordinate system, we have

$$A^n = S(D^n)S^{-1} = \begin{bmatrix} \frac{1}{2}(-1+\sqrt{5}) & \frac{1}{2}(-1-\sqrt{5}) \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -\frac{1}{2}(-1-\sqrt{5}) \\ -1 & \frac{1}{2}(-1+\sqrt{5}) \end{bmatrix}$$

This can be a written a little more nicely if we notice the occurance of $\lambda_1 = \frac{1}{2}(1+\sqrt{5})$ and $\lambda_2 = \frac{1}{2}(1-\sqrt{5})$ in the change of basis matrices, and pull the $\frac{1}{\sqrt{5}}$ term to the front. Also, in doing the calculuation, I'll take advantage of the convenient fact that $\lambda_1\lambda_2 = -1$

$$A^n = S(D^n)S^{-1} = \frac{1}{\sqrt{5}} \begin{bmatrix} -\lambda_2 & -\lambda_1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 \\ -1 & -\lambda_2 \end{bmatrix}$$

$$= \frac{1}{\sqrt{5}} \begin{bmatrix} -\lambda_2 & -\lambda_1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^n & \lambda_1^{n+1} \\ -\lambda_2^n & -\lambda_2^{n+1} \end{bmatrix}$$

$$= \frac{1}{\sqrt{5}} \begin{bmatrix} -\lambda_2\lambda_1^n + \lambda_1\lambda_2^n & -\lambda_2\lambda_1^{n+1} + \lambda_1\lambda_2^{n+1} \\ \lambda_1^n - \lambda_2^n & \lambda_1^{n+1} - \lambda_2^{n+1} \end{bmatrix}$$

$$= \frac{1}{\sqrt{5}} \begin{bmatrix} \lambda_1^{n-1} - \lambda_2^{n-1} & \lambda_1^n - \lambda_2^n \\ \lambda_1^n - \lambda_2^n & \lambda_1^{n+1} - \lambda_2^{n+1} \end{bmatrix}$$

If you compare this to the first few computations we did for A^n at the start, this gives us a very striking formula. The nth Fibonacci number has the formula

$$F_n = \frac{\lambda_1^n - \lambda_2^n}{\sqrt{5}} = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}}$$

It's not even obvious that this expression should give a rational expression, much less the Fibonacci numbers!