10 Eigenvalues and Eigenvectors Fall 2003

Introduction

To introduce the concepts of *eigenvalues* and *eigenvectors*, we consider first a three-dimensional space with a Cartesian coordinate system. Consider a vector from the origin O to a point P; call this vector **a**. The components of **a** are (a_1, a_2, a_3). Alternatively, we could say that point P has coordinates (a_1, a_2, a_3).

We can apply a linear transformation to vector **a** to obtain another vector **z**. For any linear transformation, each component of **z** is some linear combination of the components of **a**. This relationship can be expressed as

$$\begin{aligned} z_1 &= S_{11}a_1 + S_{12}a_2 + S_{13}a_3, \\ z_2 &= S_{21}a_1 + S_{22}a_2 + S_{23}a_3, \\ z_3 &= S_{31}a_1 + S_{32}a_2 + S_{33}a_3. \end{aligned} \qquad \text{or} \qquad z_i = \sum_{j=1}^{3} S_{ij}a_j. \tag{1}$$

This can be expressed more compactly by use of matrix language. We define:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad \mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix}. \tag{2}$$

Then

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} S_{11}a_1 + S_{12}a_2 + S_{13}a_3 \\ S_{21}a_1 + S_{22}a_2 + S_{23}a_3 \\ S_{31}a_1 + S_{32}a_2 + S_{33}a_3 \end{pmatrix}, \tag{3}$$

or simply $\mathbf{z} = \mathbf{S}\,\mathbf{a}$. (4)

In general, the *direction* of vector **z** is different from that of **a**. But there may be special cases where **z** has the *same* direction as **a.** For example, suppose the transformation **S** represents a rotation of vector **a** about some fixed axis. If the direction of **a** happens to coincide with this axis, then **a** is not changed by this transformation, and **z** = **a**.

More generally, if **z** has the same *direction* (but not necessarily the same magnitude) as **a**, then **z** must be a scalar multiple of **a.** That is, $\mathbf{z} = \lambda\mathbf{a}$, where λ is a *scalar*. In that case,

$$\mathbf{Sa} = \lambda\mathbf{a}, \tag{5}$$

or

$$\begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \lambda \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \end{pmatrix}. \tag{6}$$

In this case, the result of the transformation **S** applied to the vector **a** is another vector having the *same direction* as **a**. A vector **a** for which Eqs. (5) and (6) are valid is called an *eigenvector* of the transformation **S**, and the scalar λ is the corresponding *eigenvalue*.

Finding Eigenvectors

Several questions immediately arise:

(1) How do we know that eigenvectors *exist*, for any given transformation **S**?

(2) If eigenvectors (and their corresponding eigenvalues) *do* exist, how can we find them?

(3) Can there be more than one eigenvector for a given transformation **S**? If so, how are the different eigenvectors and eigenvalues related?

We'll now try to answer these questions. First, Eq. (6) can be combined with Eq. (3) and re-arranged as follows:

$$\begin{pmatrix} S_{11}a_1 + S_{12}a_2 + S_{13}a_3 \\ S_{21}a_1 + S_{22}a_2 + S_{23}a_3 \\ S_{31}a_1 + S_{32}a_2 + S_{33}a_3 \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \end{pmatrix}. \tag{7}$$

Equating corresponding elements in Eq. (7) and re-arranging, we obtain the set of three scalar equations:

$$\begin{aligned} (S_{11} - \lambda)a_1 + S_{12}a_2 + S_{13}a_3 &= 0, \\ S_{21}a_1 + (S_{22} - \lambda)a_2 + S_{23}a_3 &= 0, \\ S_{31}a_1 + S_{32}a_2 + (S_{33} - \lambda)a_3 &= 0. \end{aligned} \tag{8}$$

If an eigenvector **a** exists, its components (a_1, a_2, a_3) and its eigenvalue λ must satisfy Eqs. (8). This is a set of three simultaneous, linear, *homogeneous* equations. (I.e., every term contains an a to the first power, and there are no terms that do not contain any of a_1, a_2, or a_3.)

As was mentioned in Section 9, page 9-5, such a set of equations *always* has the trivial solution $a_1 = a_2 = a_3 = 0$. A fundamental theorem of linear algebra states that *non-trivial* solutions of Eqs. (8) exist if, and only if, the *determinant* of the system is zero. That is, the necessary and sufficient condition for the existence of non-trivial solutions is

$$\begin{vmatrix} S_{11}-\lambda & S_{12} & S_{13} \\ S_{21} & S_{22}-\lambda & S_{23} \\ S_{31} & S_{32} & S_{33}-\lambda \end{vmatrix} = 0. \tag{9}$$

For any given transformation **S**, this is a third-degree algebraic equation for λ. It has three roots, some of which may be complex. For each of the three values of λ, Eqs. (8) are a set of three simultaneous equations for the components (a_1, a_2, a_3) of the corresponding eigenvector **a**, with a different set of components for each value of λ.

Note that the components of each eigenvector are not completely determined by Eqs. (8) because the equations are *homogeneous*. For any given λ, if (a_1, a_2, a_3) is one solution, then any scalar multiple, such as $(2a_1, 2a_2, 2a_3)$, is also a solution. So for each λ, a solution of Eqs. (8) gives the *direction* of the corresponding eigenvector but not its *magnitude*. Thus when we solve these equations, we are at liberty to choose a value of one component (say a_1) arbitrarily, and then use Eqs. (8) to determine a_2 and a_3 as multiples of a_1,

Let the three eigenvalues be λ_1, λ_2, and λ_3, and let the corresponding eigenvectors be **a, b,** and **c.** Then

$$\begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \lambda_1\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \qquad \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \lambda_2\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

$$\begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \lambda_3\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}. \tag{10}$$

These may be combined into a single matrix equation. We form a square matrix **A** whose columns are the eigenvectors, and a diagonal matrix **L** whose diagonal elements are the eigenvalues: That is,

$$\mathbf{A} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}, \qquad \mathbf{L} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \tag{11}$$

From the first of Eqs. (10), the first column of the product **SA** is equal to $\lambda_1\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$.

This is also equal to the first column of the product $\mathbf{AL}$. Similarly, the other Eqs. (10) show that each of the other columns of $\mathbf{SA}$ is equal to the corresponding column of $\mathbf{AL}$. So all three of Eqs. (10) can be expressed compactly as

$$\begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad \text{or} \tag{12}$$

$$\mathbf{SA} = \mathbf{AL}. \tag{13}$$

Now something remarkable happens. Assuming the matrix $\mathbf{A}$ has an inverse $\mathbf{A}^{-1}$ (which we won't prove), we multiply both sides of Eq. (13) on the left by $\mathbf{A}^{-1}$. The result is

$$\mathbf{A}^{-1}\mathbf{SA} = \mathbf{A}^{-1}\mathbf{AL} = \mathbf{L}. \tag{14}$$

The operation $\mathbf{A}^{-1}\mathbf{SA}$ has transformed the matrix $\mathbf{S}$ into the *diagonal* matrix $\mathbf{L}$.

Normalization and Orthogonality

Because Eqs. (8) determine only the *ratios* of the components of each eigenvector, the *magnitudes* of the eigenvectors are arbitrary. It is often useful to multiply all components of an eigenvector by a scalar factor such that its magnitude is unity. Such a vector is said to be *normalized*. An eigenvector $\mathbf{a}$ is normalized if

$$\mathbf{a}^{\mathrm{T}}\mathbf{a} = 1, \qquad \text{or} \qquad a_1^{\,2} + a_2^{\,2} + a_3^{\,2} = 1, \tag{15}$$

and similarly for $\mathbf{b}$ and $\mathbf{c}$.

If $\mathbf{S}$ is a *symmetric* matrix (i.e., $\mathbf{S}^{\mathrm{T}} = \mathbf{S}$), and if λ_1, λ_2, and λ_3 are all different, then it can be shown that the eigenvectors are mutually *orthogonal* (perpendicular). That is, in this case the scalar product of any two different eigenvectors is zero:

$$\mathbf{a}^{\mathrm{T}}\mathbf{b} = \mathbf{a}^{\mathrm{T}}\mathbf{c} = \mathbf{b}^{\mathrm{T}}\mathbf{c} = 0. \tag{16}$$

If two or more eigenvalues are equal, the corresponding eigenvectors are *not* necessarily orthogonal, but in that case it is always possible to choose linear combinations that *are* orthogonal. If in addition the eigenvectors are normalized, they are said to form an *orthonormal* basis in the space.

Generalizations

***n* dimensions:** The above discussion has introduced basic concepts with reference to a space with three dimensions. But the restriction to three dimensions was not used at all in the development. Therefore these same concepts are valid in a space with any number of dimensions. In a space with n dimensions, the transformation $\mathbf{S}$ is an n by n matrix, there are n eigenvalues, and each corresponding eigenvector is a matrix with one column and n rows.

complex eigenvalues: Some of the eigenvalues may be complex numbers, even if all the elements of $\mathbf{S}$ are real. In that case the components of the eigenvectors are in general also complex, and some of the definitions need to be generalized. We generalize the definition of the scalar product of two vectors $\mathbf{a}$ and $\mathbf{b}$ to be $\mathbf{a}^+\mathbf{b}$, that is, the *adjoint* of $\mathbf{a}$ (denoted by $\mathbf{a}^+$) multiplied by $\mathbf{b}$. (The adjoint of a matrix, also called the *Hermitean conjugate*, is obtained by transposing rows and columns and taking the complex conjugate of each element.) In general, $\mathbf{a}^+\mathbf{b}$ is complex, but $\mathbf{a}^+\mathbf{a}$ is always real and nonnegative. (Can you prove this?) Also, $\mathbf{a}^+\mathbf{b}$ is the complex conjugate of $\mathbf{b}^+\mathbf{a}$. The normalization condition for eigenvectors becomes

$$\mathbf{a}^+\mathbf{a} = \mathbf{b}^+\mathbf{b} = \cdots = 1. \tag{17}$$

Hermitean matrices: A matrix that is equal to its adjoint is called a *Hermitean* matrix. (I.e., $\mathbf{S}$ is Hermitean if $\mathbf{S}^+ = \mathbf{S}$.) Note that a *symmetric* matrix is a special case of a Hermitean matrix; every symmetric matrix is Hermitean. Two important theorems about Hermitean matrices (which we won't prove) are:

(1) The eigenvalues of a Hermitean matrix $\mathbf{S}$ are always *real*, even when the elements of $\mathbf{S}$ are complex.

(2) If $\mathbf{S}$ is Hermitean, then eigenvectors corresponding to distinct (i.e., unequal) eigenvalues are orthogonal. That is, if $\mathbf{S}$ is Hermitean, and if $\mathbf{Sa} = \lambda_1\mathbf{a}$, $\mathbf{Sb} = \lambda_2\mathbf{b}$, and $\lambda_1 \neq \lambda_2$, then $\mathbf{a}^+\mathbf{b} = 0$.

unitary matrices: If the product of a matrix with its adjoint is the identity matrix, the matrix is said to be *unitary*. (I.e., $\mathbf{S}$ is unitary if $\mathbf{S}^+\mathbf{S} = \mathbf{S}\mathbf{S}^+ = \mathbf{I}$.) Two important theorems about unitary matrices (which we won't prove) are:

(1) The eigenvalues of a unitary matrix may be complex, but they always have magnitude 1. That is, if $\mathbf{S}$ is unitary, and if $\mathbf{Sa} = \lambda\mathbf{a}$, then $|\lambda| = 1$.

(2) If $\mathbf{S}$ is unitary, and if $\mathbf{b} = \mathbf{Sa}$, the vectors $\mathbf{a}$ and $\mathbf{b}$ have the same magnitude. That is, if $\mathbf{S}$ is unitary, and if $\mathbf{b} = \mathbf{Sa}$, then $\mathbf{b}^+\mathbf{b} = \mathbf{a}^+\mathbf{a}$.

Eigenvalues and Eigenvectors with Maple

To use Maple's matrix algebra tools, first load the *Linear Algebra* package with the command `with(linalg);` To find the eigenvalues of a matrix **A**, enter `eigenvals(A);` or `eigenvalues(A);` The Maple output is a *set* of the eigenvalues, in no particular order. If you enter `val := eigenvalues(A);` then you can call individual eigenvalues with `val[1];` (for the first value in the set), and so on. If you run Maple on a different computer, you may get the eigenvalues in a different order; they form a *set*, not a *list*.

The command for the eigenvectors of matrix **A** is `eigenvects(A);` or `eigenvectors(A);` The Maple output provides for the possibility that the eigenvalue equation may have multiple roots. If the eigenvalues are λ_1, λ_2, and so on, with multiplicities m_1, m_2, and so on, the Maple output has the form

$$[\lambda_1, m_1, \{[a_1, a_2, \cdots], [b_1, b_2, \cdots], \cdots\}], \quad [\lambda_2, m_2, \{[c_1, c_2, \cdots], [d_1, d_2, \cdots], \cdots\}], \cdots$$

In this expression, eigenvalue λ_1 has multiplicity m_1, and the following { } brackets enclose a *set* of m_1 orthogonal eigenvectors. And similarly for λ_2, λ_3, and so on.

More explicitly, the output expression is an entity with four levels of nested sets and lists. It is a *set* of structures, each enclosed in square brackets. Each structure is itself a *list* containing three parts: (1) an eigenvalue (denoted above by λ); (2) its multiplicity (denoted above by m); and (3) the corresponding *set* (enclosed in curly brackets) of one or more eigenvectors (each of which is a *list*, enclosed in square brackets), such as $\{[a_1, a_2, \cdots], [b_1, b_2, \cdots], \cdots\}$. Even for an eigenvalue with multiplicity one, Maple gives the eigenvector as a *set* (with one element, i.e., one list of eigenvector components).

Example 1

The command to enter the matrix $\mathbf{A} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 8 & 0 \end{pmatrix}$ is

```
A := matrix(3, 3, [5,0,0,0,0,2,0,8,0]);
```

The command `val := eigenvals(A);` then gives val := 5, 4, –4. If you want to call the second eigenvalue in the set v2, then use the command `v2 := val[2];` the result is 4.

The command `X := eigenvects(A);` gives the result

X := [4, 1, {[0, 1, 2]}], [–4, 1,{[0, 1, –2]}], [5, 1, {[1, 0, 0]}]

This confirms that the eigenvalues are 4, –4, 5 (a different order from the above result). It also shows that the multiplicity of each is one, and it gives the components of the corresponding eigenvectors. Again, if you run this code on a different computer, the eigenvector structures may come out in a different order.

As an example, suppose we want to extract the second eigenvector (corresponding to $\lambda = -4$). To do this we need to select the *second* structure in the set, then its *third* member (the set of eigenvectors), then the *first* member of the set, to get the first eigenvector corresponding to this eigenvalue (even though there's only one because this eigenvalue has multiplicity one). If we call this eigenvector x2, then the appropriate command is `x2 := X[2][3][1];` The [2] selects the structure, the [3] the third element of that structure (i.e., the set of eigenvectors), and the [1] the particular eigenvector in the set that we want. The resulting Maple output is x2 := [0, 1, −2].

Ordinarily we want to treat the eigenvectors as single-column matrices. However, Maple treats this expression for x2 as a *list*, not a matrix. So we have to take the extra step of constructing a three-row, one-column matrix, which we may call vec2, using the command `vec2 := matrix(3,1,x2);` We get $\text{vec2} := \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$. (Note that x2 doesn't have to be enclosed in square brackets because it is already a list.) Maple treats vec2 as a matrix, and we can use all the usual Maple commands for matrices. In particular, to verify that this really is an eigenvector, we compute `evalm(A &* vec2);` The result is $\begin{pmatrix} 0 \\ -4 \\ 8 \end{pmatrix}$. We recognize this as (−4) times the original vector, confirming that it is indeed an eigenvector of **A** with eigenvalue −4. Similarly, we invite you to verify that

$$\text{vec1} := \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \quad \text{and} \quad \text{vec3} := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

and that these are indeed eigenvectors of **A**, with eigenvalues 4 and 5, respectively.

Example 2

The matrix $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ is entered with `A := matrix(3, 3, [1,0,0,0,0,1,0,1,0]);` the eigenvalues are −1 (with multiplicity 1) and 1 (with multiplicity 2). The Maple output for `eigenvectors(A);` is

$$\text{eigenvectors(A)} := [-1, 1, \{[0, -1, 1]\}], \quad [1, 2, \{[0, 1, 1], [1, 0, 0]\}]$$

This says that the eigenvalue −1 has multiplicity 1 and that the corresponding eigenvector is [0, −1, 1]. The eigenvalue 1 has multiplicity 2, and the corresponding orthogonal eigenvectors are [0, 1, 1] and [1, 0, 0]. Note that in the case of multiple eigenvalues, the set of eigenvectors for each multiple eigenvalue are always *orthogonal*.

Normalization

It is often useful to *normalize* eigenvectors. A vector $\mathbf{x}$ is said to be *normalized* if the scalar product (or inner product) of the vector with itself is unity, that is, if $\mathbf{x}^T\mathbf{x}$, = 1. If the elements of the eigenvectors are complex, then we generalize the normalization requirement to be $\mathbf{x}^+\mathbf{x} = 1$. (For *any* vector $\mathbf{x}$, the product $\mathbf{x}^+\mathbf{x}$ is always real and non-negative, even when the components of $\mathbf{x}$ are complex. Can you prove this?)

Any vector can be normalized by computing $\mathbf{x}^T\mathbf{x}$ (or $\mathbf{x}^+\mathbf{x}$) and then dividing each component of $\mathbf{x}$ by the square root of the result. This can be done with Maple, but the method is a little circuitous. If you want to call the normalized vector xnorm, the Maple command is `xnorm := normalize(x);` But this command works only on *arrays* or *lists*, not on *matrices*, because in general Maple doesn't know that a particular matrix has only one column. So we have to normalize the array *first*, and then convert it to a matrix that Maple recognizes.

Referring back to Example 1, Maple doesn't understand the command `x2norm :=normalize(vec2);` But we can use `x2norm := normalize(x2);` the resulting Maple output is

$$\text{x2norm} := \left[0,\ \frac{1}{5}\sqrt{5},\ -\frac{2}{5}\sqrt{5}\right].$$

Calling the normalized eigenvector vec2norm, we convert x2norm to a matrix with

```
vec2norm := matrix(3, 1, x2norm);
```

(Note again that we don't need square brackets around `x2norm` because it is already a list.) The Maple result is $\text{vec2norm} := \begin{pmatrix} 0 \\ \frac{\sqrt{5}}{5} \\ \frac{-2}{5}\sqrt{5} \end{pmatrix}$. vec1norm and vec3norm are obtained similarly. We invite you to verify that all three eigenvectors *are* normalized.

Here's one more Maple quirk. We usually think of the scalar product (or inner product) of two vectors as a *scalar* quantity. In matrix language, it is the matrix product of $\mathbf{a}^T$ and $\mathbf{b}$ (or of $\mathbf{a}^+$ and $\mathbf{b}$ if the components are complex). But, strictly speaking, this product is a 1×1 matrix, which to Maple isn't the same thing as a single number. To extract the value from inside the matrix brackets, we can take the *determinant* of the product. Or we can designate the [1, 1] element of the matrix.

If we have defined vectors (single-column matrices) $\mathbf{a}$ and $\mathbf{b}$ and we want the numerical value of the scalar product, $c = \mathbf{a}\cdot\mathbf{b} = |\mathbf{a}^T\mathbf{b}|$, we can use

```
c :=det(transpose(a) &* b);   or   C := transpose(a) &* b;  and  c := C[1, 1];
```