Eigenvalues by row operations

Barton L. Willis
Department of Mathematics
University of Nebraska at Kearney
Kearney, Nebraska 68849

May 3, 2005

1 Introduction

There is a wonderful article, "Down with Determinants!," by Sheldon Axler that "shows how linear algebra can be done better without determinants." [1] The article received the Lester R. Ford Award for expository writing from the Mathematical Association of America.

Axler's article shows how to prove some standard results in linear algebra without using determinants, but it doesn't show how to do calculations without using determinants. My article will show you two ways to find the eigenvalues of a matrix without using determinants. The first method (Section 2) uses only row operations; the second method (Section 3) applies ideas from Axler's article. A complete description of the second method is in an article by William A. McWorter and Leroy F. Meyers [2]; I do not know of a reference for the first method. Section 4 concludes with a method for solving generalized eigenvalue problems.

This article started as class notes for my spring 2005 linear algebra class. Since then, I more than doubled the length, but I kept some of the original style of the notes.

2 The row reduction method

A number z is an eigenvalue of a square matrix A provided $A - zI$ is singular. The best way to determine if a matrix is singular is to reduce it to a triangular form. So it seems that a good scheme for finding the eigenvalues of a matrix A would be to find a triangular form of $A - zI$. We'll show that $A - zI$ has a triangular form with a "simple" representation, but finding it requires a trick. I'll show you how to do it for a 3×3 matrix. You'll need to convince yourself that the process can be done for an arbitrary square matrix. To illustrate, let's try to reduce

$$\begin{bmatrix} 1-z & 2 & 1 \\ 1 & 3-z & 1 \\ 1 & 4 & -1-z \end{bmatrix}$$

to a triangular form. If $z \neq 1$, we could use the 1,1 entry as a pivot and do the row operations $R_2 \leftarrow R_2 - \frac{1}{1-z} R_1$ and $R_3 \leftarrow R_3 - \frac{1}{1-z} R_1$. The case $z = 1$ would need to be handled in a separate calculation. This plan would result in a triangular form that is not simple (multiple cases). A better scheme would be to swap rows 1 and 2. This gives

$$\begin{bmatrix} 1 & 3-z & 1 \\ 1-z & 2 & 1 \\ 1 & 4 & -1-z \end{bmatrix}.$$

Now the 1,1 entry is a nonzero constant; we'll call 1 a good pivot[1] because it is nonzero and constant. The row operations $R_2 \leftarrow R_2 - (1-z) R_1$ and $R_3 \leftarrow R_3 - R_1$ place zeros below the diagonal in the first column; thus

$$\begin{bmatrix} 1 & 3-z & 1 \\ 0 & 2-(1-z)(3-z) & z \\ 0 & z+1 & -z-2 \end{bmatrix} = \begin{bmatrix} 1 & 3-z & 1 \\ 0 & -z^2+4z-1 & z \\ 0 & z+1 & -2-z \end{bmatrix}.$$

Now we have trouble. The 2,2 and the 3,2 entries are both bad pivots (they are non-constant). This time, a row swap will not place a good pivot in the 2,2 position. We could try swapping columns, but for this matrix it doesn't help because every entry in the 2,2 sub-matrix[2] is a bad pivot. It's time for a trick. We'll try to do a row operation of the form $R_2 \leftarrow R_2 - \theta R_3$ that makes the 2,2 entry a good pivot. What value should we choose for θ? After the row operation, the 2,2 entry is

$$(-z^2 + 4z - 1) - \theta(z+1).$$

We'd like the 2,2 entry to be constant and nonzero. Can this be arranged? Certainly, define θ to be

$$\theta = \left(-z^2 + 4z - 1\right) \div (z+1) = -z + 5,$$

where by $\div$ we mean the quotient without the remainder. The row operation $R_2 \leftarrow R_2 - (-z+5) R_3$ yields

$$\begin{bmatrix} 1 & 3-z & 1 \\ 0 & -6 & -z^2+4z+10 \\ 0 & z+1 & -z-2 \end{bmatrix}.$$

We were successful; the 2,2 entry is a good pivot. It's now possible to do $R_3 \leftarrow R_3 + \frac{z+1}{6} R_2$. The result is

$$\begin{bmatrix} 1 & 3-z & 1 \\ 0 & -6 & -z^2+4z+10 \\ 0 & 0 & -\frac{1}{6}\left(z^3 - 3z^2 - 8z + 2\right) \end{bmatrix}.$$

[1] My students will recognize the terms "good pivot" and "bad pivot."

[2] The 2,2 sub-matrix is everything on row 2 and below and everything in column 2 and to the right. This concept generalizes to the i,j sub-matrix.

This is the echelon form we desired.

The characteristic polynomial is the product of the diagonal entries times $(-1)^k$, where k is the number of row swaps. We did one row swap, so the characteristic polynomial is

$$z \mapsto -z^3 + 3z^2 + 8z - 2.$$

Let's find the eigenvectors. To start, it might seem that we should first find the roots of the characteristic polynomial. Any computer algebra system will tell us the solutions, but they are big ugly messes. Here is one root[3]

$$-\frac{\left(\sqrt{3}\,i+1\right)\left(\sqrt{29}\sqrt{31}\,i+12\sqrt{3}\right)^{2/3} - 2\sqrt{3}\sqrt[3]{\sqrt{29}\sqrt{31}\,i+12\sqrt{3}} - 11\sqrt{3}\,i + 11}{2\sqrt{3}\sqrt[3]{\sqrt{29}\sqrt{31}\,i+12\sqrt{3}}}.$$

What does a mathematician do when things get unbearably messy? We give the messy thing a name and push forward. Let's try this approach. Let z_1 be an eigenvalue. In the row reduced matrix, substitute $z \to z_1$, and use the fact that $z_1^3 - 3z_1^2 - 8z_1 + 2 = 0$. Now do back substitution on the matrix

$$\begin{bmatrix} 1 & 3-z_1 & 1 \\ 0 & -6 & -z_1^2+4z_1+10 \\ 0 & 0 & -\frac{1}{6}\left(z_1^3-3z_1^2-8z_1+2\right) \end{bmatrix} = \begin{bmatrix} 1 & 3-z_1 & 1 \\ 0 & -6 & -z_1^2+4z_1+10 \\ 0 & 0 & 0 \end{bmatrix}.$$

Naming the unknowns a, b, c, we see that c is free. Choosing $c = 0$ makes $a = 0$ and $b = 0$. Eigenvectors must be nonzero, so the choice $c = 0$ doesn't work. Choosing $c = 1$ gives

$$b = -\frac{1}{6}z_1^2 + \frac{2}{3}z_1 + \frac{5}{3},$$
$$a = -(3-z_1)\left(-\frac{1}{6}z_1^2 + \frac{2}{3}z_1 + \frac{5}{3}\right) - 1.$$

We could stop now, but the expression for a is cubic in z_1. Expanding a and substituting $z_1^3 \to 3z_1^3 + 8z_1 - 2$ gives a simpler representation for a. It is

$$a = \frac{2}{3}z_1^2 - \frac{5}{3}z_1 - \frac{17}{3}.$$

The eigenvector is

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -\frac{17}{3} \\ \frac{5}{3} \\ 1 \end{bmatrix} + z_1 \begin{bmatrix} -\frac{5}{3} \\ \frac{2}{3} \\ 0 \end{bmatrix} + z_1^2 \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{6} \\ 0 \end{bmatrix}, \tag{1}$$

where z_1 is *any* eigenvalue. Since the characteristic polynomial has three distinct roots, we were able to find *all* the eigenvectors using only one back substitution!

[3]You might guess that the imaginary part of this number is nonzero, but it's not. Remember that every cubic polynomial with real coefficients has at least one real root. Of the three roots of the polynomial, I chose to display the real root.

Does the row reduction method work when the characteristic polynomial has a degenerate root? It does, but there is a slight twist. To illustrate, let's find the eigenvectors of

$$\begin{bmatrix} 1 & -1 & -3 \\ 0 & 2 & 0 \\ 1 & 1 & 5 \end{bmatrix}.$$

We begin by subtracting zI. Swapping rows 1 and 3 gives

$$\begin{bmatrix} 1 & 1 & 5-z \\ 0 & 2-z & 0 \\ 1-z & -1 & -3 \end{bmatrix}.$$

Next, do $R_3 \leftarrow R_3 - (1-z)R_1$; thus

$$\begin{bmatrix} 1 & 1 & 5-z \\ 0 & 2-z & 0 \\ 0 & z-2 & -z^2+6z-8 \end{bmatrix}.$$

Following the trick we did in the first example, we now apply $R_2 \leftarrow R_2 - \theta R_3$, where

$$\theta = (2-z) \div (z-2) = -1.$$

The result is

$$\begin{bmatrix} 1 & 1 & 5-z \\ 0 & 0 & z^2-6z+8 \\ 0 & z-2 & -z^2+6z-8 \end{bmatrix}.$$

This time the trick doesn't make the 2,2 entry a good pivot (the slight twist). Nevertheless, swapping rows 2 and 3 gives the triangular form

$$\begin{bmatrix} 1 & 1 & 5-z \\ 0 & z-2 & -z^2+6z-8 \\ 0 & 0 & z^2-6z+8 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 5-z \\ 0 & z-2 & -(z-2)(z-4) \\ 0 & 0 & (z-4)(z-2) \end{bmatrix}$$

that we were trying to find. The characteristic polynomial is

$$z \mapsto -(z-4)(z-2)^2.$$

The eigenvalues are 2 and 4. Unlike the first example, we can't find all three eigenvectors with a single back substitution. Substituting $z \rightarrow 2$ and solving gives the eigenvectors $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$. The geometric multiplicity of this eigenvalue is 2 because the 2,2 and the 3,3 entries of the triangularized matrix have a common root. For this example, the common root was easy to identify; for more involved problems, it might be necessary to use the polynomial resultant to detect the common roots. [3] Substituting $z \rightarrow 4$ gives the eigenvector $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

3 Method of McWorter and Meyers

Our second method comes from "Computing eigenvalues and eigenvectors without determinants," by William A. McWorter and Leroy F. Meyers. [2] After reading just two pages out of ten, I had a slap-on-the-forehead moment. The article uses an idea that I've used to prove theorems, but I had never connected the idea to a computational method.

Here is the idea. Let A be a square matrix, and let $\mathbf{x}$ be any nonzero vector. Find the greatest integer m such that the set

$$\{\mathbf{x},\, A\mathbf{x},\, A^2\mathbf{x}, \ldots,\, A^{m-1}\mathbf{x}\}$$

is linearly independent. There are scalars $\alpha_0, \ldots, \alpha_{m-1}$ such that

$$\left(\alpha_0 I + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_{m-1} A^{m-1} + A^m\right)\mathbf{x} = \mathbf{0}.$$

Equivalently, there are scalars $z_0, \ldots, z_m$ such that

$$(A - z_0 I)(A - z_1 I)\cdots(A - z_m I)\mathbf{x} = \mathbf{0}.$$

Either $(A - z_1 I)\cdots(A - z_m I)\mathbf{x}$ is an eigenvector of A with eigenvalue z_0, or

$$(A - z_1 I)\cdots(A - z_m I)\mathbf{x} = \mathbf{0}.$$

The left side of this equation is a *nontrivial* linear combination of the *linearly independent* set $\{\mathbf{x},\, A\mathbf{x},\, A^2\mathbf{x}, \ldots,\, A^{m-1}\mathbf{x}\}$. Thus $(A - z_1 I)\cdots(A - z_m I)\mathbf{x} \neq \mathbf{0}$. This shows that $(A - z_1 I)\cdots(A - z_m I)\mathbf{x}$ is an eigenvector of A corresponding to the eigenvalue z_0.

Let's try this method on the example we worked using the row reduction method. Let

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 4 & -1 \end{bmatrix}.$$

Choose $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, and do row reduction on the matrix with columns $\mathbf{x}_0,\, A\mathbf{x}_0,\ A^2\mathbf{x}_0,\ A^3\mathbf{x}_0$. We have

$$\begin{bmatrix} 1 & 1 & 4 & 18 \\ 0 & 1 & 5 & 23 \\ 0 & 1 & 4 & 20 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 4 & 18 \\ 0 & 1 & 5 & 23 \\ 0 & 0 & -1 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

From this we discover that $2\mathbf{x}_0 - 8A\mathbf{x}_0 - 3A^2\mathbf{x}_0 + A^3\mathbf{x}_0 = \mathbf{0}$. Consequently,

$$\left(2I - 8A - 3A^2 + A^3\right)\mathbf{x}_0 = \mathbf{0}.$$

Factor this as

$$(A - z_1 I)(A - z_2 I)(A - z_3 I)\mathbf{x}_0 = \mathbf{0},$$

where z_1, z_2, and z_3 are zeros of the polynomial $z \mapsto 2 - 8z - 3z^2 + z^3$. The eigenvector corresponding to the eigenvalue z_1 is

$$(A - z_2 I)(A - z_3 I)\mathbf{x}_0 = \begin{bmatrix} z_2 z_3 - z_3 - z_2 + 4 \\ -z_3 - z_2 + 5 \\ -z_3 - z_2 + 4 \end{bmatrix}.$$

Showing that this vector is parallel to [see Eq. (1)]

$$\begin{bmatrix} -\frac{17}{3} \\ \frac{5}{3} \\ 1 \end{bmatrix} + z_1 \begin{bmatrix} -\frac{5}{3} \\ \frac{2}{3} \\ 0 \end{bmatrix} + z_1^2 \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{6} \\ 0 \end{bmatrix},$$

is an algebraic nightmare. I've been able to verify it numerically, but I have not been able to prove it.

4 Generalized eigenvalues

Let A and B be square matrices with the same size. A complex number z is a *generalized eigenvalue* of (A, B) provided $A - zB$ is singular. Determinants give a quick characterization of the generalized eigenvalues; z is a generalized eigenvalue of (A, B) provided $\det(A - zB) = 0$.

Can the computational method given in McWorter and Meyers be extended to find generalized eigenvalues? Can Axler's proofs be modified to develop a theory of generalized eigenvalues without resorting to determinants? I don't know. But I do know that the row operation method given in this article can be used to find generalized eigenvalues. Again, I demonstrate this by way of an example.

Let's find the generalized eigenvalues of (A, B), where[4]

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

We need to row reduce

$$\begin{bmatrix} 1 - z & 2 & 3 - z \\ 4 & 5 - z & 6 \\ 7 - z & 8 & 9 - z \end{bmatrix}.$$

Swapping rows 1 and 2 and eliminating in the first column yields

$$\begin{bmatrix} 4 & 5 - z & 6 \\ 0 & -\frac{z^2}{4} + \frac{3z}{2} + \frac{3}{4} & \frac{z}{2} + \frac{3}{2} \\ 0 & -\frac{z^2}{4} + 3z - \frac{3}{4} & \frac{z}{2} - \frac{3}{2} \end{bmatrix}.$$

[4]When B is invertible, the generalized eigenvalues of (A, B) are the eigenvalues of AB^{-1}. In this example, the matrix B is singular.

To eliminate in the second column, it would be advantageous to swap the second and third columns. Doing so would place first degree polynomials in the 2,2 and 3,2 positions. Instead of swapping columns, let's try to proceed as before. The quotient (without remainder) of the 2,2 entry divided by the 3,2 entry is 1. Thus we do the row operation $R_2 \leftarrow R_2 - R_1$. The result is

$$\begin{bmatrix} 4 & 5-z & 6 \\ 0 & \frac{3}{2}-\frac{3z}{2} & 3 \\ 0 & -\frac{z^2}{4}+3z-\frac{3}{4} & \frac{z}{2}-\frac{3}{2} \end{bmatrix}.$$

The second column still does not have a good pivot. What do we do? We use our trick again, but this time, we apply it to the third row. (The 3,2 entry has the greatest degree.) The quotient of the 3,2 and the 2,2 entries is $\frac{z-11}{6}$. The row operation $R_3 \leftarrow R_3 - \frac{z-11}{6}R_2$ gives

$$\begin{bmatrix} 4 & 5-z & 6 \\ 0 & \frac{3}{2}-\frac{3z}{2} & 3 \\ 0 & 2 & 4 \end{bmatrix}.$$

Finally, the second column has a good pivot. Swapping the second and third rows, and eliminating in the second column gives the triangular form

$$\begin{bmatrix} 4 & 5-z & 6 \\ 0 & 2 & 4 \\ 0 & 0 & 3z \end{bmatrix}.$$

The only generalized eigenvalue is 0. As a check, we have

$$\det \begin{bmatrix} 1-z & 2 & 3-z \\ 4 & 5-z & 6 \\ 7-z & 8 & 9-z \end{bmatrix} = 24z.$$

5 Acknowledgment

Sheldon Axler read a draft of this article and gave me the reference to the paper by McWorter and Meyers.

6 Conclusion

As presented in linear algebra books, all computations, except eigenvalues, rely on row reduction. Why should the eigenvalue problem be any different? This article shows how to solve both eigenvalue and generalized eigenvalue problems using a pure row reduction method.

One method that is not discussed in this article is due to Bareiss. [5] His method reduces a matrix to triangular form, but it changes the determinant. For example, applying his method to the matrix

$$\begin{bmatrix} 1-z & 2 \\ 3 & 4-z \end{bmatrix}$$

yields the triangular matrix

$$\begin{bmatrix} 1-z & 2 \\ 0 & z^2-5z-2 \end{bmatrix}.$$

Based on the triangular form, it might seem that 1 is an eigenvalue of $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, but it is not.

You may download software (written in Maxima [4]) for the row reduction method from my web page. [6]

I end with a quote from Sheldon Axler: “Down with determinants!”

References

[1] Sheldon Axler, “Down with Determinants!,” *American Mathematical Monthly* **102** (1995) 139–154.

[2] William A. McWorter and Leroy F. Meyers, “Computing eigenvalues and eigenvectors without determinants,” *Mathematics Magazine* **71** (1998) 24–33.

[3] planetmath.org/encyclopedia/Resultant.html .

[4] The Maxima Computer Algebra Project, maxima.sourceforge.net/ .

[5] E. H. Bareiss, “Sylvester’s Identity and Multistep Integer preserving Gaussian Elimination,” *Math. Comp.* **22** (1968) 565–578.

[6] Web page of Barton Willis, www.unk.edu/acad/math/people/willisb/ .