

Math 642:550 — Summer 2006

MTTh 6:00–8:30 PM Hill 705

Prof. Bumby

Supplement 4, Finding eigenvectors

1. Introduction

Important **theoretical** results are that: (1) eigenvalues are characterized as the roots of the **characteristic polynomial**; and (2) the characteristic polynomial may be expressed as a determinant. If you have an eigenvalue λ , then the eigenvectors of a matrix M with eigenvalue λ are the **nonzero** vectors in the nullspace of $M - \lambda I$ (the whole nullspace is called the **eigenspace** of M for the eigenvalue λ , but only the nonzero vectors are called eigenvectors). This eigenspace, like all vector spaces, is usually identified by giving a basis. For a one dimensional space, any nonzero vector in the space gives a basis and we sometimes say that such a vector is “the eigenvector” with the understanding that any nonzero multiple of the vector would do just as well. For spaces of at least two dimensions, the relation between different bases is harder to see, but this should not stop you from choosing a basis to describe the space.

If you are unable to find any nonzero vectors in the nullspace of $M - \lambda I$ while claiming that λ is an eigenvalue of M , then **you have made a mistake somewhere**. Mistakes in this work are common, so you should organize your work to include checks that allow you to discover your mistakes before showing them to anyone else.

Contrary to the advice at the bottom of page 236 of the textbook, this is **not the only way** to find eigenvalues or eigenvectors. There are many examples where some eigenvalues can be found **without** first finding the characteristic polynomials. There are even examples where eigenvalues are found by **first** finding the eigenvector v and then noting the scale factor relating v and Mv . Again, it should also be noted that the failure of the attempt to find such a scale factor is a **proof** that v is not an eigenvector of M .

2. Triangular matrices

As noted in Example 3 of Section 5.1 of the textbook, the determinant of a triangular matrix is the product of the diagonal entries, so that the characteristic polynomial of

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

is $(1 - \lambda)(2 - \lambda)(-4 - \lambda)$. There is **no reason to expand this product**; the roots of the characteristic polynomial are more valuable than the coefficients. The eigenvalues are 1, 2 and -4 . We now turn our attention to finding eigenvectors. This can be done by separately computing the nullspace of $M - \lambda I$ for each λ on the diagonal of M (although there are other methods that will be described in later supplements).

The eigenvectors for $\lambda = 1$ lie in the nullspace of

$$M_1 = \begin{bmatrix} 0 & -1 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & -5 \end{bmatrix}.$$

This matrix is already in echelon form, so we read the rows from the bottom to find properties of v with $M_1 v = 0$. Since -5 times the third entry of v is zero, the third entry of v is zero. The second row now tells us that the second entry of v is zero. Now, the first row tells us only that $0 = 0$. **There is no restriction on the first entry.** The nullspace consists of all multiples of $(1, 0, 0)$. We may take any nonzero multiple of this as an eigenvector. It is usually best to take $(1, 0, 0)$ itself. Although this is the easiest eigenvector to find, it is often overlooked because few matrices with a zero first column are met in a first course on Linear Algebra.

The eigenvectors for $\lambda = 2$ lie in the nullspace of

$$\begin{bmatrix} -1 & -1 & 3 \\ 0 & 0 & 5 \\ 0 & 0 & -6 \end{bmatrix}.$$

As before, the third row says that the third entry of an eigenvector is zero. This time, the second row says nothing new. Then, the first row says that the first entry of the eigenvector is the negative of the second entry. Any nonzero vector with this property may be used. To keep the expression simple, we take $(1, -1, 0)$.

The eigenvectors for $\lambda = -4$ lie in the nullspace of

$$\begin{bmatrix} 5 & -1 & 3 \\ 0 & 6 & 5 \\ 0 & 0 & 0 \end{bmatrix}.$$

Equations giving ratios of entries in the eigenvector are easily found. If the third entry is taken to be -30 , the other entries will be integers. Indeed, the eigenvector is $(23, 25, -30)$ — or any nonzero multiple. Since the scaling of an eigenvector is arbitrary, a rational vector is usually scaled to have integer entries.

2.1 Exercises

A Find the eigenvalues and eigenvectors of

$$\begin{bmatrix} 2 & 5 & -3 \\ 0 & -3 & 12 \\ 0 & 0 & 1 \end{bmatrix}.$$

B Find the eigenvalues and eigenvectors of

$$\begin{bmatrix} 3 & 5 & -2 \\ 0 & -1 & 10 \\ 0 & 0 & 0 \end{bmatrix}.$$

3. Projections Example 2 in Section 5.1 uses the geometric description of projections to identify the nullspaces of P and of $P - I$ and to show that their sum is the whole space. An approach that applies more generally is to use the geometry to show that a projection matrix P satisfies $P^2 = P$. That is, any vector Px in the range of the projection (which is the column space of P) must be projected into itself, so that $P^2x = Px$ for all x . Since the matrix $P^2 - P$ takes each vector to zero, applying it to the standard basis shows that each column is the zero vector, so every entry is zero,

If v is an eigenvector with eigenvalue λ , then $Pv = \lambda v$ and $P^2v = P(\lambda v) = \lambda(Pv) = \lambda(\lambda v) = \lambda^2v$. Thus, $\lambda^2v = \lambda v$ or $(\lambda^2 - \lambda)v = 0$. Since v must be a nonzero vector, $\lambda^2 - \lambda = 0$. The only solutions of this are $\lambda = 0$ and $\lambda = 1$.

Since $P(I - P)x = 0$ for all vectors x , if $y = (I - P)x$, then $P y = P(I - P)x = 0$, so y is either the zero vector or an eigenvector with eigenvalue 0. Similarly, each Px is either the zero vector or an eigenvector with eigenvalue 1. However x is the sum of these two vectors. Thus the sum of the two eigenspaces is the whole space, and there is a basis of eigenvectors. Moreover, applying Gaussian elimination to P leads to both a basis for the column space (the eigenspace for $\lambda = 1$) and the nullspace (the eigenspace for $\lambda = 0$), so there is no need to do a second calculation starting from $I - P$.

3.1 Example

We find the eigenspaces of the projection matrix

$$P = \frac{1}{16} \begin{bmatrix} 4 & 2 & -2 & 2 & 6 \\ 2 & 5 & -5 & 5 & -1 \\ -2 & -5 & 5 & -5 & 1 \\ 2 & 5 & -5 & 5 & -1 \\ 6 & -1 & 1 & -1 & 13 \end{bmatrix}.$$

The first step is to verify that this is a projection. This will be omitted here, but will be described in lecture.

In reducing the matrix, it is convenient to multiply by 16 to work with an integer matrix having eigenvalues 0 and 16 instead of the projection matrix.

Pivoting gives

$$\begin{bmatrix} 4 & 2 & -2 & 2 & 6 \\ 0 & 4 & -4 & 4 & -4 \\ 0 & -4 & 4 & -4 & 4 \\ 0 & 4 & -4 & 4 & -4 \\ 0 & -4 & 4 & -4 & 4 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 2 & -2 & 2 & 6 \\ 0 & 4 & -4 & 4 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This gives an LU factorization

$$16P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 & 0 \\ -\frac{1}{2} & -1 & 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 1 & 0 \\ -\frac{3}{2} & -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & -2 & 2 & 6 \\ 0 & 4 & -4 & 4 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Since the first two columns are pivot columns, a basis for the column space consists of the first two columns of P .

To find the nullspace, do the pivot steps on U corresponding to “back substitution” to obtain a matrix whose first two rows are

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

and whose remaining rows are zero. Interpreting this matrix as the coefficients of equations gives that a column whose entries are x_1, \dots, x_5 belongs to the nullspace if $x_1 = -2x_5$ and $x_2 = x_3 - x_4 + x_5$. Writing this in vector form with parameters x_3, x_4 and x_5 gives

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

The numerical column vectors appearing in this equation are a basis for the nullspace.

3.2 Related results

A generalization, equally easily proved, is that

$$Mv = \lambda v \text{ implies } f(M)v = f(\lambda)v$$

for all polynomials f . Specializing this to $f(x) = x + c$ gives the observation that adding c to the diagonal of M adds c to all eigenvalues **without changing the eigenvectors**. This is explored in exercises 5.1.3, 5.1.7, and 5.1.20. In particular, this shows that **there is nothing special about zero as an eigenvalue**, since there are closely related matrices having the same eigenvectors where one has zero as an eigenvalue and the other does not.

This also shows that, for all polynomials $f(x)$, $f(M) = 0$ implies that all roots of $f(x) = 0$ are eigenvalues of M . In addition to projections, characterized by $f(x) = x^2 - x$, there are **reflections**, characterized by $f(x) = x^2 - 1$, and **quarter turns**, characterized by $f(x) = x^2 + 1$. Examples of matrices satisfying these equations exist in all positive dimensions (except that quarter turns require at least two dimensions).

4. Markov matrices A **Markov matrix** M is an n by n matrix (all matrices in this part of the course are square matrices) with all entries nonnegative and the sum of the entries in each **column** equal to 1. The transpose of a Markov matrix M^T has the sum of each **row** equal to 1. If $v = (1, 1, \dots, 1)$, this gives $M^T v = v$. Thus, v is an eigenvector of M^T with eigenvalue 1. Since we have the eigenvector, we know that 1 is an eigenvalue of M^T .

It is always true that M and M^T have the same eigenvalues, so 1 is an eigenvalue of M . However, this time we don't know the eigenvector. However, **we are sure** that $M - I$ has a nontrivial nullspace, so we proceed to compute the nullspace. It can be shown that the eigenvector found in this way will have **all entries nonnegative**. If normalized so that the **sum of the entries** is 1, it is called the **limiting distribution** of the process represented by M . We make no attempt to find the other eigenvalues or eigenvectors. Statement **5I** of the textbook (p. 258) asserts that no eigenvalue exceeds 1 in absolute value (although some may be complex numbers). This is a special case of the Perron-Frobenius Theorem mentioned on page 261 of the textbook and described in a later supplement.

Taking each entry of a 3 by 3 matrix to be a fraction with denominator 10 allows $66^3 = 287,596$ — more than a quarter million — examples for which the computation will be manageable.

4.1 Example

Let

$$M = \begin{bmatrix} 0.2 & 0.4 & 0.6 \\ 0.5 & 0.3 & 0.1 \\ 0.3 & 0.3 & 0.3 \end{bmatrix}.$$

Then

$$M - I = \begin{bmatrix} -0.8 & 0.4 & 0.6 \\ 0.5 & -0.7 & 0.1 \\ 0.3 & 0.3 & -0.7 \end{bmatrix}.$$

To solve $(M - I)v = 0$, we first scale all rows by multiplying by 10 to get integer entries. Then, the second and third rows are multiplied by 8 to make their leading entries divisible by the pivot. This gives the array

$$\begin{bmatrix} -8 & 4 & 6 \\ 40 & -56 & 8 \\ 24 & 24 & -56 \end{bmatrix}$$

Pivoting:

$$\begin{bmatrix} -8 & 4 & 6 \\ 0 & -36 & 38 \\ 0 & 36 & -38 \end{bmatrix}$$

The second and third rows are now negatives of one another. Each gives the ratio of x_2 to x_3 and then the first row gives x_1 . In this case, we get $(x_1, x_2, x_3) = (23, 19, 18)$.

Note that only the matrix of coefficients of the left side of the equation is written when we use Gaussian elimination to solve the equations to find the eigenvector. This is because we know that the right side will **always be zero**, no matter what row operations we perform. (On the other hand, I have seen students insist on writing a right side and accidentally introduce a nonzero quantity in one of the steps.)

5. Exercises

C Find the limiting distribution of the Markov matrix

$$\begin{bmatrix} 0.1 & 0.5 & 0.2 \\ 0.4 & 0.2 & 0.7 \\ 0.5 & 0.3 & 0.1 \end{bmatrix}.$$

D Find the limiting distribution of the Markov matrix

$$\begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.3 & 0.2 & 0.7 \\ 0.4 & 0.3 & 0.1 \end{bmatrix}.$$

6. Companion matrices

Suppose that a matrix has the form

$$C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ c & b & a \end{bmatrix}.$$

Such a matrix has an eigenvector of the form $(1, x, x^2)$ since

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ c & b & a \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} = \begin{bmatrix} x \\ x^2 \\ c + bx + ax^2 \end{bmatrix} = \begin{bmatrix} x \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} x$$

if $x^3 = ax^2 + bx + c$. If this polynomial has distinct roots, then we have found a basis of eigenvectors, so the matrix can be diagonalized (as discussed in Section 5.2 of the textbook). All three roots of the polynomial are eigenvalues, so this polynomial divides the characteristic polynomial. Since the polynomials have the same degree, the quotient is a constant, and consideration of the coefficient of x^3 shows that the characteristic polynomial is $-x^3 + ax^2 + bx + c$.

The matrix C is called the **companion matrix** of the polynomial $x^3 - ax^2 - bx - c$. There is an obvious generalization giving an n by n companion matrix of a polynomial of degree n .

Remark. A related matrix is the **Leslie matrix** arising in models of population growth. (A special case makes an appearance in Exercise 5.3.2 of the textbook.) The basic idea is to consider a vector v whose i^{th} entry is the count of individuals in the i^{th} year of life at a particular time and to form a matrix M such that Mv describes a similar vector one year later. The next year's count consists of **survivors** who are one year older and **newborn** individuals in their first year of life. The survivors contribute entries between 0 and 1 to the sub-diagonal of M , and the newborns give nonnegative entries in the first row in a column indicating the age of the parent.

In general, if C is the companion matrix of a monic polynomial $P(x)$ of degree n , expansion of $\det(C - \lambda I)$ by its first column shows that the characteristic polynomial of C is **always** $(-1)^n P(x)$. Similarly, consideration of the rows of $C - \lambda I$ shows that the eigenvectors **always** have the form $(1, x, \dots, x^{n-1})$. When the eigenvalues are distinct, the matrix whose columns are the eigenvectors is a Vandermonde matrix. Since eigenvectors for different eigenvalues are **always** linearly independent, this gives another proof that the Vandermonde matrix is nonsingular.

By contrast, the companion matrix of a polynomial with repeated roots **never** has enough eigenvectors to form a basis. However, there is a simple construction leading to **generalized eigenvectors** in this case. First, note that

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ c & b & a \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} = \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ c + bx + ax^2 - x^3 \end{bmatrix}$$

for **all** x . This may be written more compactly as $Cv(x) = xv(x) + u(x)$. Since this is an identity, it can be differentiated to give $Cv'(x) = xv'(x) + v(x) + u'(x)$. Specializing x to a repeated root λ of $P(x)$ gives $Cv(\lambda) = \lambda v(\lambda)$ and $Cv'(\lambda) = \lambda v'(\lambda) + v(\lambda)$ since $u'(\lambda) = u(\lambda) = 0$ when λ is a repeated root. Thus, multiplication by $C - \lambda I$ takes $v'(\lambda)$ to $v(\lambda)$ and $v(\lambda)$ to the zero vector. This shows that $v'(\lambda)$ is a generalized eigenvector.

7. Exercise E

Let

$$C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$$

Find the characteristic polynomial of C (once you have one root, the rest of the factorization will be easy, and you should be able to find one root immediately) and all eigenvectors. Then find the generalized eigenvector $v'(\lambda)$ for the double root of the characteristic polynomial.

8. The Power Method

The basic computational method for finding eigenvalues and eigenvectors is the **power method**. Section 7.3 is devoted to this method and several of its refinements, and we will return to it later in the course. For now, it will suffice to illustrate the method. As an example, consider

$$M = \begin{bmatrix} 0.4 & 0.2 & 0.7 \\ 0.5 & 0.3 & 0.1 \\ 0.1 & 0.5 & 0.2 \end{bmatrix}.$$

This has the same rows as the matrix appearing in the Markov matrix exercise, in a different order; it is still a Markov matrix, but you will find that the limiting distribution is quite different. Iteration of $v_{n+1} = Mv_n$ for various choices of v_0 will be illustrated in lecture. Since this is a Markov matrix, the sum of the entries in v_n is constant. We can check the result of the power method for this example because we are able to find an exact solution that is rational. If we multiply by the denominator (124 in this case), our computed values should be very close to the integer numerators.

With other matrices, it would be necessary to rescale each v_n . The scale factors will converge to the dominant eigenvalue if there is one. A simple scaling is to make the first entry of the vector equal to 1. This is adequate for interactive use on a calculator, since it will be easy to detect if it is inappropriate (for example, if that entry of the eigenvalue is much smaller than the other entries). If all entries of the matrix are positive, the iteration converges reliably to a positive eigenvector. For other matrices, the eigenvalues of largest absolute value may be complex. In such cases, the iteration will not converge, although in some cases it will appear to cycle. Some examples are

$$A = \begin{bmatrix} 5 & -2 & 1 \\ 3 & 2 & -1 \\ -2 & 3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 5 & 2 & 1 \\ 3 & 2 & 1 \\ 2 & 3 & 4 \end{bmatrix}.$$

Matrix A has eigenvalues approximately $3.3 + 3.0i$, of absolute value 4.46 that slightly dominate the real eigenvalue of 4.42. Matrix B has three real eigenvalues (approximately 7.64, 2.92, and 0.45). The second iteration clearly converges, with the analysis of the method indicating that there will be an additional two places of accuracy every five iterations. For the first matrix, the lack of convergence will be clear rather quickly, but the power method calculation suggests no remedy. When we return to the topic in Chapter 7, better methods for dealing with this matrix will be met.

End of Supplement