Calculating eigenvalues

October 16, 2014

Contents

1 Introduction

We've seen that sometimes a nice linear transformation T (from a finite-dimensional vector space V to itself) can be *diagonalized*, and that doing this is closely related to finding *eigenvalues* of T. Here's the basic theorem about how to find them.

Theorem 1.1. *Suppose V is an n-dimensional vector space over a field F, and $T: V \to V$ is any linear map. Then there is polynomial*

$$p_T = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

(with coefficients in F), with the following properties:

1. *the eigenvalues of T are precisely the roots of p_T; and*
2. *$p_T(T) = 0$.*

The polynomial p_T is called the *characteristic polynomial* of T. (The two properties listed in the theorem don't specify p_T completely.) The purpose of these notes is to write down as simply as possible how to compute p_T. The longer *Notes on generalized eigenvalues* will repeat this more slowly, and draw some more serious conclusions.

2 Polynomials

Recall from pages 10 and 22–23 of the text that $\mathcal{P}(F)$ is the (F-vector space) of all polynomials with coefficients in F. The *degree* of a nonzero polynomial is the highest power of x appearing with nonzero coefficient; the polynomial zero is defined to have degree $-\infty$ (to make the formula $\deg(pq) = \deg(p) + \deg(q)$ work when one of the factors is zero). We write

$$\mathcal{P}_m(F) = \{a_m x^m + \cdots + a_1 x + a_0 \mid a_i \in F\} \tag{2.1}$$

for the $(m+1)$-dimensional subspace of polynomials of degree less than or equal to m. A polynomial is called *monic* if its leading coefficient is 1; therefore

$$1, \qquad x^2 - 7x + 1/2, \qquad x - \pi$$

are monic polynomials (with real coefficients), but $2x + 1$ is not monic.

3 Characteristic polynomial: theory

The proof of Theorem 1.1 is an algorithm for computing p_T. Here it is.

Proof. We proceed by induction on $n = \dim V$, If $n = 0$, then $V = \{0\}$, so $T = 0$. In this case we are forced to define $p_T = 1$, the unique monic polynomial of degree 0. The polynomial p_T has no roots, and T has no eigenvalues. Also $p_T(T) = I_V = 0$; so both conclusions of the theorem are true.

Suppose therefore that $n \geq 1$, and that we know the conclusion of the theorem for vector spaces of dimension strictly less than n. Since V has dimension at least 1, it is nonzero; so we can choose a nonzero vector

$$0 \neq v_0 \in V. \tag{3.1a}$$

Define

$$v_j = T^j v_0 \qquad (j = 0, 1, 2, \dots) \tag{3.1b}$$

The list $(v_0, v_1, v_2, \dots)$ is infinite, and V is finite-dimensional; so the list must be linearly dependent. According to the linear dependence lemma proved in class (very close to Lemma 2.4 on page 25 of the text), it follows that there must be an integer $m \geq 1$ with the property that

$$(v_0, \dots, v_{m-1}) \text{ is linearly independent,} \tag{3.1c}$$

and

$$v_m \in \operatorname{span}(v_0, \dots v_{m-1}). \tag{3.1d}$$

Then there is a unique expression

$$v_m = -a_0 v_0 - \cdots - a_{m-1} v_{m-1}. \tag{3.1e}$$

So far we have an m-dimensional vector space

$$U =_{\text{def}} \operatorname{span}(v_0, \dots, v_{m-1}),$$

and a degree m polynomial

$$p_U =_{\text{def}} x^m + a_{m-1} x^{m-1} + \cdots + a_0. \tag{3.1f}$$

Now extend the linearly independent list $(v_0, \dots, v_{m-1})$ to a basis

$$(v_0, \dots, v_{m-1}, e_1, \dots, e_{n-m})$$

of V. Now Tv_i is a linear combination of other v_j. More precisely,

$$Tv_j = \begin{cases} v_{j+1} & (0 \leq j < m-1) \\ -a_0 v_0 - \cdots - a_{m-1} v_{m-1} & (j = m-1) \end{cases} \tag{3.1g}$$

Consequently the matrix of T in the basis $(v_0, \dots, v_{m-1}, e_1, \dots, e_{n-m})$ is of the form

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}. \tag{3.1h}$$

Here A is an $m \times m$ matrix (of a very simple form given by (3.1g)); B is an $m \times (n-m)$ matrix; and C is an $(n-m) \times (n-m)$ matrix. Define

$$p_A(x) = x^m + a_{m-1} x^{m-1} + \cdots + a_0,$$

a monic polynomial of degree $m \geq 1$ in $\mathcal{P}(F)$.

By inductive hypothesis, we know how to compute the characteristic polynomial p_C of the matrix C; it is a monic polynomial of degree $n - m$. Define

$$p_T = p_A(x) p_C(x),$$

a monic polynomial of degree n. This is by definition the characteristic polynomial of T. I'll postpone the proof that p_T has the two desired properties to the later/longer *Notes on generalized eigenvalues.* The key ingredient is

Proposition 3.2. *Suppose*

$$T = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

is an $n \times n$ block upper-triangular matrix, with A an $m \times m$ matrix, B an $m \times (n-m)$ matrix, and C an $(n-m) \times (n-m)$ matrix.

1. *The set of eigenvalues of T is the union of the set of eigenvalues of A and the set of eigenvalues of C.*

2. *Suppose p_1 and p_2 are polynomials such that $p_1(A) = 0$ and $p_2(C) = 0$. Then $(p_1 p_2)(T) = 0$.*

(You might think about how to prove this proposition; it needs only things you already know.) $\square$

4 Characteristic polynomial: practice

Here is a description of the algorithm for computing the characteristic polynomial. We begin with a linear transformation T on an n-dimensional vector space. We need also a basis $(e_1, \ldots, e_n)$ of V; this is used as a source for various starting vectors in the algorithm. What the algorithm produces is a new basis

$$(v_0^1, v_1^1, \ldots, v_{m_1-1}^1, v_0^2, v_1^2, \ldots, v_{m_2-1}^2, v_0^3, \ldots v_{m_k-1}^k). \tag{4.1a}$$

Here each m_i is at least 1. That is, the new basis consists of k strings of positive lengths $(m_1, m_2, \ldots, m_k)$; so the sum of the m_i is equal to n. We'll choose the basis so that the action of T in the basis is very simple:

$$Tv_j^i = \begin{cases} v_{j+1}^i & (0 \le j < m_i - 1) \\ -a_0^i v_0 - \cdots - a_{m_i-1}^i v_{m_i-1}^i + (\text{junk}) & (j = m_i - 1) \end{cases} \tag{4.1b}$$

Here (junk) means a linear combination of various basis vectors $v_{j'}^{i'}$ from earlier strings; that is, with $i' < i$. The ith string defines a polynomial

$$p^i(x) = x^{m_i} + a_{m_i-1}^i x^{m_i-1} + \cdots + a_0^i, \tag{4.1c}$$

monic of degree m_i. The characteristic polynomial of T is

$$p_T(x) = p^1(x) \cdot p^2(x) \cdot \cdots \cdot p^k(x). \tag{4.1d}$$

Here is how to find such a basis.

The first string of basis vectors $(v_0^1, v_1^1, \cdots, v_{m_1-1}^1)$ is calculated as explained in the previous section: start with any nonzero vector v_0^1 in V (like the first basis vector e_1), and define

$$v_j^1 = T^j v_0^1.$$

The list (v_0^1) is linearly independent, because v_0^1 was chosen to be nonzero. For each successive j, check whether the v_j^1 is a linear combination of the preceding terms. If it is not, add it to the list, which continues to be linearly independent. If it *is* a linear combination, then declare $j = m^1$; the linear combination equation

$$v_{m_1}^1 = -a_0^1 v_0 - \cdots - a_{m_1-1}^1 v_{m_1-1}^1$$

is exactly what is claimed in (4.1b) for $i = 1$. (There is no (junk) yet because there is no $i' < 1$.)

If $m_1 = n$, then the algorithm stops here. If $m_1 < n$, then the first string $(v_0^1, \ldots, v_{m_1-1}^1)$ does *not* span V. Define

$$v_0^2 = \text{first vector in } (e_1, \ldots, e_n) \text{ not in } \operatorname{span}(v_0^1, \ldots, v_{m_1-1}^1) \tag{4.1e}$$

Now define

$$v_j^2 = T^j v_0^2.$$

The list $(v_0^1, \ldots, v_{m_1-1}^1, v_0^2)$ is linearly independent by the choice of v_0^2. For each successive j, check whether the v_j^2 is a linear combination of the preceding terms in the list. If it is not, add it to the list, which continues to be linearly independent. If it *is* a linear combination, then declare $j = m^2$; the linear combination equation

$$v_{m_2}^2 = -a_0^2 v_0^2 - \cdots - a_{m_1-1}^2 v_{m_2-1}^2 + (b_0^1 v_0^1 + b_1^1 v_1^1 + \cdots b_{m_1-1}^1 v_{m-1}^1)$$

is exactly what is claimed in (4.1b) for $i = 2$. (The term in parentheses is (junk). The coefficients b_j^1 give some entries of the matrix B from (3.1h).)

Now our two strings make a linearly independent list

$$(v_0^1, \ldots, v_{m_1-1}^1, v_0^2, \ldots, v_{m_2-1}^2)$$

of length $m_1 + m_2$. If $m_1 + m_2 = n$, the algorithm stops here. If not, we can define

$$v_0^3 = \text{first vector in } (e_i) \text{ not in } \operatorname{span}(v_0^1, \ldots, v_{m_2-1}^2). \tag{4.1f}$$

And so on. At each step, our linearly independent list gets strictly longer; so after at most n steps, we end up with a basis of V.

5 Examples

For almost all linear transformations T, and almost all choices of nonzero v_0^1, the n vectors

$$(v_0^1, Tv_0^1, \ldots, T^{n-1}v_0^1)$$

are linearly independent, and the algorithm for calculating the characteristic polynomial has just one interesting step: solving the system of n equations in n unknowns (the various a_j^1)

$$T^n v_0^1 = -a_0^1 v_0^1 - a_1^1 T v_0^1 - \cdots - a_{n-1}^1 T^{n-1} v_0^1.$$

For example, if

$$T = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad v_0^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{5.1a}$$

then

$$v_1^1 = Tv_0^1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad v_2^1 = Tv_1^1 = \begin{pmatrix} 7 \\ 15 \end{pmatrix}.$$

The system of equations we need to solve is

$$\begin{pmatrix} 7 \\ 15 \end{pmatrix} = -a_0^1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} - a_1^1 \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

or

$$-a_0^1 - a_1^1 = 7, \quad -3a_1^1 = 15.$$

Solution is $a_1^1 = -5$, $a_0^1 = -2$; so the characteristic polynomial is

$$p_T(x) = x^2 - 5x - 2. \tag{5.1b}$$

Eigenvalues are therefore $(5 \pm 2\sqrt{3})/2$, at least as long as the characteristic of the field is not 2. If the characteristic of F is 2, then

$$T = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad p_T(x) = x^2 - x,$$

and the eigenvalues are 0 and 1.

Here is an example of the "rare" case of more than one string. Take

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}. \tag{5.2a}$$

Choosing v_0^1 to be the first standard basis vector as the algorithm directs, we find

$$v_0^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad v_1^1 = Tv_0^1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \qquad v_2^1 = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 2 \end{pmatrix}.$$

The first two are linearly independent, but the third is a linear combination of them: the linear independence test equation

$$\begin{pmatrix} 1 \\ 2 \\ 2 \\ 2 \end{pmatrix} = -a_0^1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} - a_1^1 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

has unique solution

$$a_0^1 = 1, \qquad a_1^1 = -2.$$

Therefore $m_1 = 2$, and

$$p^1(x) = x^2 - 2x + 1 = (x-1)^2. \tag{5.2b}$$

(So far the only eigenvalue we've found is 1.)

Because $m_1 = 2$ is strictly smaller than 4, the algorithm continues. We choose v_0^2 to be the first standard basis vector not in the span of our list so far (as the algorithm directs) and calculate

$$v_0^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad v_1^2 = Tv_0^2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad v_2^2 = Tv_1^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

The list $(v_0^1, v_1^1, v_0^2, v_1^2)$ is linearly independent, but the fifth vector is a linear combination of the first four:

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = -a_0^2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} - a_1^2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + (b_0^1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + b_1^1 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}).$$

The last two terms in parentheses are (junk). The unique solution of this equation is

$$a_0^2 = -1, \qquad a_1^2 = 0, \qquad b_0^1 = 0, \qquad b_1^1 = 0.$$

Therefore $m_2 = 2$, and the second factor of the characteristic polynomial is

$$p^2(x) = x^2 - 1 = (x+1)(x-1). \tag{5.2c}$$

Since $m_1 + m_2 = 4$, the algorithm is finished:

$$p_T(x) = p^1(x)p^2(x) = (x-1)^2(x+1)(x-1). \tag{5.2d}$$

The complete set of eigenvalues of T is $\{1, -1\}$.

You should check that you can calculate that the $+1$ eigenspace of T has basis

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}. \tag{5.2e}$$

So far this is true for any field at all (since we didn't need to divide by anything to solve the equations along the way).

If the characteristic of F is two, then the eigenvalue -1 is equal to $+1$; so we have calculated the only eigenspace, and it's two-dimensional. In particular, T is *not* diagonalizable (since there is no basis of eigenvectors: we need four, but we have only two).

If the characteristic of F is not two, then the eigenvalue -1 is different from 1. You should check that the -1 eigenspace of T has basis

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}; \tag{5.2f}$$

solving the equations for the null space of $T+I$ involves dividing by 2, which is why this part only works if the characteristic is not two. So altogether we have three linearly independent eigenvectors: still smaller than four, so T is *not* diagonalizable even in characteristic not two, and even though the characteristic polynomial is a product of linear factors.