Methods of Mathematics

Kenneth A. Ribet

UC Berkeley

Math 10B
April 21, 2016

There is a new version of the online “textbook” file `Matrix_Algebra.pdf`.

There’s a “pop-up” lunch tomorrow at the Faculty Club at 12:30PM. All are welcome.

Course evaluations for this class are open. “Although students received an invitation email and reminders along the way, previous research demonstrates that a personal reminder from the instructor and an explanation of how evaluations are used to inform your teaching can make a positive impact on response rate and quality.”

Three remaining lectures:

- Today: Eigenvalues and eigenvectors
- Next Tuesday: Least squares, linear regression
- Next Thursday: Dynamic programming

RRR week and exam week follow.

Our exam: Monday, May 9, 11:30AM–2:30PM.

This morning’s breakfast group

Eigenvalues and eigenvectors

I can tell you what eigenvalues and eigenvectors are.

I can tell you how to compute them.

We can do some examples.

Today I will *not* tell you why eigenvalues are important or give examples of applications. To get some background, you can read the Wikipedia pages on eigenvalues and eigenvectors and on the characteristic polyonimal.

Section 4 of `Matrix_Algebra.pdf` has some material on this subject—about 11 pages (presented as slides).

You can do a google search on something like `brain imaging eigenvalues`. You'll find statements like

> *The eigenvalue magnitudes may be affected by changes in local tissue microstructure with many types of tissue injury, disease or normal physiological changes (i.e., aging).*

The Wikipedia article on Fractional anisotropy pops up as well. Oh, and see the Diffusion MRI article as well.

Suppose that A is an $n \times n$ square matrix. If v is a column vector of length n, Av is ("another") column vector of length n. If $v = 0$ then, $Av = 0$. However, for v non-zero, it is unlikely that Av will be proportional to v (i.e., a multiple of v).

For a first example, let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. If $v = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbf{R}^2$, then $Av = \begin{pmatrix} -b \\ a \end{pmatrix}$. You can convince yourself that Av is gotten from v by a 90-degree counterclockwise rotation. (Use polar coordinates if you like.) Hence there is no non-zero v in $\mathbf{R}^2$ such that Av is proportional to v.

If you're allowed to use complex numbers, you can take $v = \begin{pmatrix} 1 \\ -i \end{pmatrix}$. Then $Av = iv$ is (complex) proportional to v.

For the second example, we take $A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$. Then you can check by computation that

$$A\begin{pmatrix}2\\1\end{pmatrix} = 6\begin{pmatrix}2\\1\end{pmatrix}, \quad A\begin{pmatrix}2\\-1\end{pmatrix} = 1 \cdot \begin{pmatrix}2\\-1\end{pmatrix}.$$

The two vectors $\begin{pmatrix}2\\1\end{pmatrix}$ and $\begin{pmatrix}2\\-1\end{pmatrix}$ are called *eigenvectors*, and the proportionality constants 6 and 1 (respectively) are the corresponding *eigenvalues*. To each eigenvector, we can associate a *unique* eigenvalue.

By definition, eigenvectors are *non-zero*, though eigenvalues can be 0. For example, if A is a matrix with all 0s, then every non-zero vector is an eigenvector with eigenvalue equal to 0.

Notice that if v is an eigenvector, every non-zero multiple of v is an eigenvector with the same eigenvalue.

In the first example, we had a 2×2 matrix that had no eigenvectors (with real coordinates). In the second, the 2×2 matrix had two eigenvectors $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ that were not proportional to each other. We also mentioned the example of the 0-matrix, where all non-zero vectors are eigenvectors.

Now consider

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

For this matrix, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is an eigenvector with eigenvalue 0, and thus $\begin{pmatrix} x \\ 0 \end{pmatrix}$ is an eigenvector (with the same eigenvalue) for all non-zero real numbers x. However, this eigenvector is "unique" in the sense that all eigenvectors are multiples of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

The question becomes approximately this one: if A is a square matrix, how do we find its eigenvectors and eigenvalues?

The key to the solution is to focus on the *eigenvalues*. If v is an eigenvector of A with eigenvalue λ, then, by definition, $Av = \lambda v$. We can rewrite this:

$$(A - \lambda I)v = 0.$$

Thus λ is an eigenvalue precisely when there is a non-zero vector v such that $(A - \lambda I)v = 0$.

Now if B is a square matrix, there is a non-zero v such that $Bv = 0$ if and only if $\det B = 0$.

Aside: this is not obvious but is the sort of thing that's proved in Math 54 and Math 110. That's one of the reasons that people love their determinants.

Hence λ is an eigenvalue of A if and only if

$$\det(A - \lambda I) = 0.$$

Now what sort of condition is $\det(A - \lambda I) = 0$?

Take the example $A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$. Then

$$A - \lambda I = \begin{pmatrix} 5-\lambda & 2 \\ 2 & 2-\lambda \end{pmatrix},$$

so that

$$\det(A - \lambda I) = (5-\lambda)(2-\lambda) - 4 = \lambda^2 - 7\lambda + 6.$$

Thus the eigenvalues of A are the roots of

$$\lambda^2 - 7\lambda + 6 = (\lambda - 1)(\lambda - 6).$$

These roots are 1 and 6.

One we know the eigenvalues, it's a simple matter to find out the corresponding eigenvectors. Let's keep the A in the last example and focus on $\lambda = 6$. The corresponding eigenvectors $\begin{pmatrix} x \\ y \end{pmatrix}$ are those vectors that satisfy $(A - 6I)\begin{pmatrix} x \\ y \end{pmatrix} = 0$. This equation amounts to a pair of equations in the two unknowns x and y:

$$\begin{aligned} 5x + 2y - 6x &= 0, \\ 2x + 2y - 6y &= 0. \end{aligned}$$

In general, we might want to use Gaussian elimination to solve a system like this. In this situation, however, both equations state that $x = 2y$, so we can write down the solution by hand: the eigenvectors are the multiples of $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Summary:

- We find the eigenvalues by solving a polynomial equation;
- for each eigenvalue, we find the corresponding eigenvectors by Gaussian elimination.

The polynomial equation that's satisfied by the eigenvalues of A is

$$\det(A - \lambda I) = 0.$$

Usually we say that $p(\lambda) = 0$ where p is the polynomial

$$\det(A - t \cdot I);$$

this is a polynomial in the variable t, the *characteristic polynomial* of A.

Caution:

Some authors prefer to define the characteristic polynomial as $\det(t \cdot I - A)$ instead of $\det(A - t \cdot I)$. The two differ by multiplication by $(-1)^n$. In other words, they're the same if n is even and negatives of each other if n is odd.

We should obviously have at least one example of a 3×3 matrix. There is one on page 73 of the `Matrix_Algrebra` "textbook" and one on the Wikipedia page about finding eigenvalues. Let's do the Wikipedia example:

$$A = \begin{pmatrix} 3 & 2 & 6 \\ 2 & 2 & 5 \\ -2 & -1 & -4 \end{pmatrix}.$$

If p is the characteristic polyomial of A, then `Sage` gives

$$p(x) = x^3 - x^2 - x + 1.$$

(They're apparently using the convention $p(x) = \det(x \cdot I - A)$.)

We could obviously compute this polynomial in real time (on the document camera), but I won't do it unless you egg me on.

Obviously, if someone gives you a random cubic polynomial, you have no prayer of finding the roots of the polynomial. However, it's 99% obvious that $p(1) = 0$, so 1 is a root. In fact, $p(x) = (x-1)^2(x+1)$, as we can check on the doc camera. Hence there are two eigenvalues: 1 and -1. Because 1 "occurs twice" as a root of $p(x)$, people say that the *algebraic multiplicity* of 1 as an eigenvalue is 2.

What about the eigen*vectors*? There are eigenvectors for the eigenvalue 1 and eigenvectors for the eigenvalue -1. We can find them by Gauss (or whatever). For $+1$, we need to find the vectors v such that $(A - I)v = 0$:

$$\begin{pmatrix} 2 & 2 & 6 \\ 2 & 1 & 5 \\ -2 & -1 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0.$$

There are actually only two equations in three unknowns because the bottom of the three equations is just -1 times the middle equation!

The solution is that z can be anything; then $y = -z$, $x = -2z$. The eigenvectors for the eigenvalue 1 are thus the non-zero multiples of $\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$.

There's one more topic that we need to touch on: it's the relationship between what we've been discussing and systems of first-order equations like those in the section "Linear systems of differential equations" in `Matrix_Algebra.pdf` (pp. 119–). As an example:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} 7 & -2 \\ 15 & -4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

which appears in next week's homework.

Here, y_1 and y_2 are two functions that are linked together through their derivatives: $y_1' = 7y_1 - 2y_2$, $y_2' = 15y_1 - 4y_2$.

In our past lives, we considered DEs like

$$y'' - 3y' + 2y = 0.$$

If we put $y_1 = y$, $y_2 = y'$, then we can rewrite this as

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} y_2 \\ -2y_1 + 3y_2 \end{pmatrix}.$$

If the 2×2 matrix in the 7, -2 problem were the diagonal matrix $\begin{pmatrix} 3 & 0 \\ 0 & -5 \end{pmatrix}$, then the system would say $y_1' = 3y_1$, $y_2' = -5y_2$ and the solution would be $y_1 = C_1 e^{3t}$, $y_2 = C_2 e^{-5t}$. We could write

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = C_1 e^{3t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + C_2 e^{-5t} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

This is sort of pedantic, but 3 and -5 are the eigenvalues in this diagonal problem; also $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are the eigenvectors.

In the non-diagonal example that began the discussion, the solution is going to be

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = C_1 e^{2t} \begin{pmatrix} 2 \\ 5 \end{pmatrix} + C_2 e^{t} \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

The characteristic polynomial here is $t^2 - 3t + 2$; the roots of the polynomial are 2 and 1. The vectors $\begin{pmatrix} 2 \\ 5 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ are the eigenvectors corresponding to the eigenvalues 2 and 1. (We need to check this on the document camera.)

Thus we have a recipe, and we need only see why it works: (See also slide 120 of `Matrix_Algebra`.) If $Av = \lambda v$ and $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = e^{\lambda t} v$, then $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \lambda e^{\lambda t} v = e^{\lambda t} A v = A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$.

This is kind of formal—it goes by fast—but it's actually quite brilliant.