Linear Algebra
C Term, Sections C01-C04
W. J. Martin
February 28, 2010

Solutions: Computing Eigenvalues

NOTE: The solutions themselves begin on the third page.

The big picture

Suppose we have a linear transformation from a vector space to itself. We view $T : V \to V$ as a sort of process that moves each vector $\mathbf{x}$ to a destination $T(\mathbf{x})$. We view the vector space V together with this transformation as a "system" and we want to locate "stability" in this system.

Of course, we now know that, in the case where V has finite dimension n, the linear transformation T is entirely defined by an $n \times n$ matrix. We just fix any basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ for T and take the matrix

$$A = \left[\begin{array}{c|c|ccc|c} [\mathbf{b}_1]_{\mathcal{B}} & [\mathbf{b}_2]_{\mathcal{B}} & \cdots & & \cdots & [\mathbf{b}_n]_{\mathcal{B}} \end{array} \right] .$$

Then the eigenvalues of T are just the eigenvalues of matrix A and the eigenvectors are also in one-to-one correspondence via $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$.

Now we see that a good choice of basis $\mathcal{B}$ is necessary. If we choose a nice basis, then the matrix A may have very simple form and may reveal the essential nature of the transform T; if our of choice of basis is entirely random or uninformed, then the matrix A can very messy and unilluminating. That's our last goal in the course: finding the best basis.

Recalling the Definitions

Let A be an $n \times n$ matrix. A non-zero vector $\mathbf{v}$ in $\mathbb{R}^n$ is an *eigenvector* of A if the vector $A\mathbf{v}$ is parallel to $\mathbf{v}$:

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

for some scalar λ. This scalar is called the *eigenvalue* of A associated to $\mathbf{v}$. (We will only consider real eigenvalues and eigenvectors in this course; the complex case is entirely similar and is typically what arises in applications.)

Of course, if $\mathbf{v}$ is the zero vector, then $A\mathbf{v} = \lambda\mathbf{v}$ for every possible λ. So the zero vector clouds the issue. For this reason, we always require eigenvectors to be non-zero. (But zero is perfect fine as an eigenvalue.)

Our goal, given a square matrix A, is to find all possible eigenvalues and all eigenvectors of A. We've already had practice at the second task. Once we know λ, the set of all eigenvectors associated to λ is almost exactly the null space of the matrix $A - \lambda I$. (I say "almost" because the null space also includes the zero vector. But it's a shame to have all the ingredients of a vector space except a zero vector. So we will define the "eigenspace of A associated to λ" as $\text{Nul}\,(A - \lambda I)$.) We learned this connection as follows:

- if $A\mathbf{v} = \lambda\mathbf{v}$ ...
- then $A\mathbf{v} = \lambda(I\mathbf{v})$ since $I\mathbf{v}$ is the same as $\mathbf{v}$ for any vector $\mathbf{v}$
- so $A\mathbf{v} = (\lambda I)\mathbf{v}$ by part (d) of Theorem 2 on p113
- Now subtract the vector on the right from both sides: $A\mathbf{v} - (\lambda I)\mathbf{v} = \mathbf{0}$
- We can now factor out the $\mathbf{v}$ on the right (part (c) of that same theorem in Chapter 2): $(A - \lambda I)\mathbf{v} = \mathbf{0}$
- so every eigenvector associated to λ lives in the null space of $A - \lambda I$
- Now we can reverse all of the steps: if $\mathbf{v}$ is any vector in the null space of $A - \lambda I$ — except the zero vector — then $\mathbf{v}$ is also an eigenvector of A with eigenvalue λ

So I wanted you to understand, early on in the course, that your ability to solve these simple homogeneous linear systems automatically enables you to find eigenvectors.

But how to find the eigenvalues λ in the first place? That's where determinants come in to do their job.

Computing Eigenvalues

Let A be an $n \times n$ matrix. A scalar θ is an *eigenvalue* of A if and only if there is some non-zero vector $\mathbf{v}$ satisfying $A\mathbf{v} = \theta\mathbf{v}$.

So θ is an eigenvalue of A precisely when $A - \theta I$ has non-trivial null space. Now the Invertible Matrix Theorem gives us over a dozen different ways to proceed from here: $A - \theta I$ has fewer than n pivots; the columns are linearly dependent; the linear transformation $\mathbf{x} \mapsto (A - \theta I)\mathbf{x}$ is not one-to-one; and so on. In any given application, any one of these tests may be useful. But we stick with the traditional test: $A - \theta I$ is non-invertible if and only if its determinant is zero.

So we need to find all numbers θ for which $\det(A - \theta I) = 0$. The trick is to use a bit of abstraction and replace θ by a symbol. In the next few paragraphs, let λ be a real variable. Then for any $n \times n$ matrix A,

$$f_A(\lambda) := \det(A - \lambda I)$$

is a polynomial of degree n in variable λ. And the eigenvalues are just the *roots* of this polynomial!

So we need all of our mathematical skills, not just from linear algebra, but also from high school. We need to compute determinants, factor polynomials, and find roots. This requires serious practice.

The 2×2 case

Let's start with 2×2 matrices. Recall that the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has determinant $ad - bc$.

Example 1: Find all the eigenvalues of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

SOLUTION: We have $A - \lambda I = \begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix}$ with determinant

$$\begin{aligned} f_A(\lambda) &= (1-\lambda)(1-\lambda) - 1 \\ &= (\lambda - 1)^2 - 1 \\ &= \lambda^2 - 2\lambda \\ f_A(\lambda) &= \lambda(\lambda - 2) \end{aligned}$$

So the eigenvalues of A are $\lambda_1 = 2$ and $\lambda_2 = 0$. (It is conventional to give them in non-increasing order.)

The alert reader can easily find a basis of eigenvectors. Our methods from Chapter 1 give the basis $\mathcal{B} = \{(1,1),\ (1,-1)\}$.

Example 2: Find all the eigenvalues of $A = \begin{bmatrix} 3 & 2 \\ 8 & 3 \end{bmatrix}$

SOLUTION: We have $A - \lambda I = \begin{bmatrix} 3-\lambda & 2 \\ 8 & 3-\lambda \end{bmatrix}$ with determinant

$$\begin{aligned} f_A(\lambda) &= (3-\lambda)(3-\lambda) - 16 \\ &= (\lambda - 3)^2 - 16 \\ &= \lambda^2 - 6\lambda - 7 \\ f_A(\lambda) &= (\lambda - 7)(\lambda + 1) \end{aligned}$$

So the eigenvalues of A are $\lambda_1 = 7$ and $\lambda_2 = -1$. The corresponding eigenvectors are $(1,2)$ and $(1,-2)$.

Example 3: Now let's try $A = \begin{bmatrix} -1 & 4 \\ 4 & 5 \end{bmatrix}$

SOLUTION: We have $A - \lambda I = \begin{bmatrix} -1-\lambda & 4 \\ 4 & 5-\lambda \end{bmatrix}$ with determinant

$$\begin{aligned} f_A(\lambda) &= (-1-\lambda)(5-\lambda) - 16 \\ &= (\lambda+1)(\lambda-5) - 16 \qquad \text{I hate those minus signs!} \\ &= \lambda^2 - 4\lambda - 21 \\ f_A(\lambda) &= (\lambda-7)(\lambda+3) \end{aligned}$$

So the eigenvalues of A are $\lambda_1 = 7$ and $\lambda_2 = -3$. One can easily compute the corresponding eigenvectors: they are $(1, 2)$ and $(-2, 1)$. (It's a cool theorem that, when the matrix is symmetric, the eigenvectors associated to different eigenvalues must be orthogonal to one another.)

> Okay, those were simple enough. But I don't want to give you the impression that all matrices are so cooperative. Let's explore some anomalies.

Example 4: Here's an easy one: $A = \begin{bmatrix} -15 & 0 \\ 0 & -15 \end{bmatrix}$

SOLUTION: We have $A - \lambda I = \begin{bmatrix} -15-\lambda & 0 \\ 0 & -15-\lambda \end{bmatrix}$ with determinant

$$\begin{aligned} f_A(\lambda) &= (-15-\lambda)(-15-\lambda) \\ f_A(\lambda) &= (\lambda+15)^2 \end{aligned}$$

So the eigenvalues of A are $\lambda_1 = -15$ and $\lambda_2 = -15$. (We have an eigenvalue of "multiplicity" two.) Every non-zero vector in $\mathbb{R}^2$ is an eigenvector with this eigenvalue.

Example 5: This one looks easy at first: $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$

SOLUTION: We have $A - \lambda I = \begin{bmatrix} 1-\lambda & 3 \\ 0 & 1-\lambda \end{bmatrix}$ with determinant

$$\begin{aligned} f_A(\lambda) &= (1-\lambda)(1-\lambda) \\ f_A(\lambda) &= (\lambda-1)^2 \end{aligned}$$

So the eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 1$. (We again have an eigenvalue of multiplicity two.) But when we row reduce $A - I$, we find a null space of dimension one only. So **do not** get a basis of eigenvectors in this case.

Example 6: Find all the eigenvalues of $A = \begin{bmatrix} 5 & -2 \\ 3 & 5 \end{bmatrix}$

Solution: We have $A - \lambda I = \begin{bmatrix} 5-\lambda & -2 \\ 3 & 5-\lambda \end{bmatrix}$ with determinant

$$\begin{aligned} f_A(\lambda) &= (5-\lambda)(5-\lambda)+6 \\ &= (\lambda-5)^2+6 \\ f_A(\lambda) &= \lambda^2-10\lambda+31 \end{aligned}$$

But this stinks: the discriminant $b^2 - 4ac = (-10)^2 - 4 \cdot 1 \cdot 31 < 100 - 120$ is negative. So the eigenvalues are not real, but complex. We don't handle this case in this course, even though it is the typical case.

The 3×3 case

In order to compute eigenvalues of 3×3 matrices, we need to be able to compute these determinants. As we discussed in class, while the row reduction method gives the fastest method for finding eigenvalues of large matrices, the straightforward calculation from the definition is often most efficient (or least error-prone) for small matrices.

Here is a general 3×3 matrix and its determinant (all letters a-i are just placeholders for the nine entries:

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}, \qquad \det A = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = (aei)+(bfg)+(cdh)-(gec)-(hfa)-(idb).$$

Example 7: Find all the eigenvalues of $A = \begin{bmatrix} 1 & 3 & 2 \\ 1 & 3 & 2 \\ 0 & 5 & 1 \end{bmatrix}$

Solution: We have $A - \lambda I = \begin{bmatrix} 1-\lambda & 3 & 2 \\ 1 & 3-\lambda & 2 \\ 0 & 5 & 1-\lambda \end{bmatrix}$ with determinant

$$\begin{aligned} f_A(\lambda) &= (1-\lambda)(3-\lambda)(1-\lambda)+3\cdot2\cdot0+2\cdot1\cdot5-0\cdot(3-\lambda)\cdot2-5\cdot2\cdot(1-\lambda)-(1-\lambda)\cdot1\cdot3 \\ &= -(\lambda-1)(\lambda-3)(\lambda-1)+0+10-0+10(\lambda-1)+3(\lambda-1) \end{aligned}$$

Let's pause here to point out how critical it is to make creative use of your high school algebra. When you see the common factor $\lambda - 1$, always keep it factored that way for as long as you can. Resist the temptation to "multiply everything out". On the other hand, from the IMT, we know that A, having two identical rows, will have $\theta = 0$ as an eigenvalue, so we expect a factor of λ to appear:

$$\begin{aligned} f_A(\lambda) &= (\lambda - 1)\left[-(\lambda - 3)(\lambda - 1)\right] + 3(\lambda - 1) + 10\lambda \\ &= (\lambda - 1)\left[-\lambda^2 + 4\lambda - 3\right] + 3(\lambda - 1) + 10\lambda \\ &= (\lambda - 1)\left[-\lambda^2 + 4\lambda\right] + 10\lambda \\ &= \lambda\left[(\lambda - 1)(-\lambda + 4)\right] + 10\lambda \\ &= -\lambda\left[\lambda^2 - 5\lambda - 6\right] \\ f_A(\lambda) &= -\lambda(\lambda - 6)(\lambda + 1) \end{aligned}$$

So the eigenvalues of A are $\lambda_1 = 6$, $\lambda_2 = 0$ and $\lambda_3 = -1$. A few quick row reductions gives us the corresponding eigenvectors: $\mathbf{v}_1 = (1, 1, 1)$, $\mathbf{v}_2 = (-\frac{7}{5}, -\frac{1}{5}, 1)$ and $\mathbf{v}_3 = (-\frac{2}{5}, -\frac{2}{5}, 1)$. (Of course, I'd rather choose basis $\{(1, 1, 1),\ (7, 1, -5),\ (2, 2, -5)\}$ since I avoid fractions.)

Example 8: Find all the eigenvalues of $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$

Solution: We have $A - \lambda I = \begin{bmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & -1 \\ 1 & -1 & -\lambda \end{bmatrix}$ with determinant

$$\begin{aligned} f_A(\lambda) &= -\lambda^3 - 1 - 1 + \lambda + \lambda + \lambda \\ &= -\left(\lambda^3 - 3\lambda + 2\right) \end{aligned}$$

Okay, now we need to factor a cubic polynomial. That's a nasty problem in general, but you know that your professor gives you examples that come out nice. So look for integer roots of $\lambda^3 - 3\lambda + 2$. From the basic theory of polynomials, you know that the sum of the roots is zero (the negative of the coefficient of λ^2) and the product of the roots is -2 (the negative of the constant term, since we ignored the -1 in front of λ^3.) So we try $\theta = -2, -1, 1, 2$ and find that $\lambda_1 = 1$ and $\lambda_3 = -2$ are both eigenvalues. Since the eigenvalues must sum to zero here, we recover $\lambda_2 = 1$ also:

$$f_A(\lambda) = -(\lambda + 2)(\lambda - 1)(\lambda - 1).$$

So we have one eigenvalue of multiplicity one:

- $\lambda_3 = -2$ with corresponding eigenvector $(-1, 1, 1)$

and one eigenvalue of algebraic multiplicity two:

- $\lambda_1 = \lambda_2 = 1$ with linearly independent eigenvectors $(1, 1, 0)$ and $(1, 0, 1)$.

IMPORTANT: Once we compute the characteristic polynomial $f_A(\lambda)$ (a polynomial in λ of degree n), we often benefit from remembering that

- the coefficient of λ^{n-1} is always -1 times the **sum** of all n eigenvalues:

$$\mathrm{trace\,A} = \lambda_1 + \lambda_2 + \ldots + \lambda_\mathrm{n}$$

- the constant term is always the **product** of all n eigenvalues:

$$\det A = \lambda_1\lambda_2 \ldots \lambda_n$$

Since we found a total of three linearly independent eigenvectors for this 3×3 matrix, this matrix is *diagonalizable*.

Example 9: Find all the eigenvalues of $A = \begin{bmatrix} 3 & 0 & 7 \\ 1 & 8 & -6 \\ 1 & 0 & 9 \end{bmatrix}$ *[HINT: What is special about the second column?]*

SOLUTION: We first see that $\mathbf{e}_2$ is an eigenvector $A\mathbf{e}_2$ is just the second column of A. So $A\mathbf{e}_2 = 8\mathbf{e}_2$ and we have one eigenvalue: let's say $\lambda_3 = 8$ (even though we don't know yet whether it is biggest or smallest or neither).

As usual, we could compute $A - \lambda I = \begin{bmatrix} 3-\lambda & 0 & 7 \\ 1 & 8-\lambda & -6 \\ 1 & 0 & 9-\lambda \end{bmatrix}$ and grind out its determinant. But we instead use the two critical facts about eigenvalues in the box above. We have

$$\begin{aligned} \mathrm{trace\,A} &= 20 = \lambda_1 + \lambda_2 + \lambda_3 \\ \det A &= 3 \cdot 8 \cdot 9 + 0 + 0 - 1 \cdot 8 \cdot 7 - 0 - 0 \\ &= 8 \cdot (3 \cdot 9 - 1 \cdot 7) \\ 8 \cdot 20 &= \lambda_1\lambda_2\lambda_3 \end{aligned}$$

So the two missing eigenvalues satisfy

$$\lambda_1 + \lambda_2 = 12, \qquad \lambda_1\lambda_2 = 20.$$

It is now elementary to see that the eigenvalues, in decreasing order, are

$$\lambda_1 = 10, \qquad \lambda_2 = 8, \qquad \lambda_3 = 2.$$

Example 10: Find all the eigenvalues of

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

SOLUTION: The eigenvalues are just the entries on the main diagonal. we sort them in decreasing order:

$$\lambda_1 = 4, \qquad \lambda_2 = 3, \qquad \lambda_3 = 2, \qquad \lambda_4 = 1 \; .$$

Example 11: Find all the eigenvalues of

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

SOLUTION: Again, for an upper- or lower-triangular matrix, the eigenvalues are exactly the entries on the main diagonal:

$$\lambda_1 = \lambda_2 = 5, \qquad \lambda_3 = 4, \qquad \lambda_4 = \lambda_5 = 1, \qquad \lambda_6 = 0 \; .$$

Example 12: Find all the eigenvalues of

$$A = \begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 7 \end{bmatrix}$$

SOLUTION: The eigenvalues are just the entries on the main diagonal:

$$\lambda_1 = 7, \qquad \lambda_2 = \lambda_3 = \lambda_4 = 5, \qquad \lambda_5 = \lambda_6 = 3 \; .$$

(I include this example because this matrix, in "almost diagonal form", is in "Jordan canonical form". This basic form is very important in the theory of eigenvectors and eigenvalues of linear transformations.)