Unit 5

Eigenvalues and eigenvectors

Introduction

Consider the following simplified migration problem.

The towns Exton and Wyeville have a regular interchange of population: each year, one-tenth of Exton's population migrates to Wyeville, while one-fifth of Wyeville's population migrates to Exton (see Figure 1). Other changes in population, such as births, deaths and other migrations, cancel each other thus can be ignored. If x_n and y_n denote, respectively, the populations of Exton and Wyeville at the beginning of year n, then the corresponding populations at the beginning of year $n+1$ are given by

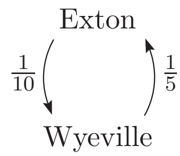


Figure 1 Population flow

$$\begin{aligned} x_{n+1} &= 0.9x_n + 0.2y_n, \\ y_{n+1} &= 0.1x_n + 0.8y_n, \end{aligned}$$

or, in matrix form,

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix}.$$

A matrix such as $\begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}$ is called a *transition matrix*.

The elements of such a matrix are all non-negative, and the elements in each column sum to 1.

This is an example of an *iterative process*, in which the values associated with the $(n+1)$th iterate can be determined from the values associated with the nth iterate.

Suppose that initially the population of Exton is 10 000 and that of Wyeville is 8000, that is, $x_0 = 10\,000$ and $y_0 = 8000$. Then after one year the populations are given by

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix} \begin{pmatrix} 10\,000 \\ 8\,000 \end{pmatrix} = \begin{pmatrix} 10\,600 \\ 7\,400 \end{pmatrix},$$

and after two years they are given by

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix} \begin{pmatrix} 10\,600 \\ 7\,400 \end{pmatrix} = \begin{pmatrix} 11\,020 \\ 6\,980 \end{pmatrix};$$

we can continue this process as far as we wish, provided that the underlying assumptions do not change.

What happens in the long term? Do the populations eventually stabilise? Using a computer algebra package, we can verify that the populations after 30 years are $x_{30} = 12\,000$ and $y_{30} = 6000$. It follows that after 31 years the populations are given by

$$\begin{aligned} \begin{pmatrix} x_{31} \\ y_{31} \end{pmatrix} &= \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix} \begin{pmatrix} 12\,000 \\ 6\,000 \end{pmatrix} \\ &= \begin{pmatrix} 0.9 \times 12\,000 + 0.2 \times 6000 \\ 0.1 \times 12\,000 + 0.8 \times 6000 \end{pmatrix} = \begin{pmatrix} 12\,000 \\ 6\,000 \end{pmatrix}. \end{aligned}$$

So if $x_{30} = 12\,000$ and $y_{30} = 6000$, then $x_{31} = x_{30}$ and $y_{31} = y_{30}$, and the sizes of the populations of the two towns do not change. Moreover, the sizes of the populations will not change in any subsequent year.

There are situations, such as in the above migration problem, where a particular non-zero vector does not change under a linear transformation. However, this is more the exception than the rule. It is more useful to investigate vectors that are transformed to *scalar multiples* of themselves – geometrically, this means that each such vector is transformed into another vector in the same or opposite direction. Such vectors are called *eigenvectors*, and the corresponding scalar multipliers are called *eigenvalues*.

Section 1 considers the eigenvectors and eigenvalues associated with various linear transformations of the plane. We also outline situations where eigenvectors and eigenvalues are useful.

In many problems, such as in the above migration problem, the appropriate linear transformation is given in matrix form. In Sections 2 and 3 we investigate the various types of eigenvalue that can arise for 2×2 and 3×3 matrices.

Here 'steady-state' refers to the fact that the populations no longer change.

In the migration problem, we iterated to obtain the 'steady-state' population. Section 4 uses a similar method in order to calculate the eigenvalues and eigenvectors of a matrix. Although our discussion is mainly about 2×2 and 3×3 matrices, many of the ideas can be extended to the much larger matrices that often arise from practical applications.

1 Introducing eigenvectors

We first investigate the eigenvectors and eigenvalues that arise from various linear transformations of the plane – in particular, scaling, reflection and rotation. Do not worry about how we construct the transformation matrices – that is not important. It is the geometric properties of these matrices, and the relevance of these properties to eigenvectors and eigenvalues, that are important here. Later in this section we outline another use of eigenvectors and eigenvalues.

1.1 Eigenvectors in the plane

Consider the linear transformation of the plane specified by the matrix

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}.$$

Using matrix multiplication, we can find the image under this linear transformation of any given vector in the plane. For example, to find the image of the vector $\mathbf{v} = (2 \quad 1)^T$, we calculate

$$\mathbf{A}\mathbf{v} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 8 \\ 6 \end{pmatrix},$$

so the required image is $(8 \quad 6)^T$, as shown in Figure 2.

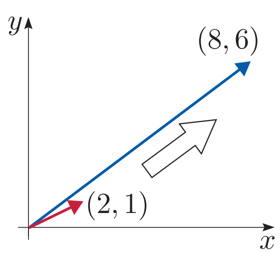


Figure 2 A linear transformation acting on the vector $(2 \quad 1)^T$

In general, to find the image under this linear transformation of the vector $\mathbf{v} = (x \quad y)^T$, we calculate

$$\mathbf{Av} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x + 2y \\ x + 4y \end{pmatrix},$$

so the required image is $(3x + 2y \quad x + 4y)^T$.

This is a linear transformation because each component of the vector $(3x + 2y \quad x + 4y)^T$ is a linear function of x and y. Recall the definition of a linear transformation from Subsection 2.2 of Unit 4.

Exercise 1

Find the image of each of the vectors

$$\mathbf{w} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad \mathbf{z} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

under the transformation matrix $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$.

In Exercise 1 you saw that the image of $\mathbf{x} = (1 \quad 1)^T$ is $(5 \quad 5)^T$. This image is a vector in the same direction as $\mathbf{x}$, but with five times the magnitude, as shown in Figure 3. In symbols, we write this transformation as $\mathbf{Ax} = 5\mathbf{x}$.

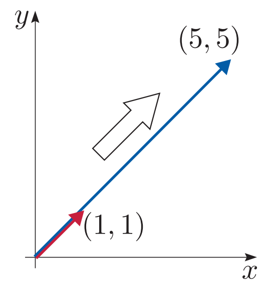


Figure 3 The vector $(1, 1)$ is transformed to the vector $(5, 5)$

In fact, as you can easily check, *any* vector in the same direction as $\mathbf{x} = (1 \quad 1)^T$ (or in the direction opposite to $\mathbf{x}$) is transformed into another vector in the same (or opposite) direction, but with five times the magnitude, as shown in Figure 4. Non-zero vectors that are transformed to scalar multiples of themselves are called *eigenvectors*; the corresponding scalar multipliers are called *eigenvalues*. For the 2×2 matrix $\mathbf{A}$ in Exercise 1, $(1 \quad 1)^T$ is an eigenvector, with corresponding eigenvalue 5, since

$$\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

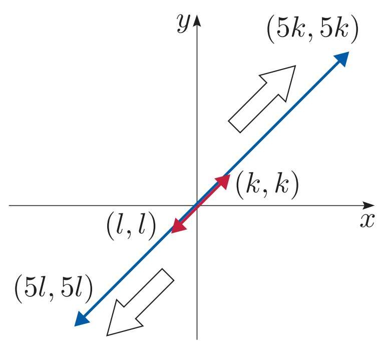


Figure 4 Any scalar multiple of $(1, 1)$ is transformed to the same scalar multiple of $(5, 5)$

Similarly, since (from Exercise 1)

$$\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} -2 \\ 1 \end{pmatrix},$$

we can see that $(-2 \quad 1)^T$ is also an eigenvector, with corresponding eigenvalue 2.

Eigenvalues and eigenvectors

Let $\mathbf{A}$ be any square matrix. A non-zero vector $\mathbf{v}$ is an **eigenvector** of $\mathbf{A}$ if $\mathbf{Av} = \lambda\mathbf{v}$ for some number λ. The number λ is the corresponding **eigenvalue**.

Because $\mathbf{A}(k\mathbf{v}) = k(\mathbf{Av}) = k(\lambda\mathbf{v}) = \lambda(k\mathbf{v})$ for any scalar k, any non-zero multiple of an eigenvector is also an eigenvector with the same eigenvalue.

For example, if $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$, then for any non-zero number k, each vector $(k \quad k)^T$ is an eigenvector corresponding to the eigenvalue 5, because

$$\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} k \\ k \end{pmatrix} = 5 \begin{pmatrix} k \\ k \end{pmatrix};$$

also, each vector $(-2k \quad k)^T$ is an eigenvector corresponding to the eigenvalue 2, because

$$\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} -2k \\ k \end{pmatrix} = 2 \begin{pmatrix} -2k \\ k \end{pmatrix}.$$

We defined linear dependence for the rows of a matrix in Unit 4. The concept may be extended to any collection of vectors as follows.

Linear dependence and linear independence

The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are **linearly dependent** if a non-trivial linear combination of these vectors is equal to the zero vector, that is,

$$q_1\mathbf{v}_1 + q_2\mathbf{v}_2 + \cdots + q_n\mathbf{v}_n = \mathbf{0}, \qquad (1)$$

where the numbers $q_1, q_2, \ldots, q_n$ are not all zero.

Vectors that are not linearly dependent are **linearly independent**.

Example 1

Consider the two-dimensional vectors

Recall that the two-dimensional Cartesian unit vectors $\mathbf{i}$ and $\mathbf{j}$ may be written as column vectors as $\mathbf{i} = (1 \quad 0)^T$ and $\mathbf{j} = (0 \quad 1)^T$.

$$\mathbf{v}_1 = \mathbf{i}, \quad \mathbf{v}_2 = \mathbf{i} + \mathbf{j}, \quad \mathbf{v}_3 = 5\mathbf{i} - 3\mathbf{j}.$$

(a) Show that $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent.

(b) Show that $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly dependent.

Solution

(a) For linear dependence, we need to find q_1 and q_2, not both zero, such that $q_1\mathbf{v}_1 + q_2\mathbf{v}_2 = \mathbf{0}$. But

$$q_1\mathbf{v}_1 + q_2\mathbf{v}_2 = (q_1 + q_2)\mathbf{i} + q_2\mathbf{j},$$

and it is not possible to find q_1 and q_2, not both zero, such that the right-hand side of this equation is equal to $\mathbf{0}$. Therefore $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent.

(b) We need to find q_1, q_2 and q_3, not all zero, such that $q_1\mathbf{v}_1 + q_2\mathbf{v}_2 + q_3\mathbf{v}_3 = \mathbf{0}$. But

$$q_1\mathbf{v}_1 + q_2\mathbf{v}_2 + q_3\mathbf{v}_3 = (q_1 + q_2 + 5q_3)\mathbf{i} + (q_2 - 3q_3)\mathbf{j}.$$

Start by choosing $q_3 = 1$. Then for the coefficient of $\mathbf{j}$ to be zero we require $q_2 = 3$, and for the coefficient of $\mathbf{i}$ to be zero we require $q_1 + 3 + 5 = 0$, so $q_1 = -8$.

Hence

$$-8\mathbf{v}_1 + 3\mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0},$$

so $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly dependent.

Example 2

Consider the three-dimensional vectors

$$\mathbf{v}_1 = \mathbf{i} + \mathbf{j} + \mathbf{k}, \quad \mathbf{v}_2 = 2\mathbf{i} + \mathbf{j} + \mathbf{k}, \quad \mathbf{v}_3 = 3\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}.$$

Recall that the three-dimensional Cartesian unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ may be written as column vectors as $\mathbf{i} = (1 \quad 0 \quad 0)^T$, $\mathbf{j} = (0 \quad 1 \quad 0)^T$ and $\mathbf{k} = (0 \quad 0 \quad 1)^T$.

Show that $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent.

Solution

If $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ were linearly dependent, it would be possible to find q_1, q_2 and q_3, not all zero, such that $q_1\mathbf{v}_1 + q_2\mathbf{v}_2 + q_3\mathbf{v}_3 = \mathbf{0}$. But

$$\begin{aligned} q_1\mathbf{v}_1 + q_2\mathbf{v}_2 + q_3\mathbf{v}_3 = {} & (q_1 + 2q_2 + 3q_3)\mathbf{i} + (q_1 + q_2 + 2q_3)\mathbf{j} \\ & + (q_1 + q_2 + 3q_3)\mathbf{k}. \end{aligned}$$

It is not possible to set all the coefficients of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ to zero without having $q_1 = q_2 = q_3 = 0$. To see this, note that for the coefficient of $\mathbf{j}$ to be zero, we require $q_1 + q_2 = -2q_3$. Then for the coefficient of $\mathbf{k}$ to be zero, we require $q_1 + q_2 + 3q_3 = -2q_3 + 3q_3 = q_3 = 0$, which means that $q_1 + q_2 = -2q_3 = 0$, so $q_1 = -q_2$. Finally, for the coefficient of $\mathbf{i}$ to be zero, we require $q_1 + 2q_2 + 3q_3 = -q_2 + 2q_2 + 3 \times 0 = q_2 = 0$, hence $q_1 = -q_2 = 0$; and we already have $q_3 = 0$. Thus $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent.

The nature of the eigenvectors of a matrix is closely bound to the notion of linear dependence.

Consider two (non-zero) vectors $\mathbf{v}_1$ and $\mathbf{v}_2$. If these vectors are linearly dependent, then we can write

$$q_1\mathbf{v}_1 + q_2\mathbf{v}_2 = \mathbf{0}$$

with q_1 and q_2 not both zero. If $q_1 \neq 0$, then

$$\mathbf{v}_1 = -\frac{q_2}{q_1}\mathbf{v}_2.$$

Similarly, if $q_2 \neq 0$, then

$$\mathbf{v}_2 = -\frac{q_1}{q_2}\mathbf{v}_1.$$

Thus one of the vectors is a multiple of the other.

On the other hand, if one of the vectors is a multiple of the other, say $\mathbf{v}_2 = k\mathbf{v}_1$ ($k \neq 0$), then

$$q_1\mathbf{v}_1 + q_2\mathbf{v}_2 = (q_1 + kq_2)\mathbf{v}_1,$$

and choosing $q_1 = k$, $q_2 = -1$ shows that the vectors are linearly dependent.

So we have shown that two (non-zero) vectors are linearly dependent if and only if one is a multiple of the other.

Thus the vectors $(1 \quad 1)^T$ and $(2 \quad 2)^T$ are linearly dependent, as are all the eigenvectors of the form $(k \quad k)^T$ corresponding to the eigenvalue 5 of the matrix $\mathbf{A}$ considered in Exercise 1. Similarly, all the eigenvectors of the form $(-2k \quad k)^T$ corresponding to the eigenvalue 2 of $\mathbf{A}$ are linearly dependent. However, the eigenvectors $(1 \quad 1)^T$ and $(-2 \quad 1)^T$ are linearly independent, since one is clearly not a multiple of the other.

The following exercise, concerning the eigenvectors of $\mathbf{A}$, illustrates a property of linearly independent vectors that will prove useful later.

Exercise 2

The vectors $\mathbf{v}_1 = (1 \quad 1)^T$ and $\mathbf{v}_2 = (-2 \quad 1)^T$ are linearly independent. Show that for any vector $\mathbf{v} = (x \quad y)^T$, there are numbers α and β such that $\mathbf{v} = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$. (In other words, show that any two-dimensional vector $\mathbf{v}$ can be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.)

The result in Exercise 2 generalises to *any* two linearly independent vectors. That is, if $\mathbf{v}_1$ and $\mathbf{v}_2$ are any pair of linearly independent two-dimensional vectors and $\mathbf{v}$ is any two-dimensional vector, then we can find numbers α and β such that $\mathbf{v} = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$.

The result generalises further. If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are any n linearly independent n-dimensional vectors and $\mathbf{v}$ is any n-dimensional vector, then we can find numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n.$$

We do not prove this result here.

Returning to eigenvectors, we require an eigenvector to be non-zero. This is because $\mathbf{A0} = \mathbf{0}$ for *every* square matrix $\mathbf{A}$, so $\mathbf{0}$ would otherwise be an eigenvector for *all* square matrices. Thus there is no point in including $\mathbf{0}$ as an eigenvector. However, it is possible for an eigenvalue to be 0. For example, if we choose $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$, then

$$\mathbf{Av} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0\begin{pmatrix} 1 \\ -2 \end{pmatrix},$$

so $(1 \quad -2)^T$ is an eigenvector of $\mathbf{A}$ with eigenvalue 0.

Exercise 3

In each of the following cases, verify that $\mathbf{v}$ is an eigenvector of $\mathbf{A}$, and write down the corresponding eigenvalue.

(a) $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

(c) $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 0 \\ 6 \end{pmatrix}$ (d) $\mathbf{A} = \begin{pmatrix} 12 & 3 \\ -8 & -2 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 1 \\ -4 \end{pmatrix}$

Exercise 4

Write down an eigenvector and the corresponding eigenvalue for the matrix associated with the migration problem in the Introduction, that is,

$$\begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}.$$

Exercise 5

Show that $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ are eigenvectors of the matrix $\begin{pmatrix} 1 & 4 \\ 9 & 1 \end{pmatrix}$, and find the corresponding eigenvalues.

Very occasionally it is possible to determine information about the eigenvectors and eigenvalues of a given matrix from its geometric properties. This is so for each of the cases in Figure 5 (where the image of the unit square is used in each case to indicate the behaviour of the transformation).

The unit square has vertices $(0, 0)$, $(1, 0)$, $(1, 1)$ and $(0, 1)$.

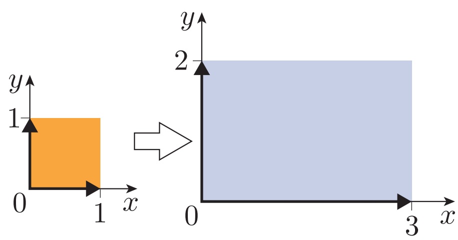
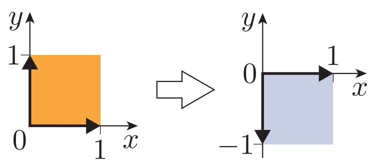
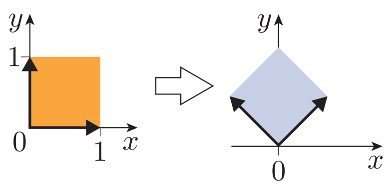
Matrix	Comment	Transformation of the unit square	Eigenvectors	Eigenvalues
$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$	A scaling by 3 in the x-direction and by 2 in the y-direction (i.e. a $(3, 2)$ scaling)		$(1 \quad 0)^T$ $(0 \quad 1)^T$	3 2
$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	A reflection in the x-axis		$(1 \quad 0)^T$ $(0 \quad 1)^T$	1 -1
$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$	A rotation through $\frac{\pi}{4}$ anticlockwise about the origin		No real eigenvalues	No real eigenvectors

Figure 5 Three matrices representing transformations of the plane, together with their real eigenvectors and corresponding eigenvalues

In the case of the matrix $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$, it is clear from the geometric properties of the linear transformation that vectors along the coordinate axes are eigenvectors, as these are transformed to vectors in the same directions.

In the case of the matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, we see that reflection in the x-axis leaves the vector $(1 \quad 0)^T$ unchanged, and reverses the direction of $(0 \quad 1)^T$, so these must be eigenvectors.

In the third case (rotation through $\pi/4$ anticlockwise about the origin), we do not expect to find any real eigenvectors because the direction of every vector is changed by the linear transformation. Surprisingly, even this matrix has eigenvectors of a kind, namely complex ones, but we would have to adopt the algebraic approach of Section 2 in order to find them.

Exercise 6

The matrix $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ represents reflection in a line through the origin at an angle $\pi/4$ to the x-axis. What are the eigenvectors of $\mathbf{A}$ and their corresponding eigenvalues?

(*Hint*: Find two lines through the origin that are transformed to themselves, then consider what happens to a point on each line.)

Exercise 7

The matrix $\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ represents reflection in the y-axis. Write down two eigenvectors and the corresponding eigenvalues.

(*Hint*: Find two lines through the origin that are transformed to themselves, then consider what happens to a point on each line.)

1.2 Simultaneous differential equations

We conclude this section by outlining another application of eigenvectors and eigenvalues that you will meet again in Unit 6.

Many mathematical models give rise to systems of simultaneous differential equations relating two or more functions and their derivatives. A simple case is shown in Example 3; we will discuss the relevance of eigenvectors and eigenvalues to these equations after we have solved them.

Example 3

Determine the general solution of the pair of differential equations

$\dot{x}$ is dx/dt, $\dot{y}$ is dy/dt.

$$\dot{x} = 3x + 2y, \tag{2}$$
$$\dot{y} = x + 4y, \tag{3}$$

where $x(t)$ and $y(t)$ are functions of the independent variable t.

Solution

In order to solve such a system of equations, we need to find specific formulas for x and y in terms of t.

From equation (3) we have

$$x = \dot{y} - 4y, \tag{4}$$

and differentiating this equation with respect to t gives

$$\dot{x} = \ddot{y} - 4\dot{y}. \tag{5}$$

Substituting for x and $\dot{x}$ in equation (2) using equations (4) and (5), we obtain $\ddot{y} - 4\dot{y} = 3(\dot{y} - 4y) + 2y$, which simplifies to

$$\ddot{y} - 7\dot{y} + 10y = 0.$$

This is a type of differential equation that you should recognise from Unit 1.

The general solution of this equation is obtained by solving the auxiliary equation $\lambda^2 - 7\lambda + 10 = 0$. This has solutions $\lambda = 5$ and $\lambda = 2$, so the required general solution for y is

$$y = Ae^{5t} + Be^{2t},$$

where A and B are arbitrary constants.

It follows from equation (4) that

$$\begin{aligned} x &= \dot{y} - 4y \\ &= (5Ae^{5t} + 2Be^{2t}) - 4(Ae^{5t} + Be^{2t}) \\ &= Ae^{5t} - 2Be^{2t}. \end{aligned}$$

Thus the general solution of the system of equations is

$$\begin{cases} x = Ae^{5t} - 2Be^{2t}, \\ y = Ae^{5t} + Be^{2t}. \end{cases} \tag{6}$$

What is the connection between eigenvalues and eigenvectors, and the equations of Example 3? The differential equations (2) and (3) can be written in matrix form as

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{7}$$

The 2×2 matrix in equation (7) has eigenvectors $(1 \quad 1)^T$ and $(-2 \quad 1)^T$ with corresponding eigenvalues 5 and 2, respectively, as we saw in Subsection 1.1.

These eigenvectors and eigenvalues appear in the matrix form of the general solution given in equations (6):

$$\begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{5t} + B \begin{pmatrix} -2 \\ 1 \end{pmatrix} e^{2t}.$$

The first term on the right-hand side involves $(1 \quad 1)^T$, an eigenvector of the matrix of coefficients $\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$, and a term e^{5t}, where 5 is the corresponding eigenvalue.

The second term on the right-hand side is of the same form, but it involves the other eigenvector and corresponding eigenvalue. That the general solution of a system of differential equations can be written explicitly in terms of eigenvectors and eigenvalues is no coincidence, as we will explain in Unit 6.

2 Eigenvalues and eigenvectors of 2×2 matrices

In Section 1 we considered the linear transformation of the plane specified by the matrix $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$. We saw that $(1 \quad 1)^T$ and $(-2 \quad 1)^T$ are eigenvectors of $\mathbf{A}$, with corresponding eigenvalues 5 and 2, respectively.

We did not show in Section 1 that these are the *only* possible eigenvalues of $\mathbf{A}$, nor did we show you how to find the eigenvectors and corresponding eigenvalues for an arbitrary matrix. In this section we use algebraic techniques to show you how to calculate *all* the eigenvalues and eigenvectors of any given 2×2 matrix. We also investigate the three situations that arise when the eigenvalues are:

- distinct real numbers
- one real number repeated
- complex numbers.

For example:

5 and 2,

2 and 2,

i and $-i$, where $i^2 = -1$.

Before we begin the discussion, we need to remind you of some results from Unit 4.

Often, a pair of equations with zero right-hand sides, such as

$$\begin{cases} 2x + 3y = 0, \\ x - y = 0, \end{cases} \tag{8}$$

has just one solution $x = 0$, $y = 0$. But this is not the case for every pair of equations of this type. For example, the equations

$$\begin{cases} x - 2y = 0, \\ -2x + 4y = 0, \end{cases} \tag{9}$$

have a solution $x = 2$, $y = 1$, and another solution $x = 6$, $y = 3$. In fact, $x = 2k$, $y = k$ is a solution for every value of k, that is, there is an infinite number of solutions. This is because the equations $x - 2y = 0$ and $-2x + 4y = 0$ are essentially the same equation, since either one of the equations can be obtained from the other by multiplying by a suitable factor.

Example 4

Find a solution (other than $x = y = 0$) of the equations

$$\begin{aligned} 3x - 5y &= 0, \\ -6x + 10y &= 0. \end{aligned}$$

The second equation can be obtained from the first equation by multiplying it by -2.

Solution

There is an infinite number of solutions. For example, choosing $x = 5$, we find (from either equation) that $y = 3$, which gives us a solution of the pair of equations. (Any other non-zero choice of x would give another solution.)

The general solution is $x = 5k$, $y = 3k$, for any constant k.

You saw in Unit 4 that a system of linear equations $\mathbf{Ax} = \mathbf{b}$ has a unique solution if and only if the matrix of coefficients $\mathbf{A}$ is invertible. So if the system has an infinite number of solutions, then the matrix $\mathbf{A}$ must be non-invertible.

Thus the coefficient matrix $\begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}$ of equations (9) must be non-invertible, as must be the coefficient matrix $\begin{pmatrix} 3 & -5 \\ -6 & 10 \end{pmatrix}$ of Example 4, whereas the coefficient matrix $\begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix}$ of equations (8) must be invertible. Recall from Unit 4 that a matrix $\mathbf{A}$ is non-invertible if and only if $\det \mathbf{A} = 0$. You may like to check that the determinants of the first two matrices are zero, but the determinant of the last matrix is non-zero.

2.1 Basic method

We now look at a method for finding the eigenvalues and corresponding eigenvectors of an arbitrary 2×2 matrix. Consider the following example.

Example 5

Find the eigenvalues and corresponding eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}.$$

This matrix was discussed in Subsection 1.1.

Solution

We wish to find those non-zero vectors $\mathbf{v}$ that satisfy the equation

$$\mathbf{Av} = \lambda \mathbf{v}, \tag{10}$$

for some number λ. Writing $\mathbf{v} = (x \quad y)^T$, we have

$$\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix},$$

that is, x and y must satisfy the simultaneous linear equations

$$\begin{cases} 3x + 2y = \lambda x, \\ x + 4y = \lambda y. \end{cases}$$

We are interested in only non-zero vectors $\mathbf{v}$, so x and y are not both zero. These equations will have such a solution for only certain values of λ – the eigenvalues of the matrix – and our first task is to find these values of λ. The above equations are a pair of linear equations in the unknowns x and y, where λ is a constant that has yet to be determined.

Transferring the terms λx and λy to the left-hand side, we obtain

$$\begin{cases} (3-\lambda)x + \qquad 2y = 0, \\ \qquad x + (4-\lambda)y = 0, \end{cases} \tag{11}$$

To see this, notice that equation (10) can be rewritten as $\mathbf{A}\mathbf{v} = \lambda\mathbf{I}\mathbf{v}$ (where $\mathbf{I}$ is the 2×2 identity matrix), which in turn can be rewritten as $\mathbf{A}\mathbf{v} - \lambda\mathbf{I}\mathbf{v} = \mathbf{0}$ and hence as $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$.

which can be written as

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}. \tag{12}$$

It is convenient to refer to equations (11) as the *eigenvector equations.*

Do not lose sight of the fact that we require $\mathbf{v} \neq \mathbf{0}$. In Unit 4 we saw that if such a non-zero solution $\mathbf{v}$ exists, then the matrix $\mathbf{A} - \lambda\mathbf{I}$ in equation (12) must be non-invertible, so $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, that is,

$$\begin{vmatrix} 3-\lambda & 2 \\ 1 & 4-\lambda \end{vmatrix} = 0. \tag{13}$$

Recall that $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$

Our conclusion at this stage is that equation (10) can hold for only certain values of λ, and these values must satisfy equation (13). Expanding the determinant in equation (13) gives

$$(3-\lambda)(4-\lambda) - 2 = 0,$$

which simplifies to

$$\lambda^2 - 7\lambda + 10 = 0.$$

Since $\lambda^2 - 7\lambda + 10 = (\lambda - 5)(\lambda - 2)$, we deduce that $\lambda = 5$ or $\lambda = 2$.

Thus it is only when $\lambda = 5$ or $\lambda = 2$ that equation (10) has a non-zero solution $\mathbf{v}$, that is, 5 and 2 are the *only* eigenvalues of $\mathbf{A}$.

In order to find the corresponding eigenvectors, we substitute each of these eigenvalues into equations (11) and solve the resulting eigenvector equations, as follows.

The matrix $\mathbf{A} - \lambda\mathbf{I}$, i.e.
$$\begin{pmatrix} 3-\lambda & 2 \\ 1 & 4-\lambda \end{pmatrix},$$
is non-invertible, as stated above, so the equivalent eigenvector equations are linearly dependent – hence represent a *single* equation. This is equivalent to the property of non-invertible matrices having linearly dependent rows, as explained in Unit 4.

- For $\lambda = 5$, the eigenvector equations (11) become $-2x + 2y = 0$ and $x - y = 0$, and both are equivalent to the single equation $y = x$. It follows that an eigenvector corresponding to $\lambda = 5$ is $(1 \quad 1)^T$.
- For $\lambda = 2$, the eigenvector equations (11) become $x + 2y = 0$ and $x + 2y = 0$, and both are equivalent to the single equation $-2y = x$. It follows that an eigenvector corresponding to $\lambda = 2$ is $(-2 \quad 1)^T$.

Thus $(1 \quad 1)^T$ is an eigenvector of $\mathbf{A}$ with corresponding eigenvalue 5, and $(-2 \quad 1)^T$ is an eigenvector of $\mathbf{A}$ with corresponding eigenvalue 2.

In this case, the matrix $\mathbf{A}$ has two distinct eigenvalues, and these correspond to two linearly independent eigenvectors.

In Example 5 we found an eigenvector $(1 \quad 1)^T$ corresponding to the eigenvalue 5, and an eigenvector $(-2 \quad 1)^T$ corresponding to the eigenvalue 2, but these vectors are not unique. We could have chosen $(2 \quad 2)^T$ as the eigenvector corresponding to the eigenvalue 5, and $(6 \quad -3)^T$ as the eigenvector corresponding to the eigenvalue 2. However, *all* the eigenvectors corresponding to 5 are multiples of $(1 \quad 1)^T$, and *all* the eigenvectors corresponding to 2 are multiples of $(-2 \quad 1)^T$.

Thus, in a sense, $(1 \quad 1)^T$ and $(2 \quad 2)^T$ represent the 'same' eigenvector, as do $(-2 \quad 1)^T$ and $(6 \quad -3)^T$.

Exercise 8

Use the method of Example 5 to find the eigenvalues and corresponding eigenvectors of the matrix $\mathbf{A} = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}$.

As we will see shortly, there is a result involving the diagonal elements of a matrix that provides a useful check on calculations. To state this result succinctly, we will need the following definition.

Trace of a matrix

The sum of the elements on the leading diagonal of a square matrix $\mathbf{A}$ is known as the **trace** of $\mathbf{A}$ and is written as $\operatorname{tr} \mathbf{A}$.

In Example 5 we found the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$$

by solving the equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, that is,

$$\begin{vmatrix} 3-\lambda & 2 \\ 1 & 4-\lambda \end{vmatrix} = 0,$$

which gives a quadratic equation in λ.

More generally, we find the eigenvalues of the 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

by solving the equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, that is,

$$\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = 0,$$

for λ. Expanding this determinant gives

$$(a-\lambda)(d-\lambda) - bc = 0,$$

which is equivalent to the quadratic equation

$$\lambda^2 - (a+d)\lambda + (ad-bc) = 0. \tag{14}$$

Writing this equation in the form

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = 0$$

gives

$$\lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2 = 0.$$

So the roots λ_1 and λ_2 of equation (14) are the eigenvalues of $\mathbf{A}$ and satisfy

$$\lambda_1\lambda_2 = ad - bc = \det \mathbf{A} \quad \text{and} \quad \lambda_1 + \lambda_2 = a + d = \operatorname{tr} \mathbf{A}. \tag{15}$$

Equation (14) is called the **characteristic equation** of $\mathbf{A}$. Note that by combining the results in equations (15) with equation (14), the characteristic equation for the 2×2 matrix $\mathbf{A}$ can be expressed as

$$\lambda^2 - (\operatorname{tr} \mathbf{A})\,\lambda + \det \mathbf{A} = 0.$$

This can have distinct real roots, a repeated real root or complex roots, depending on the values of a, b, c and d. We investigate these three possibilities in the next subsection.

Using Example 5 and the above discussion as a model, we can now give the following procedures for determining the eigenvalues and eigenvectors of a given 2×2 matrix.

Procedure 1 Finding eigenvalues of a 2×2 matrix

Let $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. To find the *eigenvalues* of $\mathbf{A}$, write down the *characteristic equation* $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. Expand this as

$$\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = \lambda^2 - (a+d)\lambda + (ad - bc) = 0. \tag{16}$$

Solve this quadratic equation for λ.

Using equations (15) to determine the characteristic equation leads to the same quadratic equation.

Procedure 2 Finding eigenvectors of a 2×2 matrix

Let $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. To find an *eigenvector* corresponding to the eigenvalue λ, write down the **eigenvector equations**

$$\begin{cases} (a-\lambda)x + by = 0, \\ cx + (d-\lambda)y = 0. \end{cases} \tag{17}$$

These equations reduce to a single equation of the form $py = qx$, with solution $x = p$, $y = q$, so $(p \quad q)^T$ is an eigenvector. Any non-zero multiple $(kp \quad kq)^T$ is also an eigenvector corresponding to the same eigenvalue.

Note that in the equation $py = qx$, the coefficient of y gives the first component of the eigenvector, and the coefficient of x gives the second component.

Example 6

Find the eigenvalues and corresponding eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix},$$

following Procedures 1 and 2.

Solution

The characteristic equation is

$$\begin{vmatrix} 2-\lambda & 3 \\ 2 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda) - 6 = 0,$$

which simplifies to $\lambda^2 - 3\lambda - 4 = 0$. Using the formula for solving a quadratic equation, we deduce that the eigenvalues are $\lambda = 4$ and $\lambda = -1$.

Alternatively, you may have noticed that
$$\lambda^2 - 3\lambda - 4 = (\lambda - 4)(\lambda + 1).$$

- For $\lambda = 4$, the eigenvector equations become $-2x + 3y = 0$ and $2x - 3y = 0$, which reduce to the single equation $3y = 2x$ (so $p = 3$ and $q = 2$ in Procedure 2). Thus an eigenvector corresponding to $\lambda = 4$ is $(3 \quad 2)^T$.
- For $\lambda = -1$, the eigenvector equations become $3x + 3y = 0$ and $2x + 2y = 0$, which reduce to the single equation $-y = x$ (so $p = -1$ and $q = 1$ in Procedure 2). Thus an eigenvector corresponding to $\lambda = -1$ is $(-1 \quad 1)^T$.

We check that $\operatorname{tr} \mathbf{A} = 2 + 1 = 3$ and $\lambda_1 + \lambda_2 = 4 - 1 = 3$, as required, and $\det \mathbf{A} = 2 \times 1 - 3 \times 2 = -4$ and $\lambda_1\lambda_2 = 4 \times (-1) = -4$, as required.

In Example 6 we have two distinct real eigenvalues, and these correspond to two linearly independent eigenvectors.

Exercise 9

Use Procedures 1 and 2 to find the eigenvalues and corresponding eigenvectors of each of the following matrices.

(a) $\mathbf{A} = \begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix} 8 & -5 \\ 10 & -7 \end{pmatrix}$ (c) $\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 9 & 2 \end{pmatrix}$

From equations (15), the sum of the eigenvalues is $a + d = \operatorname{tr} \mathbf{A}$ and the product is $ad - bc = \det \mathbf{A}$. It is useful to check that these properties hold whenever you have calculated the eigenvalues of a given matrix. If they do not hold, then you have made a mistake, which you should rectify before proceeding to calculate the eigenvectors.

Similar results are true for all square matrices, of any size.

Exercise 10

Verify the properties in equations (15) for the matrices in the following.

(a) Example 5 (b) Exercise 8 (c) Exercise 9

2.2 Three types of eigenvalue

The characteristic equation of a matrix is a quadratic equation, and this has distinct real, repeated real or distinct complex roots depending on whether the discriminant of the quadratic equation is positive, zero or negative. We illustrate each of these three cases below.

See equation (14).

See Unit 1, Section 2.

Distinct real eigenvalues

In the following exercises, the eigenvalues are real and distinct.

Exercise 11

Calculate the eigenvalues and corresponding eigenvectors of the matrix $\begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix}$ with $p \neq q$. Check that your answers agree with those obtained by geometric considerations in Section 1 (see Figure 5).

Exercise 12

This matrix is associated with the migration problem discussed in the Introduction.

Calculate the eigenvalues and corresponding eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}.$$

This is true in general, though we do not prove it here except for the case of real symmetric matrices, which is left until Subsection 3.3.

Notice that whenever we have two distinct real eigenvalues, we also have two linearly independent eigenvectors.

Repeated eigenvalue

In the following exercise, the eigenvalue is repeated.

Exercise 13

Calculate the eigenvalues and corresponding eigenvectors of the following matrices.

(a) $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$, where $a \neq 0$ (b) $\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$, where $a \neq 0$

In fact, it is always true that if a 2×2 matrix has a repeated eigenvalue, then either we can select any two linearly independent vectors as eigenvectors, as is the case for diagonal matrices (Exercise 13(a)), or there is only one linearly independent eigenvector, as exemplified by Exercise 13(b). Indeed, one can show that if any *non-diagonal* 2×2 matrix has a repeated eigenvalue, then it has only one linearly independent eigenvector, although we do not prove this here.

Complex eigenvalues

Although we discuss complex eigenvalues, in this module the *elements* of a matrix $\mathbf{A}$ will always be real, so that $\mathbf{A}$ is a *real* matrix.

For 2×2 matrices, complex eigenvalues arise when there are no fixed directions under the corresponding linear transformation as, for example, in an anticlockwise rotation through $\pi/4$ about the origin (see Figure 5 and the discussion about it). They can occur for matrices of all sizes, but for real matrices they always occur in complex conjugate pairs (i.e. if $\lambda_1 = a + ib$ is an eigenvalue, then $\lambda_2 = a - ib$ is also an eigenvalue). The following example illustrates one of the simplest cases.

Example 7

Calculate the eigenvalues and corresponding eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Solution

The characteristic equation is $\lambda^2 + 1 = 0$. This equation has no real solutions, so there are no real eigenvalues. However, it has complex solutions

$$\lambda = i \quad \text{and} \quad \lambda = -i, \quad \text{where } i^2 = -1.$$

Check: sum $= \operatorname{tr} \mathbf{A} = 0$; product $= \det \mathbf{A} = 1$.

Thus the matrix $\mathbf{A}$ has the complex eigenvalues i and $-i$.

The eigenvector equations are $-\lambda x - y = 0$ and $x - \lambda y = 0$ (using equations (17)).

- For $\lambda = i$, the eigenvector equations become $-ix - y = 0$ and $x - iy = 0$, which reduce to the single equation $y = -ix$ (so $p = 1$ and $q = -i$ in Procedure 2). Thus an eigenvector corresponding to the eigenvalue $\lambda = i$ is $(1 \quad -i)^T$.

  The second equation is i times the first.

- For $\lambda = -i$, the eigenvector equations become $ix - y = 0$ and $x + iy = 0$, which reduce to the single equation $y = ix$ (so $p = 1$ and $q = i$ in Procedure 2). Thus an eigenvector corresponding to the eigenvalue $\lambda = -i$ is $(1 \quad i)^T$.

  The second equation is $-i$ times the first.

In Example 7 the eigenvectors are complex, but nevertheless are linearly independent. It remains true that any two eigenvectors corresponding to distinct eigenvalues, whether real or complex, are linearly independent.

Exercise 14

Calculate the eigenvalues and corresponding eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix}.$$

2.3 Some results on eigenvalues

In this subsection we list some general results that will be needed later. Although we introduce them in the context of 2×2 matrices, they hold for square matrices of any size.

We consider the eigenvalues of various types of matrix: triangular, symmetric and non-invertible.

Triangular matrices

See Unit 4.

A matrix is *triangular* if all the elements above (or below) the leading diagonal are 0. Thus a 2×2 triangular matrix has one of the forms

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}, \qquad \begin{pmatrix} a & 0 \\ c & d \end{pmatrix}, \qquad \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}.$$

(upper triangular) (lower triangular) (diagonal)

The above triangular matrices all have characteristic equation

$$(a-\lambda)(d-\lambda) = 0,$$

and eigenvalues $\lambda = a$ and $\lambda = d$.

Thus the eigenvalues of a triangular matrix are the diagonal entries.

Exercise 15

Find the eigenvalues and corresponding eigenvectors of the upper triangular matrix $\begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$.

Symmetric matrices

See Unit 4.

A matrix is *symmetric* if it is equal to its transpose – that is, if the entries are symmetric about the leading diagonal. Thus a 2×2 symmetric matrix has the form

$$\mathbf{A} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}.$$

It has characteristic equation

$$\lambda^2 - (a+d)\lambda + (ad - b^2) = 0,$$

and eigenvalues

$$\lambda = \tfrac{1}{2}\left(a + d \pm \sqrt{(a+d)^2 - 4(ad-b^2)}\right).$$

The discriminant is

In this module, $\mathbf{A}$ is always a real matrix, so a, b and d are real numbers.

$$(a+d)^2 - 4(ad-b^2) = (a^2 + 2ad + d^2) - 4ad + 4b^2 = (a-d)^2 + 4b^2,$$

which is the sum of two squares and therefore cannot be negative.

It follows that the eigenvalues of a real symmetric matrix are real. In Subsection 3.3, this result will be shown to hold for general $n \times n$ real symmetric matrices, along with an important property of their eigenvectors.

Exercise 16

Under what circumstances can a symmetric matrix $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$ have a repeated eigenvalue?

Non-invertible matrices

A matrix is *non-invertible* if and only if its determinant is 0. Thus a 2×2 non-invertible matrix has the form

See Unit 4.

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{where } ad - bc = 0.$$

However, we know that if λ_1 and λ_2 are the eigenvalues of $\mathbf{A}$, then

$$\lambda_1 \lambda_2 = \det \mathbf{A} = 0.$$

See equations (15).

It follows that a matrix is non-invertible if and only if at least one of its eigenvalues is 0. Also, a matrix is invertible if and only if all its eigenvalues are non-zero.

For example, if $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$, then $\det \mathbf{A} = 0$ and hence 0 is an eigenvalue of $\mathbf{A}$, as we saw in the discussion immediately before Exercise 3.

Since $\lambda_1 + \lambda_2 = \operatorname{tr} \mathbf{A} = 4$, the other eigenvalue is 4.

Non-invertible matrices

A square matrix $\mathbf{A}$ is non-invertible if and only if any of the following equivalent conditions holds:

- its determinant is zero
- its rows are linearly dependent
- its columns are linearly dependent
- the equation $\mathbf{Ax} = \mathbf{0}$ has a non-zero solution
- at least one of its eigenvalues is zero.

The third condition follows immediately from the second, since $\det \mathbf{A} = \det(\mathbf{A}^T)$.

Summary

We can summarise the results about eigenvalues in this section so far.

Eigenvalues – summary

- The eigenvalues of a triangular matrix are the diagonal entries.
- The eigenvalues of a real symmetric matrix are real.
- The sum of the eigenvalues of $\mathbf{A}$ is $\operatorname{tr} \mathbf{A}$.
- The product of the eigenvalues of $\mathbf{A}$ is $\det \mathbf{A}$.

Exercise 17

Without solving the characteristic equation, what can you say about the eigenvalues of each of the following matrices?

(a) $\mathbf{A} = \begin{pmatrix} 67 & 72 \\ 72 & -17 \end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix} 67 & 72 \\ 0 & -17 \end{pmatrix}$ (c) $\mathbf{A} = \begin{pmatrix} 288 & 72 \\ 72 & 18 \end{pmatrix}$

2.4 Eigenvalues of related matrices

For the rest of this section, we compare the eigenvalues of a matrix $\mathbf{A}$ with the eigenvalues of some related matrices. The results of our investigations are needed in Section 4. The following exercise leads our discussion.

Exercise 18

Let $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$. In Example 5 you saw that $(1 \quad 1)^T$ is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue 5, and that $(-2 \quad 1)^T$ is an eigenvector corresponding to the eigenvalue 2.

(a) Evaluate the following matrices. In each case, solve the characteristic equation, determine the eigenvalues, and compare these eigenvalues with the eigenvalues of $\mathbf{A}$.

(i) $\mathbf{A}^2$ (ii) $\mathbf{A}^{-1}$ (iii) $\mathbf{A} + 2\mathbf{I}$ (iv) $(\mathbf{A} - 4\mathbf{I})^{-1}$ (v) $3\mathbf{A}$

(b) Verify that the given eigenvectors of $\mathbf{A}$ are also eigenvectors of the matrices in part (a).

Exercise 18 illustrates some general results.

- If $\mathbf{v}$ is an eigenvector of a matrix $\mathbf{A}$ with eigenvalue λ, then $\mathbf{Av} = \lambda\mathbf{v}$. It follows that

  $$\mathbf{A}^2\mathbf{v} = \mathbf{A}(\mathbf{Av}) = \mathbf{A}(\lambda\mathbf{v}) = \lambda(\mathbf{Av}) = \lambda(\lambda\mathbf{v}) = \lambda^2\mathbf{v},$$

  which shows that $\mathbf{v}$ is an eigenvector of $\mathbf{A}^2$, with eigenvalue λ^2.

  More generally, if $\mathbf{v}$ is an eigenvector of a matrix $\mathbf{A}$ with eigenvalue λ, then $\mathbf{v}$ is also an eigenvector of $\mathbf{A}^k$ (for any positive integer k), with eigenvalue λ^k.

- If $\mathbf{v}$ is an eigenvector of a matrix $\mathbf{A}$ with eigenvalue λ, then $\mathbf{Av} = \lambda\mathbf{v}$. It follows that (provided that $\mathbf{A}$ is invertible, so $\mathbf{A}^{-1}$ exists)

  $$\mathbf{A}^{-1}(\mathbf{Av}) = \mathbf{A}^{-1}(\lambda\mathbf{v}) = \lambda\mathbf{A}^{-1}\mathbf{v},$$

  so $\mathbf{v} = \lambda\mathbf{A}^{-1}\mathbf{v}$. Dividing by λ (which cannot be zero since $\mathbf{A}$ is invertible) gives

  $$\mathbf{A}^{-1}\mathbf{v} = (1/\lambda)\mathbf{v},$$

  which shows that $\mathbf{v}$ is an eigenvector of $\mathbf{A}^{-1}$, with eigenvalue $1/\lambda$.

- If $\mathbf{v}$ is an eigenvector of a matrix $\mathbf{A}$ with eigenvalue λ, then $\mathbf{Av} = \lambda\mathbf{v}$. If p is any number, it follows that

  $$(\mathbf{A} - p\mathbf{I})\mathbf{v} = \mathbf{Av} - p\mathbf{Iv} = \lambda\mathbf{v} - p\mathbf{v} = (\lambda - p)\mathbf{v},$$

  which shows that $\mathbf{v}$ is an eigenvector of $\mathbf{A} - p\mathbf{I}$, with eigenvalue $\lambda - p$.

If the number p is *not* an eigenvalue of $\mathbf{A}$, then the eigenvalues $\lambda - p$ of $\mathbf{A} - p\mathbf{I}$ are non-zero, so $\mathbf{A} - p\mathbf{I}$ is invertible. Therefore we can multiply both sides of the equation

$$(\mathbf{A} - p\mathbf{I})\mathbf{v} = (\lambda - p)\mathbf{v}$$

on the left by $(\mathbf{A} - p\mathbf{I})^{-1}$ to obtain

$$\mathbf{v} = (\mathbf{A} - p\mathbf{I})^{-1}(\lambda - p)\mathbf{v},$$

and, dividing by $\lambda - p$,

$$(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{v} = (\lambda - p)^{-1}\mathbf{v}.$$

This shows that $\mathbf{v}$ is also an eigenvector of $(\mathbf{A} - p\mathbf{I})^{-1}$, with eigenvalue $(\lambda - p)^{-1}$.

- If $\mathbf{v}$ is an eigenvector of a matrix $\mathbf{A}$ with eigenvalue λ, then $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. It follows that for any number p, $p(\mathbf{A}\mathbf{v}) = p\lambda\mathbf{v}$, so $(p\mathbf{A})\mathbf{v} = (p\lambda)\mathbf{v}$, which shows that $\mathbf{v}$ is an eigenvector of $p\mathbf{A}$ with eigenvalue $p\lambda$.

Eigenvalues and eigenvectors of associated matrices

If $\mathbf{A}$ is an arbitrary matrix and λ is one of its eigenvalues, then:

- λ^k is an eigenvalue of $\mathbf{A}^k$ for any positive integer k
- if $\mathbf{A}$ is invertible, then λ^{-1} is an eigenvalue of $\mathbf{A}^{-1}$
- $\lambda - p$ is an eigenvalue of $\mathbf{A} - p\mathbf{I}$, for any number p
- $(\lambda - p)^{-1}$ is an eigenvalue of $(\mathbf{A} - p\mathbf{I})^{-1}$, for any number p that is not an eigenvalue of $\mathbf{A}$
- $p\lambda$ is an eigenvalue of $p\mathbf{A}$ for any number p.

In each case, an eigenvector of $\mathbf{A}$ is also an eigenvector of the associated matrix.

We will need these results in Section 4 and in later units of the module.

Exercise 19

(a) The eigenvalues of $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$ are 4 and -1. Write down the eigenvalues of each of the following.

(i) $\mathbf{A}^3$ (ii) $\mathbf{A}^{-1}$ (iii) $\mathbf{A} - 6\mathbf{I}$ (iv) $(\mathbf{A} + 3\mathbf{I})^{-1}$

(b) What can you say about the eigenvalues of $\mathbf{A} - 4\mathbf{I}$? What can you say about the inverse of $\mathbf{A} - 4\mathbf{I}$?

Exercise 20

(a) Find the eigenvalues and corresponding eigenvectors of the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix}$. Write down the eigenvalues and corresponding eigenvectors of the matrix $\mathbf{A}^{10}$.

(b) Write down the eigenvalues and corresponding eigenvectors of the matrix $\begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}$.

(*Hint*: Compare this matrix with the matrix $\mathbf{A}$ in part (a).)

Exercise 21

(a) Let θ be an angle that is not an integer multiple of π. Calculate the eigenvalues and eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

which represents an anticlockwise rotation through the angle θ.

(b) If θ is an integer multiple of π, what can be stated about the matrix $\mathbf{A}$? What can be deduced about the eigenvalues and eigenvectors of $\mathbf{A}$?

3 Eigenvalues and eigenvectors of 3×3 matrices

In this section we extend the ideas of Section 2 to deal with 3×3 matrices. In fact, it is possible to extend the treatment to $n \times n$ matrices, and with this in mind it is convenient to use the notation x_1, x_2 and x_3 (rather than x, y and z).

So we are interested in finding the eigenvalues and corresponding eigenvectors of a matrix such as

$$\begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix}.$$

As in the case of 2×2 matrices, we can verify that a given vector is an eigenvector. For example, $(1 \quad 0 \quad -1)^T$ is an eigenvector of the above matrix, since

$$\begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix} = -2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

and the corresponding eigenvalue is -2.

Exercise 22

Verify that $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ are eigenvectors of $\begin{pmatrix} 5 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, and find the corresponding eigenvalues.

The above exercise should have convinced you that it is easy to *verify* that a given vector is an eigenvector, but how are we to *find* such eigenvectors? Moreover, how are we to deal with the more general case of an $n \times n$ matrix? This section examines these questions.

3.1 Basic method

In this subsection we look at a method for finding the eigenvalues and corresponding eigenvectors of an arbitrary 3×3 matrix. Consider the following example.

Example 8

Find the eigenvalues and corresponding eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{pmatrix}.$$

Solution

We wish to find those non-zero vectors $\mathbf{v}$ that satisfy the equation

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

for some number λ. Writing $\mathbf{v} = (x_1 \quad x_2 \quad x_3)^T$, we have

$$\begin{pmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

which gives

$$\begin{cases} (3-\lambda)x_1 + 2x_2 + 2x_3 = 0, \\ 2x_1 + (2-\lambda)x_2 = 0, \\ 2x_1 + (4-\lambda)x_3 = 0, \end{cases} \tag{18}$$

These are the *eigenvector equations.*

which can be written as $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$. We are interested in non-zero solutions of these equations (i.e. solutions in which x_1, x_2 and x_3 are not all 0). The condition for such solutions to exist is that the determinant of the coefficient matrix $\mathbf{A} - \lambda\mathbf{I}$ is 0, that is,

$$\begin{vmatrix} 3-\lambda & 2 & 2 \\ 2 & 2-\lambda & 0 \\ 2 & 0 & 4-\lambda \end{vmatrix} = 0.$$

Expanding this determinant by the middle row gives

$$-2\begin{vmatrix}2 & 2\\ 0 & 4-\lambda\end{vmatrix} + (2-\lambda)\begin{vmatrix}3-\lambda & 2\\ 2 & 4-\lambda\end{vmatrix} - 0\begin{vmatrix}3-\lambda & 2\\ 2 & 0\end{vmatrix} = 0,$$

hence

$$-4(4-\lambda) + (2-\lambda)[(3-\lambda)(4-\lambda) - 4] = 0,$$

This is the *characteristic equation* of $\mathbf{A}$. For *any* 3×3 matrix, it will be a cubic equation.

which simplifies to the cubic equation

$$\lambda^3 - 9\lambda^2 + 18\lambda = 0.$$

We must now solve this equation. In general, it is difficult to solve a cubic equation algebraically, unless you use a computer algebra package. However, if you can spot one of the roots, then the task becomes considerably easier. In this case,

$$\lambda^3 - 9\lambda^2 + 18\lambda = \lambda(\lambda^2 - 9\lambda + 18),$$

One way of solving the quadratic equation is to notice that
$\lambda^2 - 9\lambda + 18 = (\lambda-3)(\lambda-6)$.

so $\lambda = 0$ is one of the roots. The others are obtained by solving the quadratic equation $\lambda^2 - 9\lambda + 18 = 0$ to give the roots 3 and 6. So the eigenvalues are 0, 3 and 6.

In order to find the corresponding eigenvectors, we substitute each of these eigenvalues into the eigenvector equations (18) as follows.

- For $\lambda = 6$, the matrix form of the eigenvector equations (18) is

$$\begin{pmatrix}-3 & 2 & 2\\ 2 & -4 & 0\\ 2 & 0 & -2\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} = \begin{pmatrix}0\\ 0\\ 0\end{pmatrix}.$$

While it may appear that we have three equations in the three unknowns x_1, x_2 and x_3, only two of the equations are distinct. You may recall a similar occurrence in Exercise 7(c) of Unit 4.

We saw in Subsection 2.3 that an equation of the form $\mathbf{Ax} = \mathbf{0}$ has a non-zero solution (and hence infinitely many solutions) if and only if the rows of $\mathbf{A}$ are linearly dependent. Here, this means that at least one of the following equations may be obtained from the other two:

$$\begin{aligned}-3x_1 + 2x_2 + 2x_3 &= 0,\\ 2x_1 - 4x_2 \phantom{{}+2x_3} &= 0,\\ 2x_1 \phantom{{}-4x_2} - 2x_3 &= 0.\end{aligned}$$

So we can find our desired solution by considering just two of the equations. Using the second and third, we have $x_1 - 2x_2 = 0$ and $x_1 - x_3 = 0$. Putting $x_3 = k$ and solving for x_1 and x_2, we obtain $x_1 = k$, $x_2 = \frac{1}{2}k$. Choosing $k = 2$ gives an answer avoiding fractions: $x_1 = 2$, $x_2 = 1$, $x_3 = 2$. This means that $(2\quad 1\quad 2)^T$ is an eigenvector corresponding to the eigenvalue $\lambda = 6$.

We should check that these values satisfy the original equations.

Alternatively, we may solve the eigenvector equations by Gaussian elimination. In this case, we have the following augmented matrix.

$$\left(\begin{array}{ccc|c}-3 & 2 & 2 & 0\\ 2 & -4 & 0 & 0\\ 2 & 0 & -2 & 0\end{array}\right)\begin{array}{l}\mathbf{R}_1\\ \mathbf{R}_2\\ \mathbf{R}_3\end{array}$$

Gaussian elimination is a more cumbersome process than the previous method, but it is illuminating. At the final stage of the elimination process, the bottom row of the augmented matrix consists entirely of zeros – and this must be the case because the rows of $\mathbf{A} - \lambda\mathbf{I}$ are linearly dependent.

Stage 1(a) We reduce to zero the elements below the leading diagonal in column 1.

$$\begin{array}{r} \\ \mathbf{R}_2 + \frac{2}{3}\mathbf{R}_1 \\ \mathbf{R}_3 + \frac{2}{3}\mathbf{R}_1 \end{array} \left(\begin{array}{ccc|c} -3 & 2 & 2 & 0 \\ 0 & -\frac{8}{3} & \frac{4}{3} & 0 \\ 0 & \frac{4}{3} & -\frac{2}{3} & 0 \end{array}\right) \begin{array}{l} \mathbf{R}_1 \\ \mathbf{R}_{2a} \\ \mathbf{R}_{3a} \end{array}$$

Stage 1(b) We reduce to zero the element below the leading diagonal in column 2.

$$\begin{array}{r} \\ \\ \mathbf{R}_{3a} + \frac{1}{2}\mathbf{R}_{2a} \end{array} \left(\begin{array}{ccc|c} -3 & 2 & 2 & 0 \\ 0 & -\frac{8}{3} & \frac{4}{3} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Stage 2 Back substitution gives $x_2 = \frac{1}{2}x_3$ and $x_1 = \frac{2}{3}(x_2 + x_3) = x_3$. We are free to choose x_3 as we please, so putting $x_3 = 2$, we find that $(2 \quad 1 \quad 2)^T$ is an eigenvector (as before).

- For $\lambda = 3$, substituting $\lambda = 3$ into equations (18), the eigenvector equations become $2x_2 + 2x_3 = 0$, $2x_1 - x_2 = 0$ and $2x_1 + x_3 = 0$. The first and second equations reduce to $x_3 = -x_2$ and $x_2 = 2x_1$. It follows that $(1 \quad 2 \quad -2)^T$ is an eigenvector corresponding to the eigenvalue 3.

Choose, for example, $x_1 = 1$, and use this to calculate x_2 and x_3.

- For $\lambda = 0$, substituting $\lambda = 0$ into equations (18), the eigenvector equations become $3x_1 + 2x_2 + 2x_3 = 0$, $2x_1 + 2x_2 = 0$ and $2x_1 + 4x_3 = 0$. These equations reduce to $x_2 = -x_1$ and $x_1 = -2x_3$. It follows that $(-2 \quad 2 \quad 1)^T$ is an eigenvector corresponding to the eigenvalue 0.

Choose, for example, $x_3 = 1$, and use this to calculate x_1 and x_2.

It is always worth checking that the values of x_1, x_2 and x_3 satisfy each equation.

In Example 8, we first found the eigenvalues by solving a cubic equation, then used these eigenvalues and the eigenvector equations to find the corresponding eigenvectors. We found the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{pmatrix}$$

by solving the equation

$$\begin{vmatrix} 3-\lambda & 2 & 2 \\ 2 & 2-\lambda & 0 \\ 2 & 0 & 4-\lambda \end{vmatrix} = 0.$$

Since the left-hand side of this equation is $\det(\mathbf{A} - \lambda\mathbf{I})$, it can be written in the form $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$.

As in the case of a 2×2 matrix, the values of $\operatorname{tr}\mathbf{A}$ and $\det\mathbf{A}$ provide a useful check on the calculations.

For Example 8, we see that $\operatorname{tr}\mathbf{A} = 3+2+4 = 9$, and the sum of the eigenvalues is also 9. Also, expanding by the middle row, we have

$$\det\mathbf{A} = -2\begin{vmatrix} 2 & 2 \\ 0 & 4 \end{vmatrix} + 2\begin{vmatrix} 3 & 2 \\ 2 & 4 \end{vmatrix} - 0\begin{vmatrix} 3 & 2 \\ 2 & 0 \end{vmatrix} = -16 + 16 = 0,$$

and the product of the eigenvalues is also 0.

We can extend and generalise Procedures 1 and 2 to $n \times n$ matrices.

Procedure 3 Finding eigenvalues and eigenvectors of an $n \times n$ matrix

To find the eigenvalues and eigenvectors of an $n \times n$ matrix $\mathbf{A}$, carry out the following steps.

1. Solve the characteristic equation

   $$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

   to determine the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.
2. Solve the corresponding eigenvector equations $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{v}_i = \mathbf{0}$ for each eigenvalue λ_i, to find a corresponding eigenvector $\mathbf{v}_i$.

The characteristic equation of an $n \times n$ matrix has n solutions (some of which may be repeated). If there are n distinct solutions, then there will be n linearly independent eigenvectors (but there may be fewer if any eigenvalue is repeated).

Exercise 23

The eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix}$$

are $\lambda = 1$, $\lambda = 2$ and $\lambda = 3$. Write down the eigenvector equations, and determine corresponding eigenvectors.

Exercise 24

Determine the characteristic equation of the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 6 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{pmatrix}.$$

Verify that $\lambda = 1$ is an eigenvalue of $\mathbf{A}$, and find a corresponding eigenvector. (You do not need to find any other eigenvalues.)

3.2 Finding eigenvalues by hand

For the matrices that arise from real applications, it is often impossible to calculate the eigenvalues and eigenvectors by hand (as opposed to using a computer), and we would generally use numerical techniques such as those discussed in the next section. However, finding eigenvalues and eigenvectors by hand is an important part of the learning process. In this module you can divide the exercises into two types: those for which it is easy to find the eigenvalues by hand, and the rest (for which you will probably need to use a computer algebra package). The examples and exercises in this section have been carefully chosen to be of the former kind.

Example 9

Find the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 5 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

Solution

The characteristic equation is

$$\begin{vmatrix} 5-\lambda & 0 & 0 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} = 0.$$

Expanding the determinant by the first row gives

$$(5-\lambda)\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0,$$

hence $(5-\lambda)(\lambda^2 - 4\lambda + 3) = 0$. So $5 - \lambda = 0$ or $\lambda^2 - 4\lambda + 3 = 0$. Hence one solution is $\lambda = 5$. The quadratic equation $\lambda^2 - 4\lambda + 3 = 0$ has roots 1 and 3.

Thus the eigenvalues are $\lambda = 5$, $\lambda = 3$ and $\lambda = 1$.

Exercise 25

Find the eigenvalues of each of the following matrices.

(a) $\mathbf{A} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 6 & 2 & 0 \\ 5 & 4 & 3 \end{pmatrix}$

Often the most arduous part of such problems is the expansion of the determinant, but a judicious choice of which row or column to expand by can reduce the work, as the following example shows.

Example 10

Find the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 5 & 2 & 0 \\ 2 & 5 & 0 \\ -3 & 4 & 6 \end{pmatrix}.$$

Solution

The characteristic equation is

$$\begin{vmatrix} 5-\lambda & 2 & 0 \\ 2 & 5-\lambda & 0 \\ -3 & 4 & 6-\lambda \end{vmatrix} = 0.$$

Since the third column contains two zeros, we expand the determinant by the third column to give

$$(6-\lambda)\begin{vmatrix} 5-\lambda & 2 \\ 2 & 5-\lambda \end{vmatrix} = 0,$$

hence $(6-\lambda)((5-\lambda)^2-4)=0$. So $\lambda-6=0$ or $(5-\lambda)^2-4=0$. Hence one solution is $\lambda=6$. The quadratic equation $(5-\lambda)^2-4=0$ can be rewritten as $(5-\lambda)^2=4$, so $5-\lambda=\pm 2$, and the two other solutions are $\lambda=7$ and $\lambda=3$.

Thus the eigenvalues are $\lambda=6$, $\lambda=7$ and $\lambda=3$.

Exercise 26

Find the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 8 & 0 & -5 \\ 9 & 3 & -6 \\ 10 & 0 & -7 \end{pmatrix}.$$

Exercise 27

See Subsection 2.3.

Verify that the eigenvalues of the triangular matrices

$$\begin{pmatrix} a & 0 & 0 \\ d & b & 0 \\ e & f & c \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a & d & e \\ 0 & b & f \\ 0 & 0 & c \end{pmatrix}$$

are the diagonal entries a, b and c.

Exercise 28

Verify that the sum of the eigenvalues is $\operatorname{tr}\mathbf{A}$ for the matrices $\mathbf{A}$ in the following. In each case, find the value of $\det\mathbf{A}$, and verify that this is the product of the eigenvalues.

(a) Example 9 (b) Example 10

(c) Exercise 23 (d) Exercise 25 (e) Exercise 26

The eigenvalues of a 3×3 matrix can be real and distinct (as in Exercise 23), or real and repeated (as in Exercise 26), or one may be real and the other two form a complex conjugate pair – as in the following example.

Repeated real eigenvalues may be repeated once,

$$\lambda_1 = \lambda_2 \neq \lambda_3,$$

or twice,

$$\lambda_1 = \lambda_2 = \lambda_3.$$

Example 11

Find the eigenvalues and corresponding eigenvectors of the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Solution

The characteristic equation is

$$\begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & -1 & -\lambda \end{vmatrix} = 0.$$

Expanding the determinant by the first row gives

$$(1-\lambda)\begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = (1-\lambda)(\lambda^2+1) = 0,$$

so $1-\lambda = 0$ or $\lambda^2 + 1 = 0$. The quadratic equation $\lambda^2 + 1 = 0$ has roots $\lambda = i$ and $\lambda = -i$, where $i^2 = -1$. Thus the eigenvalues are $\lambda = 1$, $\lambda = i$ and $\lambda = -i$.

The eigenvector equations are

$$\begin{aligned} (1-\lambda)x_1 \qquad\qquad\qquad &= 0, \\ -\lambda x_2 + \quad x_3 &= 0, \\ -\quad x_2 - \lambda x_3 &= 0. \end{aligned}$$

- For $\lambda = 1$, the eigenvector equations become $0 = 0$, $-x_2 + x_3 = 0$ and $-x_2 - x_3 = 0$, which reduce to the equations $x_2 = 0$ and $x_3 = 0$. However, there is no constraint on the value of x_1, so we may choose it as we please provided that it is non-zero. It follows that a corresponding eigenvector is $(1 \quad 0 \quad 0)^T$.
- For $\lambda = i$, the eigenvector equations become $(1-i)x_1 = 0$, $-ix_2 + x_3 = 0$ and $-x_2 - ix_3 = 0$, which reduce to the equations $x_1 = 0$ and $ix_2 = x_3$. It follows that a corresponding eigenvector is $(0 \quad 1 \quad i)^T$.
- For $\lambda = -i$, the eigenvector equations become $(1+i)x_1 = 0$, $ix_2 + x_3 = 0$ and $-x_2 + ix_3 = 0$, which reduce to the equations $x_1 = 0$ and $-ix_2 = x_3$. It follows that a corresponding eigenvector is $(0 \quad 1 \quad -i)^T$.

In Example 11 we have three distinct eigenvalues and three distinct eigenvectors, but only the real eigenvalue $\lambda = 1$ gives rise to a real eigenvector.

Exercise 29

Calculate the eigenvalues and corresponding eigenvectors of the following matrices.

(a) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 1 & -1 \\ 0 & -3 & 2 \\ 0 & 0 & 4 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & 2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (e) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & -2 & 5 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & -2 \end{pmatrix}$

We have dealt with 3×3 matrices for which the eigenvalues can be found easily, and it is reasonable to solve such problems by hand. Generally, the larger matrices that arise in practical problems are better dealt with in an entirely different fashion, as we will see in Section 4. However, our experience of finding the eigenvalues and eigenvectors of the simpler kinds of 2×2 and 3×3 matrices will not be wasted, for it will allow us to see what types of solution we may expect for larger matrices.

In each of the cases considered in this section, if a 3×3 matrix has three distinct eigenvalues, then the corresponding eigenvectors are linearly independent. More generally, any $n \times n$ matrix with n distinct eigenvalues has n linearly independent eigenvectors, although we do not prove this here.

If eigenvalues are repeated, then the situation becomes more complicated. We leave that discussion to the next unit, when we return to this topic in the context of solving systems of differential equations.

3.3 Real symmetric matrices

The eigenvalues and eigenvectors of any real symmetric $n \times n$ matrix have some important properties, which we now investigate: namely, that their eigenvalues are real and that their eigenvectors are mutually *orthogonal* – that is, perpendicular to one another – or can be made orthogonal.

Orthogonal vectors

If two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$ satisfy the condition

$$\mathbf{u}^T\mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n = 0,$$

then they are said to be **orthogonal** to one another.

You may think of $\mathbf{u}^T\mathbf{v}$ as a $1 \times n$ matrix multiplying an $n \times 1$ matrix.

Recall from Unit 2 that the component form of the dot product (generalised to n-component vectors) is

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n,$$

so

$$\mathbf{u}^T \mathbf{v} = \mathbf{u} \cdot \mathbf{v} = |\mathbf{u}|\,|\mathbf{v}| \cos\theta,$$

where θ ($0 \leq \theta \leq \pi$) is the angle between $\mathbf{u}$ and $\mathbf{v}$. Thus the orthogonality condition $\mathbf{u}^T\mathbf{v} = \mathbf{u} \cdot \mathbf{v} = 0$ implies that $\theta = \pi/2$, that is, $\mathbf{u}$ and $\mathbf{v}$ are perpendicular.

We begin by showing that the eigenvalues of a real symmetric matrix are real. This was already demonstrated for 2×2 real symmetric matrices in Subsection 2.3, but we now show that the property holds for general $n \times n$ real symmetric matrices.

We start from the eigenvalue equation

$$\mathbf{A}\mathbf{v} = \lambda \mathbf{v}, \tag{19}$$

where $\mathbf{A}$ is a real symmetric $n \times n$ matrix, and assume that the eigenvalues are in general complex, but we will eventually arrive at an expression that can be satisfied only if they are real.

Premultiply both sides of equation (19) by the transpose of the complex conjugate of $\mathbf{v}$ to obtain

$$\overline{\mathbf{v}}^T \mathbf{A}\mathbf{v} = \lambda \overline{\mathbf{v}}^T \mathbf{v}. \tag{20}$$

Note that $\overline{\mathbf{v}}$ denotes the complex conjugate of $\mathbf{v}$, i.e. if $\mathbf{v} = \mathbf{a} + i\mathbf{b}$, then $\overline{\mathbf{v}} = \mathbf{a} - i\mathbf{b}$, where $\mathbf{a}$ and $\mathbf{b}$ are real vectors.

Now transpose both sides of equation (19), which gives

$$\mathbf{v}^T \mathbf{A}^T = \lambda \mathbf{v}^T,$$

where we have used the transpose property $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ as explained in Subsection 2.1 of Unit 4 (rules for transposes of matrices). Taking the complex conjugate of this last equation, and noting that $\mathbf{A}$ is real so that $\overline{\mathbf{A}} = \mathbf{A}$, gives

$$\overline{\mathbf{v}}^T \mathbf{A}^T = \overline{\lambda}\,\overline{\mathbf{v}}^T,$$

and then postmultiplying both sides by $\mathbf{v}$ yields

$$\overline{\mathbf{v}}^T \mathbf{A}^T \mathbf{v} = \overline{\lambda}\,\overline{\mathbf{v}}^T \mathbf{v}. \tag{21}$$

In this module, we only ever consider real matrices (apart from complex vectors). However, if $\mathbf{A}$ is complex, then its complex conjugate $\overline{\mathbf{A}}$ is obtained by replacing every element in $\mathbf{A}$ by the element's complex conjugate. An example of a complex matrix is $\mathbf{A} = \begin{pmatrix} 2-3i & 2 \\ 4i & 0 \end{pmatrix}$, in which case $\overline{\mathbf{A}} = \begin{pmatrix} 2+3i & 2 \\ -4i & 0 \end{pmatrix}$.

Subtracting equation (21) from equation (20) gives

$$\overline{\mathbf{v}}^T \mathbf{A}\mathbf{v} - \overline{\mathbf{v}}^T \mathbf{A}^T \mathbf{v} = \lambda \overline{\mathbf{v}}^T \mathbf{v} - \overline{\lambda}\,\overline{\mathbf{v}}^T \mathbf{v}. \tag{22}$$

But $\mathbf{A}$ is symmetric, which means that $\mathbf{A}^T = \mathbf{A}$, so the left-hand side of equation (22) is zero, and

$$\lambda \overline{\mathbf{v}}^T \mathbf{v} - \overline{\lambda}\,\overline{\mathbf{v}}^T \mathbf{v} = (\lambda - \overline{\lambda})\overline{\mathbf{v}}^T \mathbf{v} = 0. \tag{23}$$

Finally, by writing $\mathbf{v} = \mathbf{a} + i\mathbf{b} \neq \mathbf{0}$, where $\mathbf{a}$ and $\mathbf{b}$ are real vectors, and noting that $(\mathbf{a} - i\mathbf{b})^T = \mathbf{a}^T - i\mathbf{b}^T$ and $\mathbf{a}^T\mathbf{b} = \mathbf{b}^T\mathbf{a}$, we have

$$\begin{aligned} \overline{\mathbf{v}}^T \mathbf{v} = (\mathbf{a} - i\mathbf{b})^T(\mathbf{a} + i\mathbf{b}) &= \mathbf{a}^T\mathbf{a} - i\mathbf{b}^T\mathbf{a} + i\mathbf{a}^T\mathbf{b} + \mathbf{b}^T\mathbf{b} \\ &= \mathbf{a}^T\mathbf{a} + \mathbf{b}^T\mathbf{b} > 0, \end{aligned}$$

so equation (23) can be satisfied only if $\overline{\lambda} = \lambda$, which holds *only* if λ is real.

At least one of $\mathbf{a}$ and $\mathbf{b}$ is not the zero vector, so at least one of $\mathbf{a}^T\mathbf{a}$ and $\mathbf{b}^T\mathbf{b}$ is non-zero (and in fact positive, since $\mathbf{u}^T\mathbf{u} = \mathbf{u} \cdot \mathbf{u} = |\mathbf{u}|^2 > 0$ for any non-zero vector $\mathbf{u}$.

Having established that the eigenvalues of general real symmetric matrices are real, we now move on to the second property by first showing that the eigenvectors corresponding to *distinct* eigenvalues of a real symmetric matrix are orthogonal.

Let λ_i be an eigenvalue of the real symmetric matrix $\mathbf{A}$ with corresponding eigenvector $\mathbf{v}_i$, and let λ_j be an eigenvalue with corresponding eigenvector $\mathbf{v}_j$ such that $\lambda_i \neq \lambda_j$. Thus we have

$$\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i, \tag{24}$$

$$\mathbf{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j. \tag{25}$$

Premultiplying both sides of equation (24) by $\mathbf{v}_j^T$ gives

$$\mathbf{v}_j^T \mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_j^T \mathbf{v}_i, \tag{26}$$

and taking the transpose of equation (25) followed by postmultiplying by $\mathbf{v}_i$ gives

$$\mathbf{v}_j^T \mathbf{A}^T \mathbf{v}_i = \lambda_j \mathbf{v}_j^T \mathbf{v}_i. \tag{27}$$

Finally, subtracting equation (27) from equation (26), and using the fact that $\mathbf{A}^T = \mathbf{A}$, leads to

$$(\lambda_i - \lambda_j)\mathbf{v}_j^T \mathbf{v}_i = 0,$$

and since $\lambda_i \neq \lambda_j$, it must follow that $\mathbf{v}_j^T \mathbf{v}_i = 0$ or, in other words, $\mathbf{v}_i$ and $\mathbf{v}_j$ are orthogonal.

Now, we have established that eigenvectors corresponding to *distinct* eigenvalues of a symmetric matrix are orthogonal. It turns out that eigenvectors corresponding to a repeated eigenvalue (of general multiplicity) of a symmetric matrix are linearly independent and can be selected to be mutually orthogonal. The proof of this for general $n \times n$ matrices is beyond the scope of this module. However, recall from Exercise 16 that a 2×2 symmetric matrix with a repeated eigenvalue must necessarily be a scalar multiple of the 2×2 identity matrix, for which any two linearly independent vectors are eigenvectors and we are free to select two that are orthogonal, such as $(1 \quad 0)^T$ and $(0 \quad 1)^T$.

Eigenvalues and eigenvectors of real symmetric matrices

If $\mathbf{A}$ is a real symmetric $n \times n$ matrix, then:

- the eigenvalues of $\mathbf{A}$ are real
- the eigenvectors corresponding to distinct eigenvalues of $\mathbf{A}$ are orthogonal, and those corresponding to repeated eigenvalues are linearly independent and can be chosen to be orthogonal.

These properties of the eigenvalues and eigenvectors of real symmetric matrices are important in their own right and have many applications in applied mathematics. For this module, they are used in Unit 7 for classifying the stationary points of functions of several variables.

You saw an example of a real symmetric matrix having real eigenvalues and orthogonal eigenvectors in Exercise 8. Some more examples are given in the following exercises.

Exercise 30

Determine the eigenvalues and eigenvectors of the following matrices, and verify that their eigenvalues are real and that their eigenvectors are orthogonal.

(a) $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}$ (c) $\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

Exercise 31

Determine the eigenvalues and eigenvectors of the matrix

$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix},$$

and verify that the eigenvalues are real. In the case of a repeated eigenvalue, show that it has two linearly independent eigenvectors that can be chosen to be orthogonal, and that these are orthogonal to the eigenvector corresponding to the other eigenvalue.

4 Iterative methods

Finding the eigenvalues and eigenvectors of a 3×3 matrix using the method of Section 3 can be a quite laborious process, and the calculations become progressively more difficult for larger matrices.

In this section we show how we can often find approximations to real eigenvectors and their corresponding eigenvalues by iteration – that is, by choosing a vector and applying the matrix repeatedly.

4.1 Approximating eigenvectors

In the Introduction we considered a migration problem in which the towns Exton and Wyeville have a regular interchange of population. We saw that if x_n and y_n denote the respective populations of Exton and Wyeville at the beginning of year n, then the corresponding populations at the beginning of year $n+1$ are given by the matrix equation

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix}.$$

Using this equation, we saw that if the initial populations are $x_0 = 10\,000$ and $y_0 = 8000$, then the populations in successive years are

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 10\,000 \\ 8\,000 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 10\,600 \\ 7\,400 \end{pmatrix}, \quad \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 11\,020 \\ 6\,980 \end{pmatrix}, \quad \ldots .$$

As n increases, the sequence of vectors $(x_n \quad y_n)^T$ converges to the vector $(x \quad y)^T = (12\,000 \quad 6000)^T$, which is an eigenvector of the above 2×2 matrix.

More generally, suppose that we wish to find the eigenvectors of a given matrix $\mathbf{A}$, and that we have an initial estimate $\mathbf{e}_0$ for an eigenvector.

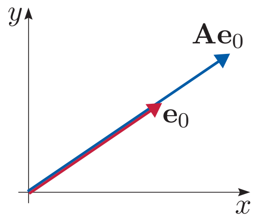


Figure 6 Finding $\mathbf{Ae}_0$ when $\mathbf{e}_0$ is an eigenvector

It may happen that $\mathbf{e}_0$ is an eigenvector of $\mathbf{A}$, and $\mathbf{Ae}_0$ is in the same direction as $\mathbf{e}_0$, as in Figure 6.

If, as usually happens, $\mathbf{e}_0$ is not an eigenvector, then we calculate the vector $\mathbf{e}_1 = \mathbf{Ae}_0$ and then another vector $\mathbf{e}_2$, defined by $\mathbf{e}_2 = \mathbf{Ae}_1 = \mathbf{A}^2\mathbf{e}_0$. Continuing in this way, we obtain a sequence of vectors

$$\mathbf{e}_0, \quad \mathbf{e}_1 = \mathbf{Ae}_0, \quad \mathbf{e}_2 = \mathbf{A}^2\mathbf{e}_0, \quad \mathbf{e}_3 = \mathbf{A}^3\mathbf{e}_0, \quad \mathbf{e}_4 = \mathbf{A}^4\mathbf{e}_0, \quad \ldots,$$

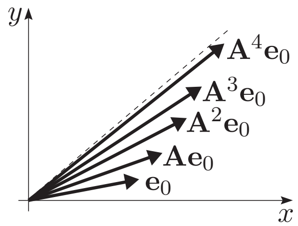


Figure 7 Repeatedly applying $\mathbf{A}$ to $\mathbf{e}_0$ converges to an eigenvector (along the dashed line)

as shown in Figure 7, where each vector in the sequence is obtained from the previous one by multiplying by the matrix $\mathbf{A}$. Often this simple method of repeatedly applying the matrix $\mathbf{A}$ produces a sequence of vectors that converges to an eigenvector.

Exercise 32

Given

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}, \quad \mathbf{e}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_{n+1} = \mathbf{Ae}_n, \; n = 0, 1, 2, \ldots,$$

calculate $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and $\mathbf{e}_4$.

From Example 5 we know that for $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$, $\mathbf{v}_1 = (1 \quad 1)^T$ is an eigenvector corresponding to the eigenvalue 5, and $\mathbf{v}_2 = (-2 \quad 1)^T$ is an eigenvector corresponding to the eigenvalue 2. We may suspect that the sequence $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \ldots$ in Exercise 32 converges to a scalar multiple of the eigenvector $(1 \quad 1)^T$, but how can we be *sure* that it does?

Since $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent eigenvectors, we can express any other vector as a linear combination of them (see Exercise 2 and the text following that exercise). So let us express our initial vector $\mathbf{e}_0 = (1 \quad 0)^T$ as $\mathbf{e}_0 = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$ for some numbers α and β. Then

α and β are known as the *components* of $\mathbf{v}_1$ and $\mathbf{v}_2$ in $\mathbf{e}_0$.

$$\mathbf{e}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

so

$$\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Multiplying both sides of this equation on the left by the inverse of the matrix on the left-hand side, we see that

A matrix formed from linearly independent vectors in this way always has an inverse.

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \frac{1}{3}\begin{pmatrix}1 & 2\\ -1 & 1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} = \begin{pmatrix}\frac{1}{3}\\ -\frac{1}{3}\end{pmatrix},$$

so $\alpha = \frac{1}{3}$, $\beta = -\frac{1}{3}$ and

$$\mathbf{e}_0 = \tfrac{1}{3}\mathbf{v}_1 - \tfrac{1}{3}\mathbf{v}_2.$$

We can now express $\mathbf{e}_1, \mathbf{e}_2, \ldots$ in terms of $\mathbf{v}_1$ and $\mathbf{v}_2$. Since $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors, and we know that $\mathbf{A}\mathbf{v}_1 = 5\mathbf{v}_1$ and $\mathbf{A}\mathbf{v}_2 = 2\mathbf{v}_2$, we have

$$\mathbf{e}_1 = \mathbf{A}\mathbf{e}_0 = \tfrac{1}{3}\mathbf{A}\mathbf{v}_1 - \tfrac{1}{3}\mathbf{A}\mathbf{v}_2 = \tfrac{1}{3}(5\mathbf{v}_1) - \tfrac{1}{3}(2\mathbf{v}_2).$$

Applying $\mathbf{A}$ repeatedly gives

$$\mathbf{e}_2 = \mathbf{A}^2\mathbf{e}_0 = \tfrac{1}{3}(5\mathbf{A}\mathbf{v}_1) - \tfrac{1}{3}(2\mathbf{A}\mathbf{v}_2) = \tfrac{1}{3}(5^2\mathbf{v}_1) - \tfrac{1}{3}(2^2\mathbf{v}_2),$$

$$\mathbf{e}_3 = \mathbf{A}^3\mathbf{e}_0 = \tfrac{1}{3}(5^2\mathbf{A}\mathbf{v}_1) - \tfrac{1}{3}(2^2\mathbf{A}\mathbf{v}_2) = \tfrac{1}{3}(5^3\mathbf{v}_1) - \tfrac{1}{3}(2^3\mathbf{v}_2),$$

$\mathbf{e}_2 = \mathbf{A}\mathbf{e}_1 = \mathbf{A}^2\mathbf{e}_0$

and so on.

In general, we can write

$$\mathbf{e}_n = \mathbf{A}^n\mathbf{e}_0 = \tfrac{1}{3}(5^n\mathbf{v}_1) - \tfrac{1}{3}(2^n\mathbf{v}_2).$$

For example,

$$\begin{aligned}\mathbf{e}_{10} = \mathbf{A}^{10}\mathbf{e}_0 &= \tfrac{1}{3}(5^{10}\mathbf{v}_1) - \tfrac{1}{3}(2^{10}\mathbf{v}_2)\\ &\simeq 3\,255\,208\mathbf{v}_1 - 341\mathbf{v}_2\\ &\simeq 3\,255\,208(\mathbf{v}_1 - 0.000\,105\mathbf{v}_2).\end{aligned}$$

As you can see, the powers of 5 rapidly become much larger than the powers of 2, and for large values of n we can ignore the latter, to give the approximation

$$\mathbf{e}_n = \mathbf{A}^n\mathbf{e}_0 \simeq \tfrac{1}{3}(5^n\mathbf{v}_1).$$

This is a scalar multiple of an eigenvector that corresponds to the eigenvalue of larger magnitude.

Thus repeatedly applying $\mathbf{A}$ leads to an approximation of an eigenvector – an eigenvector corresponding to the eigenvalue of larger magnitude.

Taking another example, you can show that the matrix

$$\begin{pmatrix}4 & 2\\ -7 & -5\end{pmatrix}$$

has eigenvalues -3 and 2, and that the above method will give approximations to an eigenvector corresponding to the eigenvalue -3 because powers of -3 eventually dominate powers of 2.

Exercise 33

Consider

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{e}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_{n+1} = \mathbf{A}\mathbf{e}_n,\ n = 0, 1, 2, \ldots.$$

(a) Calculate $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$.

(b) Given $\mathbf{e}_{10} = (29\,525 \quad 29\,524)^T$, calculate $\mathbf{e}_{11}$.

(c) Use Procedures 1 and 2 to find the eigenvalues λ_1 and λ_2, and corresponding eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$, of $\mathbf{A}$.

(d) Express $\mathbf{e}_0$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

(e) Express $\mathbf{e}_1$, $\mathbf{e}_2$, and hence $\mathbf{e}_n$, as linear combinations of $\mathbf{v}_1$ and $\mathbf{v}_2$.

(f) To which eigenvector does the sequence $\{\mathbf{e}_n\}$ provide an approximation?

The above technique provides us with increasingly accurate approximations to one of the eigenvectors of a 2×2 matrix. But the most significant aspect of the method is that it is possible to extend it to matrices of *any* size. However, there are difficulties, and you should be aware of them before we proceed. You may have noticed in the previous exercise that the components of an approximation to an eigenvector may be quite large and, had we attempted to calculate $\mathbf{e}_{20}$, we would have found that $\mathbf{e}_{20} = (1\,743\,392\,201 \quad 1\,743\,392\,200)^T$. For larger values of n, $\mathbf{e}_n$ involves even larger numbers. We will see that this difficulty is easily overcome, but there are other difficulties. Table 1 shows five examples, each exhibiting a different problem.

The $\mathbf{e}_0$ are normally chosen rather arbitrarily.

Table 1

	Matrix	Eigenvalues	Corresponding eigenvectors	Initial vector and nth approximation
(a)	$\begin{pmatrix} 1 & 4 \\ 9 & 1 \end{pmatrix}$	7 -5	$\mathbf{v}_1 = (2 \quad 3)^T$ $\mathbf{v}_2 = (2 \quad -3)^T$	$\mathbf{e}_0 = (1 \quad 0)^T$ $\mathbf{e}_n = \frac{1}{4}(7)^n\mathbf{v}_1 + \frac{1}{4}(-5)^n\mathbf{v}_2$
(b)	$\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$	2 -2	$\mathbf{v}_1 = (1 \quad 0)^T$ $\mathbf{v}_2 = (0 \quad 1)^T$	$\mathbf{e}_0 = (1 \quad 1)^T$ $\mathbf{e}_n = 2^n\mathbf{v}_1 + (-2)^n\mathbf{v}_2$
(c)	$\begin{pmatrix} 1.2 & -0.2 \\ 0.3 & 0.7 \end{pmatrix}$	1 0.9	$\mathbf{v}_1 = (1 \quad 1)^T$ $\mathbf{v}_2 = (2 \quad 3)^T$	$\mathbf{e}_0 = (1 \quad 0)^T$ $\mathbf{e}_n = 3\mathbf{v}_1 - (0.9)^n\mathbf{v}_2$
(d)	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	i $-i$	$\mathbf{v}_1 = (1 \quad -i)^T$ $\mathbf{v}_2 = (1 \quad i)^T$	$\mathbf{e}_0 = (1 \quad 0)^T$ $\mathbf{e}_n = \frac{1}{2}(i)^n\mathbf{v}_1 + \frac{1}{2}(-i)^n\mathbf{v}_2$
(e)	$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	3 2 1	$\mathbf{v}_1 = (1 \quad 0 \quad 0)^T$ $\mathbf{v}_2 = (0 \quad 1 \quad 0)^T$ $\mathbf{v}_3 = (0 \quad 0 \quad 1)^T$	$\mathbf{e}_0 = (0 \quad 1 \quad 1)^T$ $\mathbf{e}_n = 2^n\mathbf{v}_2 + \mathbf{v}_3$

In row (a) of the table, the elements of $\mathbf{e}_n$ become very large: for example,

$$\mathbf{e}_5 = (6841 \quad 14\,949)^T$$

and

$$\mathbf{e}_{10} = (146\,120\,437 \quad 204\,532\,218)^T.$$

This may cause difficulties in the calculations. You may already suspect how this difficulty may be overcome. We are interested in only the *directions* of the vectors, and rescaling a vector does not change its direction. So dividing both components of $\mathbf{e}_{10}$ by 204 532 218, we obtain the vector $(0.7144 \quad 1)^T$ (to four decimal places) as the estimate of an eigenvector.

In row (b) we have a rather more fundamental problem. The eigenvalues have the same magnitude, so as n increases, neither the term involving $\mathbf{v}_1$ nor the term involving $\mathbf{v}_2$ becomes dominant, thus the iteration does not converge.

The eigenvalues in row (c) are certainly not equal, but they are similar in magnitude. This means that we need to choose a very large value of n in order to obtain a good approximation to $\mathbf{v}_1$ (the eigenvector corresponding to the eigenvalue of larger magnitude).

In row (d) we have complex eigenvalues, and as you might expect, a sequence of real vectors cannot converge to a complex eigenvector.

You always generate a real sequence (unless you start with a complex $\mathbf{e}_0$).

In row (e) we see that the sequence $\mathbf{e}_n$ converges to the eigenvector $\mathbf{v}_2$, when we might expect it to converge to $\mathbf{v}_1$ (the eigenvector corresponding to the eigenvalue of largest magnitude). This is because the original estimate $\mathbf{e}_0$ contains no component of $\mathbf{v}_1$, so the same is true of all subsequent estimates. (Such a difficulty is unlikely to arise in practice because rounding errors in the calculation of the iterates will normally ensure that the component of $\mathbf{v}_1$ is not exactly zero.)

Exercise 34

In the migration problem first discussed in the Introduction, we have

$$\mathbf{A} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}, \quad \mathbf{e}_0 = \begin{pmatrix} 10\,000 \\ 8\,000 \end{pmatrix}.$$

Eigenvectors of $\mathbf{A}$ are $\mathbf{v}_1 = (2 \quad 1)^T$ with corresponding eigenvalue 1, and $\mathbf{v}_2 = (1 \quad -1)^T$ with corresponding eigenvalue 0.7.

You found these eigenvalues and eigenvectors in Exercise 12.

(a) Write $\mathbf{e}_0$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

(b) Obtain an expression for $\mathbf{e}_n = \mathbf{A}^n\mathbf{e}_0$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

(c) Explain what happens as n becomes large.

4.2 Iterative techniques

Suppose that we are given a square matrix whose eigenvalues are known to be real and distinct in magnitude. Though we do not know their values, assume that the eigenvalues are listed in decreasing order of magnitude. For example, eigenvalues 5, -1 and -4 would be listed in the order

$|5| > |-4| > |-1|$

$$\lambda_1 = 5, \quad \lambda_2 = -4, \quad \lambda_3 = -1.$$

Using the ideas of the previous subsection, together with the results at the end of Section 2, we can approximate all the eigenvalues and corresponding eigenvectors. We start with the eigenvector corresponding to the eigenvalue of largest magnitude, then we show how the other eigenvectors and eigenvalues are approximated.

Eigenvalue of largest magnitude: direct iteration

In order to approximate an eigenvector corresponding to the eigenvalue λ_{max} of largest magnitude, we use the approach that we employed at the beginning of the previous subsection. We start with a square matrix $\mathbf{A}$ and a vector $\mathbf{e}_0$, and successively calculate the new vectors

$$\mathbf{e}_1 = \mathbf{A}\mathbf{e}_0, \quad \mathbf{e}_2 = \mathbf{A}\mathbf{e}_1, \quad \mathbf{e}_3 = \mathbf{A}\mathbf{e}_2, \quad \ldots,$$

which is equivalent to writing $\mathbf{e}_n = \mathbf{A}^n\mathbf{e}_0$, $n = 1, 2, 3, \ldots$.

For example, if

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \quad \text{and} \quad \mathbf{e}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

then, as you saw in Exercise 32,

$$\mathbf{e}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 11 \\ 7 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 47 \\ 39 \end{pmatrix}, \quad \mathbf{e}_4 = \begin{pmatrix} 219 \\ 203 \end{pmatrix}, \quad \ldots.$$

The main difficulty with this method is that the components of $\mathbf{e}_n$ can rapidly become very large (or very small). But we can overcome this problem by setting α_n to be the component of largest magnitude in $\mathbf{e}_n$. Then dividing the vector $\mathbf{e}_n$ by α_n ensures that the vector $\mathbf{e}_n/\alpha_n$ has components that are less than, or equal to, 1 in magnitude. This process is called *scaling*, giving a *scaled vector*.

For the above vectors, $\alpha_1 = 3$, $\alpha_2 = 11$, $\alpha_3 = 47$ and $\alpha_4 = 219$, and we obtain the sequence of vectors

$$\begin{pmatrix} 1 \\ \frac{1}{3} \end{pmatrix}, \quad \begin{pmatrix} 1 \\ \frac{7}{11} \end{pmatrix}, \quad \begin{pmatrix} 1 \\ \frac{39}{47} \end{pmatrix}, \quad \begin{pmatrix} 1 \\ \frac{203}{219} \end{pmatrix}$$

(with the same directions as the sequence found in Exercise 32).

The components of our final answer will now be of reasonable size. If this scaling process is applied only at the final stage, we may still encounter very large components in the intermediate calculations. This difficulty is avoided by applying the scaling process at each step of the calculation, as in the following procedure.

Procedure 4 Direct iteration

For any square matrix $\mathbf{A}$ for which the eigenvalue $\lambda_{\max}$ of largest magnitude is real (and distinct in magnitude from any other eigenvalue), choose any vector $\mathbf{e}_0$.

For $n = 0, 1, 2, \ldots$, carry out the following steps.

1. Calculate $\mathbf{z}_{n+1} = \mathbf{A}\mathbf{e}_n$.
2. Find α_{n+1}, the component of largest magnitude of $\mathbf{z}_{n+1}$.
3. Put $\mathbf{e}_{n+1} = \mathbf{z}_{n+1}/\alpha_{n+1}$.

For sufficiently large n, $\mathbf{e}_n$ will be a good approximation to an eigenvector corresponding to the eigenvalue of largest magnitude, provided that $\mathbf{e}_0$ has a non-zero component of the required eigenvector. If the sequence α_n converges, then it converges to $\lambda_{\max}$.

For example, a matrix with eigenvalues 3 and -3 would not qualify, because the eigenvalues are equal in magnitude.

If $\lambda_{\max}$ is complex, then its complex conjugate will also be an eigenvalue of the same magnitude, so this procedure works only when the eigenvalue of largest magnitude is real.

The final sentence of the above procedure can be deduced from the fact that at each stage of the calculation we have $\mathbf{A}\mathbf{e}_n = \mathbf{z}_{n+1} = \alpha_{n+1}\mathbf{e}_{n+1}$. If $\mathbf{e}_n$ converges to a vector $\mathbf{e}$ and α_n converges to a number α, then, in the limit, we have $\mathbf{A}\mathbf{e} = \alpha\mathbf{e}$, so $\mathbf{e}$ is an eigenvector corresponding to the eigenvalue α. But we know that $\mathbf{e}$ is an eigenvector corresponding to the eigenvalue of largest magnitude, so α must be this eigenvalue.

Example 12

Given

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \quad \text{and} \quad \mathbf{e}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

use Procedure 4 to find $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$. (From Example 5, the eigenvalues of $\mathbf{A}$ are 5 and 2, with corresponding eigenvectors $(1 \quad 1)^T$ and $(-2 \quad 1)^T$, respectively.)

Solution

First iteration:

$$\mathbf{z}_1 = \mathbf{A}\mathbf{e}_0 = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

$$\alpha_1 = 3,$$

$$\mathbf{e}_1 = \frac{\mathbf{z}_1}{\alpha_1} = \tfrac{1}{3}\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{3} \end{pmatrix}.$$

Second iteration:

$$\mathbf{z}_2 = \mathbf{A}\mathbf{e}_1 = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{11}{3} \\ \frac{7}{3} \end{pmatrix},$$

$$\alpha_2 = \tfrac{11}{3},$$

$$\mathbf{e}_2 = \frac{\mathbf{z}_2}{\alpha_2} = \tfrac{3}{11} \begin{pmatrix} \frac{11}{3} \\ \frac{7}{3} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{7}{11} \end{pmatrix}.$$

Third iteration:

$$\mathbf{z}_3 = \mathbf{A}\mathbf{e}_2 = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{7}{11} \end{pmatrix} = \begin{pmatrix} \frac{47}{11} \\ \frac{39}{11} \end{pmatrix},$$

$$\alpha_3 = \tfrac{47}{11},$$

$$\mathbf{e}_3 = \frac{\mathbf{z}_3}{\alpha_3} = \tfrac{11}{47} \begin{pmatrix} \frac{47}{11} \\ \frac{39}{11} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{39}{47} \end{pmatrix}.$$

If we were to continue the process in Example 12, we would find that the sequence of vectors $\mathbf{e}_n$ converges to the eigenvector $(1 \quad 1)^T$, and the sequence α_n converges to the corresponding eigenvalue 5.

Exercise 35

Given

$$\mathbf{A} = \begin{pmatrix} 3 & 6 \\ 4 & 8 \end{pmatrix} \quad \text{and} \quad \mathbf{e}_0 = (1 \quad 0)^T,$$

use Procedure 4 to calculate $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$. What can you deduce about the eigenvalues and eigenvectors of $\mathbf{A}$?

Eigenvalue of smallest magnitude: inverse iteration

In order to approximate an eigenvector corresponding to the eigenvalue of *smallest* magnitude, we adapt the above method of direct iteration.

Consider an invertible matrix $\mathbf{A}$ whose eigenvalues are real and distinct in magnitude. Suppose again that the eigenvalues of $\mathbf{A}$ are listed in decreasing order of magnitude. It follows that their reciprocals must appear in increasing order of magnitude. But by the results of Subsection 2.4, the numbers λ^{-1} are the eigenvalues of the inverse matrix $\mathbf{A}^{-1}$ with the same eigenvectors as the matrix $\mathbf{A}$. So the problem of approximating the eigenvalue of smallest magnitude λ of $\mathbf{A}$ is the same as that of approximating the eigenvalue of largest magnitude λ^{-1} of $\mathbf{A}^{-1}$.

If $\mathbf{A}$ is invertible, then $\mathbf{A}^{-1}$ exists and the eigenvalues are non-zero (see Subsection 2.3).

It follows that repeatedly applying the matrix $\mathbf{A}^{-1}$ produces an eigenvector corresponding to the eigenvalue of smallest magnitude, assuming that it is real and distinct in magnitude from any other eigenvalue of $\mathbf{A}$.

For any square matrix $\mathbf{A}$ with non-zero real eigenvalues, we find an eigenvector corresponding to an eigenvalue of smallest magnitude by choosing a vector $\mathbf{e}_0$ and successively calculating the vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots$ using the formula

$$\mathbf{e}_{n+1} = \mathbf{A}^{-1}\mathbf{e}_n, \quad n = 0, 1, 2, \ldots.$$

In practice, such calculations can suffer from the same difficulties as direct iteration. We solve the problem of very large or very small vectors as before, by scaling, writing

$$\mathbf{z}_{n+1} = \mathbf{A}^{-1}\mathbf{e}_n, \quad \mathbf{e}_{n+1} = \frac{\mathbf{z}_{n+1}}{\alpha_{n+1}}, \quad n = 0, 1, 2, \ldots,$$

where α_{n+1} is the component of largest magnitude of $\mathbf{z}_{n+1}$. However, there is a further complication – the calculation of the inverse matrix can be very time consuming for large matrices. A more practical approach is based on solving the equations $\mathbf{A}\mathbf{z}_{n+1} = \mathbf{e}_n$ for $\mathbf{z}_{n+1}$ by Gaussian elimination and then putting $\mathbf{e}_{n+1} = \mathbf{z}_{n+1}/\alpha_{n+1}$.

Procedure 5 Inverse iteration

For any invertible (square) matrix $\mathbf{A}$ for which the eigenvalue $\lambda_{\min}$ of smallest magnitude is real and distinct in magnitude from any other eigenvalue, choose any vector $\mathbf{e}_0$.

The fact that $\mathbf{A}$ is invertible ensures that $\lambda_{\min} \neq 0$ and that $\mathbf{A}^{-1}$ exists.

(a) For $n = 0, 1, 2, \ldots$, carry out the following steps.

1. Calculate $\mathbf{z}_{n+1} = \mathbf{A}^{-1}\mathbf{e}_n$.
2. Find α_{n+1}, the component of largest magnitude of $\mathbf{z}_{n+1}$.
3. Put $\mathbf{e}_{n+1} = \mathbf{z}_{n+1}/\alpha_{n+1}$.

(b) The above method is inefficient for large matrices due to the difficulty of calculating $\mathbf{A}^{-1}$. In such cases, for $n = 0, 1, 2, \ldots$, carry out the following steps.

1. Calculate $\mathbf{z}_{n+1}$ by solving the equation $\mathbf{A}\mathbf{z}_{n+1} = \mathbf{e}_n$.
2. Find α_{n+1}, the component of largest magnitude of $\mathbf{z}_{n+1}$.
3. Put $\mathbf{e}_{n+1} = \mathbf{z}_{n+1}/\alpha_{n+1}$.

For sufficiently large n, $\mathbf{e}_n$ will be a good approximation to an eigenvector corresponding to the eigenvalue of smallest magnitude, provided that $\mathbf{e}_0$ has a non-zero component of the required eigenvector. If the sequence α_n converges, then it converges to $1/\lambda_{\min}$.

Example 13

Given the matrix $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 3 & 4 \end{pmatrix}$, calculate $\mathbf{A}^{-1}$. Given $\mathbf{e}_0 = (1 \quad 1)^T$, use Procedure 5(a) to calculate $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$.

Solution

We have

$$\mathbf{A}^{-1} = \tfrac{1}{6}\begin{pmatrix} 4 & -2 \\ -3 & 3 \end{pmatrix}.$$

First iteration:

$$\mathbf{z}_1 = \mathbf{A}^{-1}\mathbf{e}_0 = \tfrac{1}{6}\begin{pmatrix} 4 & -2 \\ -3 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ 0 \end{pmatrix},$$

$$\alpha_1 = \tfrac{1}{3},$$

$$\mathbf{e}_1 = \frac{\mathbf{z}_1}{\alpha_1} = 3\begin{pmatrix} \frac{1}{3} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Second iteration:

$$\mathbf{z}_2 = \mathbf{A}^{-1}\mathbf{e}_1 = \tfrac{1}{6}\begin{pmatrix} 4 & -2 \\ -3 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{2} \end{pmatrix},$$

$$\alpha_2 = \tfrac{2}{3},$$

$$\mathbf{e}_2 = \frac{\mathbf{z}_2}{\alpha_2} = \tfrac{3}{2}\begin{pmatrix} \frac{2}{3} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{3}{4} \end{pmatrix}.$$

Third iteration:

$$\mathbf{z}_3 = \mathbf{A}^{-1}\mathbf{e}_2 = \tfrac{1}{6}\begin{pmatrix} 4 & -2 \\ -3 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ -\frac{3}{4} \end{pmatrix} = \begin{pmatrix} \frac{11}{12} \\ -\frac{7}{8} \end{pmatrix},$$

$$\alpha_3 = \tfrac{11}{12},$$

$$\mathbf{e}_3 = \frac{\mathbf{z}_3}{\alpha_3} = \tfrac{12}{11}\begin{pmatrix} \frac{11}{12} \\ -\frac{7}{8} \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{21}{22} \end{pmatrix}.$$

If we were to continue the process in Example 13, we would find that the sequence of vectors $\mathbf{e}_n$ converges to the eigenvector $(1 \quad -1)^T$, and the sequence α_n converges to 1, corresponding to the eigenvalue $1/1 = 1$.

Exercise 36

Use Procedure 5(a) with $\mathbf{A} = \begin{pmatrix} 7 & 3 \\ 8 & 5 \end{pmatrix}$ and $\mathbf{e}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ to obtain $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$. Hence find an approximation to the eigenvalue of $\mathbf{A}$ of smallest magnitude, and a corresponding eigenvector.

Specific eigenvalues: modified inverse iteration

Procedures 4 and 5 are restricted to calculating the eigenvalue of largest or smallest magnitude. Sometimes we do not need either of these, but rather we need the eigenvalue closest to a given value. The following method will allow us to find it. We assume that there is just one eigenvalue of $\mathbf{A}$ closest to the given value p.

If λ is an eigenvalue of a matrix $\mathbf{A}$ (and p is not an eigenvalue of $\mathbf{A}$), then $(\lambda - p)^{-1}$ is an eigenvalue of the matrix $(\mathbf{A} - p\mathbf{I})^{-1}$, and the corresponding eigenvectors are unchanged. Suppose that λ_1 is the real eigenvalue closest to p. In other words, $|\lambda_1 - p|$ is the smallest of all possible choices of $|\lambda - p|$, thus $1/|\lambda_1 - p|$ is the largest of all possible choices of $1/|\lambda - p|$. It follows that repeatedly applying the matrix $(\mathbf{A} - p\mathbf{I})^{-1}$ to a chosen vector $\mathbf{e}_0$ produces an eigenvector corresponding to the eigenvalue closest to p. Thus the sequence $\mathbf{e}_{n+1} = (\mathbf{A} - p\mathbf{I})^{-1}\mathbf{e}_n$ $(n = 0, 1, 2, \ldots)$ should produce a sequence of approximations to the eigenvector.

See Subsection 2.4.

This method suffers from the deficiencies mentioned for inverse iteration, so we make similar refinements.

Procedure 6 Modified inverse iteration

Suppose that $\mathbf{A}$ is a square matrix for which one distinct real eigenvalue λ_1 is closest to a given real number p. To find an eigenvector corresponding to the eigenvalue closest to p, choose any vector $\mathbf{e}_0$.

The procedure breaks down if p is an eigenvalue.

(a) For $n = 0, 1, 2, \ldots$, carry out the following steps.

1. Calculate $\mathbf{z}_{n+1} = (\mathbf{A} - p\mathbf{I})^{-1}\mathbf{e}_n$.
2. Find α_{n+1}, the component of largest magnitude of $\mathbf{z}_{n+1}$.
3. Put $\mathbf{e}_{n+1} = \mathbf{z}_{n+1}/\alpha_{n+1}$.

(b) The above method is inefficient for large matrices due to the difficulty of calculating $(\mathbf{A} - p\mathbf{I})^{-1}$. In such cases, for $n = 0, 1, 2, \ldots$, carry out the following steps.

1. Calculate $\mathbf{z}_{n+1}$ by solving the equation $(\mathbf{A} - p\mathbf{I})\mathbf{z}_{n+1} = \mathbf{e}_n$.
2. Find α_{n+1}, the component of largest magnitude of $\mathbf{z}_{n+1}$.
3. Put $\mathbf{e}_{n+1} = \mathbf{z}_{n+1}/\alpha_{n+1}$.

For sufficiently large n, $\mathbf{e}_n$ will be a good approximation to an eigenvector corresponding to the eigenvalue closest to p, provided that $\mathbf{e}_0$ has a non-zero component of the required eigenvector. If the sequence α_n converges, then it converges to $1/(\lambda_1 - p)$.

In the following exercise, much of the work has been done for you.

Exercise 37

We wish to obtain an approximation to an eigenvector corresponding to the eigenvalue closest to $p = -1$ for the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}.$$

This inverse could most easily be found using a computer algebra package. Alternatively, Gaussian elimination, as described in Unit 4, could be used.

We have

$$(\mathbf{A} - p\mathbf{I})^{-1} = (\mathbf{A} + \mathbf{I})^{-1} = \begin{pmatrix} \frac{3}{2} & -\frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{4} & \frac{5}{8} & -\frac{1}{4} \\ -\frac{1}{2} & -\frac{1}{4} & \frac{1}{2} \end{pmatrix}.$$

Applying Procedure 6(a) with $\mathbf{e}_0 = (1 \quad 0 \quad 0)^T$ gives $\alpha_{20} \simeq 1.724\,96$ and $\mathbf{e}_{20} \simeq (1 \quad -0.141\,05 \quad -0.379\,39)^T$.

Calculate $\mathbf{e}_{21}$ with the components given to five decimal places so as to check convergence to three decimal places, and obtain an estimate for the corresponding eigenvalue to three decimal places.

Note that Procedure 5 is essentially the same as Procedure 4, but with $\mathbf{A}^{-1}$ in place of $\mathbf{A}$, and Procedure 6 is also essentially the same as Procedure 4, but with $(\mathbf{A} - p\mathbf{I})^{-1}$ in place of $\mathbf{A}$.

Procedures 4, 5 and 6 can be used to find individual eigenvalues and eigenvectors of a matrix. If we require *all* the eigenvalues and/or eigenvectors, then there are more efficient methods that can be used, though we do not discuss them here.

The rate of convergence of each of the methods depends on the relative closeness in magnitude of the other eigenvalues of $\mathbf{A}$ (or $\mathbf{A}^{-1}$, or $(\mathbf{A} - p\mathbf{I})^{-1}$) to the required eigenvalue. For example, in Exercise 37 the eigenvalues are 10.187, −0.420 and 0.234, to three decimal places. The direct iteration method applied to this problem would converge very rapidly since the largest eigenvalue, 10.187, is much larger in magnitude than the other two. On the other hand, inverse iteration would be slower, since the eigenvalues of $\mathbf{A}^{-1}$ are 4.281, −2.379 and 0.098 (to three decimal places), and the second largest eigenvalue in magnitude, −2.379, is just over half the magnitude of the largest eigenvalue.

A judicious choice of p for the modified inverse iteration method can significantly increase the rate of convergence. For example, choosing $p = 0.2$ for the matrix in Exercise 37 gives the eigenvalues of $\mathbf{A} - p\mathbf{I}$ as 9.987, −0.620 and 0.034, so the eigenvalues of $(\mathbf{A} - p\mathbf{I})^{-1}$ are 29.781, −1.612 and 0.100. We expect that the modified inverse iteration method with this value of p would converge very rapidly to the eigenvalue 0.234 of the matrix $\mathbf{A}$.

Procedures 4 and 5 cannot be used to determine complex eigenvalues, since complex eigenvalues of real matrices occur in complex conjugate pairs, and a complex eigenvalue and its complex conjugate have the same magnitude. Both procedures fail to find the required eigenvalue when there is a second distinct eigenvalue of the same magnitude.

However, it is possible to use Procedure 6 to find complex eigenvalues, provided that a complex value for p is chosen so that p is closer in magnitude to one of the complex eigenvalues than to any of the others.

Exercise 38

In this exercise, $\mathbf{A} = \begin{pmatrix} 0.5 & 0.6 \\ 1.4 & -0.3 \end{pmatrix}$.

These calculations are intended to be done by hand (with the aid of a scientific calculator).

(a) Given $\mathbf{e}_0 = (1 \quad 0)^T$, use direct iteration to calculate $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$.

(b) Use the methods of Section 2 to find the eigenvalues and corresponding eigenvectors of $\mathbf{A}$.

(c) To which eigenvector would you expect the sequence $\mathbf{e}_n$ of part (a) to converge?

(d) If $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ, express $\mathbf{A}^8\mathbf{v}$ in terms of λ and $\mathbf{v}$.

(e) Designate the eigenvectors found in part (b) as $\mathbf{v}_1$ and $\mathbf{v}_2$. Express $\mathbf{e}_0$ in terms of $\mathbf{v}_1$ and $\mathbf{v}_2$, then calculate $\mathbf{e}_8$. (If this seems like hard work, then look for an easier method.)

(f) Find $\mathbf{A}^{-1}$, then use inverse iteration to calculate $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$, given $\mathbf{e}_0 = (-0.4 \quad 1)^T$. To which eigenvector would you expect this sequence to converge?

(g) Comment on the rates of convergence for direct iteration and inverse iteration applied to this problem.

Exercise 39

Suppose that you wish to find all the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}.$$

Before you start this exercise, consider how the third eigenvalue can be obtained from the other two eigenvalues. Applying this method will reduce the work.

Further, suppose that you have used direct iteration to find an eigenvector

$$\mathbf{v}_1 \simeq (0.477 \quad 0.689 \quad 1)^T,$$

and inverse iteration to find another eigenvector

$$\mathbf{v}_2 \simeq (-0.102 \quad 1 \quad -0.641)^T.$$

Use this information to find approximations to all three eigenvalues.

Learning outcomes

After studying this unit, you should be able to:

- explain the meaning of the terms eigenvalue, eigenvector and characteristic equation
- calculate the eigenvalues of a given 2×2 matrix, and find the corresponding eigenvectors
- calculate the eigenvalues and corresponding eigenvectors of a 3×3 matrix, where one of the eigenvalues is 'obvious'
- appreciate that an $n \times n$ matrix with n distinct eigenvalues gives rise to n linearly independent eigenvectors
- appreciate that the eigenvalues of a matrix may be real and distinct, real and repeated, or complex
- recall that the sum of the eigenvalues of an $n \times n$ matrix $\mathbf{A}$ is $\operatorname{tr} \mathbf{A}$, and that the product of the eigenvalues of $\mathbf{A}$ is $\det \mathbf{A}$, and use these properties as a check in hand calculations
- write down the eigenvalues of a triangular matrix
- write down the eigenvalues of the matrices $\mathbf{A}^k$, $\mathbf{A}^{-1}$, $\mathbf{A} + p\mathbf{I}$, $(\mathbf{A} - p\mathbf{I})^{-1}$ and $p\mathbf{A}$, given the eigenvalues of $\mathbf{A}$;
- appreciate that the eigenvalues of a real symmetric matrix are real and the corresponding eigenvectors are orthogonal, and in the case of repeated eigenvalues, the corresponding eigenvectors can be chosen to be orthogonal
- appreciate the use of direct, inverse and modified inverse iteration in approximating individual eigenvalues and corresponding eigenvectors of a square matrix
- use iterative methods and hand calculation to estimate an eigenvalue and corresponding eigenvector in simple cases.

Solutions to exercises

Solution to Exercise 1

$$\mathbf{Aw} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 7 \\ -1 \end{pmatrix},$$

$$\mathbf{Ax} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix},$$

$$\mathbf{Ay} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \end{pmatrix},$$

$$\mathbf{Az} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solution to Exercise 2

If $(x \quad y)^T = \alpha(1 \quad 1)^T + \beta(-2 \quad 1)^T$, then

$$\begin{aligned} \alpha - 2\beta &= x, \\ \alpha + \ \ \beta &= y. \end{aligned}$$

Solving these equations, we obtain $\alpha = (2y + x)/3$ and $\beta = (y - x)/3$.

Solution to Exercise 3

(a) $$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 8 \end{pmatrix} = 4 \begin{pmatrix} 3 \\ 2 \end{pmatrix},$$

so $(3 \quad 2)^T$ is an eigenvector with eigenvalue 4.

(b) $$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = (-1) \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

so $(1 \quad -1)^T$ is an eigenvector with eigenvalue -1.

(c) $$\begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 12 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 6 \end{pmatrix},$$

so $(0 \quad 6)^T$ is an eigenvector with eigenvalue 2.

(d) $$\begin{pmatrix} 12 & 3 \\ -8 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ -4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ -4 \end{pmatrix},$$

so $(1 \quad -4)^T$ is an eigenvector with eigenvalue 0.

Solution to Exercise 4

Since $\mathbf{v} = (12\,000 \quad 6000)^T$ is transformed to itself, this is an eigenvector with eigenvalue 1.

You may have noticed that there are many other eigenvectors with the same eigenvalue, for example $(12 \quad 6)^T$ and $(2 \quad 1)^T$.

There is another eigenvector $(1 \quad -1)^T$, with corresponding eigenvalue 0.7 (although we do not expect you to have found it).

Solution to Exercise 5

$$\begin{pmatrix} 1 & 4 \\ 9 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ -3 \end{pmatrix} = \begin{pmatrix} -10 \\ 15 \end{pmatrix} = -5\begin{pmatrix} 2 \\ -3 \end{pmatrix},$$

so $(2 \quad -3)^T$ is an eigenvector with eigenvalue -5. Also,

$$\begin{pmatrix} 1 & 4 \\ 9 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 14 \\ 21 \end{pmatrix} = 7\begin{pmatrix} 2 \\ 3 \end{pmatrix},$$

so $(2 \quad 3)^T$ is an eigenvector with eigenvalue 7.

Solution to Exercise 6

The eigenvectors act along the line of reflection $y = x$ and perpendicular to it, so they are the scalar multiples of $(1 \quad 1)^T$ and $(1 \quad -1)^T$. The vector $(1 \quad 1)^T$ is scaled by a factor of 1 by the transformation, while for $(1 \quad -1)^T$ the scale factor is -1; these scale factors are the corresponding eigenvalues.

We may check our conclusion by evaluating

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

so $(1 \quad 1)^T$ corresponds to the eigenvalue 1, and

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -1\begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

so $(1 \quad -1)^T$ corresponds to the eigenvalue -1.

Solution to Exercise 7

$(0 \quad 1)^T$ is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue 1, and $(1 \quad 0)^T$ is an eigenvalue of $\mathbf{A}$ corresponding to the eigenvalue -1.

Solution to Exercise 8

The equation $\mathbf{Av} = \lambda\mathbf{v}$ becomes

$$\begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \end{pmatrix}.$$

Thus x and y satisfy the simultaneous equations

$$\begin{aligned} 5x + 2y &= \lambda x, \\ 2x + 5y &= \lambda y, \end{aligned}$$

which can be rewritten as the eigenvector equations

$$\begin{aligned} (5-\lambda)x + \qquad 2y &= 0, \\ 2x + (5-\lambda)y &= 0. \end{aligned}$$

These equations have a non-zero solution only if

$$\begin{vmatrix} 5-\lambda & 2 \\ 2 & 5-\lambda \end{vmatrix} = 0.$$

So $(5-\lambda)(5-\lambda)-4=0$, that is, $5-\lambda=\pm 2$, so the eigenvalues are $\lambda=7$ and $\lambda=3$.

- For $\lambda=7$, the eigenvector equations become

  $$\begin{aligned} -2x+2y&=0,\\ 2x-2y&=0. \end{aligned}$$

  These equations reduce to the single equation $y=x$, so an eigenvector corresponding to $\lambda=7$ is $(1 \quad 1)^T$.

- For $\lambda=3$, the eigenvector equations become

  $$\begin{aligned} 2x+2y&=0,\\ 2x+2y&=0. \end{aligned}$$

  These equations reduce to the single equation $y=-x$, so an eigenvector corresponding to $\lambda=3$ is $(1 \quad -1)^T$.

Solution to Exercise 9

(a) The characteristic equation is

$$\begin{vmatrix} 1-\lambda & 4 \\ 1 & -2-\lambda \end{vmatrix} = 0.$$

Expanding gives $(1-\lambda)(-2-\lambda)-4=0$, which simplifies to $\lambda^2+\lambda-6=(\lambda-2)(\lambda+3)=0$. (Alternatively, you could have calculated $a+d=-1$ and $ad-bc=-6$, and so obtained this equation directly.) So the eigenvalues are $\lambda=2$ and $\lambda=-3$.

The eigenvector equations are

$$\begin{aligned} (1-\lambda)x+\qquad\quad 4y&=0,\\ x+(-2-\lambda)y&=0. \end{aligned}$$

- For $\lambda=2$, the eigenvector equations become

  $$-x+4y=0 \quad\text{and}\quad x-4y=0,$$

  which reduce to the single equation $4y=x$. So an eigenvector corresponding to $\lambda=2$ is $(4 \quad 1)^T$.

- For $\lambda=-3$, the eigenvector equations become

  $$4x+4y=0 \quad\text{and}\quad x+y=0,$$

  which reduce to the single equation $y=-x$. So an eigenvector corresponding to $\lambda=-3$ is $(1 \quad -1)^T$.

(b) The characteristic equation is

$$\begin{vmatrix} 8-\lambda & -5 \\ 10 & -7-\lambda \end{vmatrix} = 0.$$

Expanding gives $(8-\lambda)(-7-\lambda)+50=0$, which simplifies to $\lambda^2-\lambda-6=(\lambda-3)(\lambda+2)=0$. (Alternatively, you could have calculated $a+d=1$ and $ad-bc=-6$, and so obtained this equation directly.) So the eigenvalues are $\lambda=3$ and $\lambda=-2$.

The eigenvector equations are

$$\begin{aligned}(8-\lambda)x - \quad\quad 5y &= 0,\\ 10x - (7+\lambda)y &= 0.\end{aligned}$$

- For $\lambda = 3$, the eigenvector equations become

  $$5x - 5y = 0 \quad \text{and} \quad 10x - 10y = 0,$$

  which reduce to the single equation $y = x$. So an eigenvector corresponding to $\lambda = 3$ is $(1 \quad 1)^T$.
- For $\lambda = -2$, the eigenvector equations become

  $$10x - 5y = 0 \quad \text{and} \quad 10x - 5y = 0,$$

  which reduce to the single equation $y = 2x$. So an eigenvector corresponding to $\lambda = -2$ is $(1 \quad 2)^T$.

(c) The characteristic equation is

$$\begin{vmatrix}-1-\lambda & 0\\ 9 & 2-\lambda\end{vmatrix} = 0.$$

Expanding gives $(-1-\lambda)(2-\lambda) = 0$, so the eigenvalues are $\lambda = -1$ and $\lambda = 2$.

The eigenvector equations are

$$\begin{aligned}(-1-\lambda)x + \quad\quad 0y &= 0,\\ 9x + (2-\lambda)y &= 0.\end{aligned}$$

- For $\lambda = -1$, the eigenvector equations become

  $$0x + 0y = 0 \quad \text{and} \quad 9x + 3y = 0,$$

  which reduce to the single equation $y = -3x$. So an eigenvector corresponding to $\lambda = -1$ is $(1 \quad -3)^T$.
- For $\lambda = 2$, the eigenvector equations become

  $$-3x + 0y = 0 \quad \text{and} \quad 9x + 0y = 0,$$

  which reduce to the single equation $x = 0$. So an eigenvector corresponding to $\lambda = 2$ is $(0 \quad 1)^T$.

Solution to Exercise 10

(a) Sum of eigenvalues $= 5 + 2 = 7$, and $\operatorname{tr} \mathbf{A} = 3 + 4 = 7$.

Product of eigenvalues $= 5 \times 2 = 10$, and $\det \mathbf{A} = (3 \times 4) - (2 \times 1) = 10$.

(b) Sum of eigenvalues $= 7 + 3 = 10$, and $\operatorname{tr} \mathbf{A} = 5 + 5 = 10$.

Product of eigenvalues $= 7 \times 3 = 21$, and $\det \mathbf{A} = (5 \times 5) - (2 \times 2) = 21$.

(c) For part (a):

Sum of eigenvalues $= 2 + (-3) = -1$, and $\operatorname{tr} \mathbf{A} = 1 + (-2) = -1$.

Product of eigenvalues $= 2 \times (-3) = -6$, and $\det \mathbf{A} = (1 \times (-2)) - (4 \times 1) = -6$.

For part (b):

Sum of eigenvalues $= 3 + (-2) = 1$, and $\operatorname{tr} \mathbf{A} = 8 + (-7) = 1$.

Product of eigenvalues $= 3 \times (-2) = -6$, and
$\det \mathbf{A} = (8 \times (-7)) - ((-5) \times 10) = -6$.

For part (c):

Sum of eigenvalues $= -1 + 2 = 1$, and $\operatorname{tr} \mathbf{A} = -1 + 2 = 1$.

Product of eigenvalues $= (-1) \times 2 = -2$, and
$\det \mathbf{A} = ((-1) \times 2) - (0 \times 9) = -2$.

Solution to Exercise 11

The characteristic equation is

$$\begin{vmatrix} p-\lambda & 0 \\ 0 & q-\lambda \end{vmatrix} = (p-\lambda)(q-\lambda) = 0.$$

Thus the eigenvalues are $\lambda = p$ and $\lambda = q$.

The eigenvector equations are

$$\begin{aligned} (p-\lambda)x \qquad\qquad &= 0, \\ (q-\lambda)y &= 0. \end{aligned}$$

- For $\lambda = p$, the eigenvector equations become

  $$0 = 0 \quad \text{and} \quad (q-p)y = 0,$$

  which reduce to the single equation $y = 0$ (since $p \neq q$), so a corresponding eigenvector is $(1 \quad 0)^T$.
- For $\lambda = q$, the eigenvector equations become

  $$(p-q)x = 0 \quad \text{and} \quad 0 = 0,$$

  which reduce to the single equation $x = 0$ (since $p \neq q$), so a corresponding eigenvector is $(0 \quad 1)^T$.

These results agree with the eigenvalues and eigenvectors in Figure 5.

Solution to Exercise 12

The characteristic equation is

$$\lambda^2 - 1.7\lambda + 0.7 = 0.$$

The eigenvalues are $\lambda = 1$ and $\lambda = 0.7$.

The eigenvector equations are

$$\begin{aligned} (0.9-\lambda)x + \qquad 0.2y &= 0, \\ 0.1x + (0.8-\lambda)y &= 0. \end{aligned}$$

- For $\lambda = 1$, the eigenvector equations become

  $$-0.1x + 0.2y = 0 \quad \text{and} \quad 0.1x - 0.2y = 0,$$

  which reduce to the single equation $2y = x$, so a corresponding eigenvector is $(2 \quad 1)^T$.

(In the migration problem, where the total population was 18 000, an eigenvector corresponding to $\lambda = 1$ was found to be $(12\,000 \quad 6000)^T$. This is a multiple of $(2 \quad 1)^T$, as expected, giving stable populations of 12 000 in Exton and 6000 in Wyeville.)

- For $\lambda = 0.7$, the eigenvector equations become

  $$0.2x + 0.2y = 0 \quad \text{and} \quad 0.1x + 0.1y = 0,$$

  which reduce to the single equation $y = -x$, so a corresponding eigenvector is $(1 \quad -1)^T$.

  (Since populations cannot be negative, this solution has no relevance for the migration problem.)

Solution to Exercise 13

(a) The characteristic equation is

$$\begin{vmatrix} a-\lambda & 0 \\ 0 & a-\lambda \end{vmatrix} = (a-\lambda)^2 = 0,$$

for which $\lambda = a$ is a repeated root.

The eigenvector equations are

$$\begin{aligned} (a-\lambda)x \qquad\qquad &= 0, \\ (a-\lambda)y &= 0, \end{aligned}$$

which for $\lambda = a$ become

$$0 = 0 \quad \text{and} \quad 0 = 0,$$

which are satisfied by all values of x and y, so the eigenvectors are all the non-zero vectors of the form $(k \quad l)^T$.

(Any non-zero vector is an eigenvector, but it is possible to choose two eigenvectors that are linearly independent, for example $(1 \quad 0)^T$ and $(0 \quad 1)^T$.)

(b) The characteristic equation is

$$\begin{vmatrix} a-\lambda & 1 \\ 0 & a-\lambda \end{vmatrix} = (a-\lambda)^2 = 0,$$

for which $\lambda = a$ is a repeated root.

The eigenvector equations are

$$\begin{aligned} (a-\lambda)x + \qquad\quad y &= 0, \\ (a-\lambda)y &= 0, \end{aligned}$$

which for $\lambda = a$ reduce to a single equation $y = 0$, and a corresponding eigenvector is $(1 \quad 0)^T$.

(In this case, there is only one linearly independent eigenvector.)

Solution to Exercise 14

The characteristic equation is $\lambda^2 - 4\lambda + 5 = 0$. The eigenvalues are

$$\lambda = \tfrac{1}{2}\left(4 \pm \sqrt{16-20}\right) = \tfrac{1}{2}(4 \pm 2i) = 2 \pm i,$$

that is, $\lambda = 2 + i$ and $\lambda = 2 - i$.

The eigenvector equations are

$$\begin{aligned}(3-\lambda)x - y &= 0,\\ 2x + (1-\lambda)y &= 0.\end{aligned}$$

- For $\lambda = 2 + i$, the eigenvector equations become

  $$(1-i)x - y = 0 \quad \text{and} \quad 2x - (1+i)y = 0,$$

  which reduce to the single equation $y = (1-i)x$ (as can be seen by multiplying the second equation by $(1-i)/2$ and noting that $(1-i)(1+i) = 2$), so a corresponding eigenvector is $(1 \quad 1-i)^T$.
- For $\lambda = 2 - i$, the eigenvector equations become

  $$(1+i)x - y = 0 \quad \text{and} \quad 2x - (1-i)y = 0,$$

  which reduce to the single equation $y = (1+i)x$ (as can be seen by multiplying the second equation by $(1+i)/2$ and noting that $(1+i)(1-i) = 2$), so a corresponding eigenvector is $(1 \quad 1+i)^T$.

Solution to Exercise 15

The eigenvalues are 1 and 2.

- For $\lambda = 1$, the eigenvector equations become

  $$3y = 0 \quad \text{and} \quad y = 0,$$

  which reduce to the single equation $y = 0$, so $(1 \quad 0)^T$ is a corresponding eigenvector.
- For $\lambda = 2$, the eigenvector equations become

  $$-x + 3y = 0 \quad \text{and} \quad 0 = 0,$$

  which reduce to the single equation $3y = x$, so $(3 \quad 1)^T$ is a corresponding eigenvector.

Solution to Exercise 16

For the eigenvalue to be repeated, we require $\sqrt{(a+d)^2 - 4(ad-b^2)} = 0$, that is, $(a-d)^2 + 4b^2 = 0$. This is true only if $a = d$ and $b = 0$, so the only symmetric 2×2 matrices with a repeated eigenvalue are of the form $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$.

Solution to Exercise 17

(a) The eigenvalues are real, since $\mathbf{A}$ is real and symmetric. One is positive and the other negative, since $\lambda_1\lambda_2 = \det \mathbf{A} < 0$. Also, $\lambda_1 + \lambda_2 = \operatorname{tr} \mathbf{A} = 50$.

(b) The eigenvalues are the diagonal entries 67 and -17, since $\mathbf{A}$ is triangular.

(c) The eigenvalues are real, since $\mathbf{A}$ is real and symmetric. In fact, $\mathbf{A}$ is non-invertible, since $\det \mathbf{A} = 0$. Thus one eigenvalue is 0. Hence the other is 306, since $0 + \lambda_2 = \operatorname{tr} \mathbf{A} = 306$.

Solution to Exercise 18

(a) (i) $$\mathbf{A}^2 = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 11 & 14 \\ 7 & 18 \end{pmatrix}.$$

The characteristic equation of $\mathbf{A}^2$ is

$$\lambda^2 - 29\lambda + 100 = 0.$$

So the eigenvalues of $\mathbf{A}^2$ are $\lambda = 25$ and $\lambda = 4$. These are the squares of the eigenvalues of $\mathbf{A}$.

(ii) $$\mathbf{A}^{-1} = \tfrac{1}{10}\begin{pmatrix} 4 & -2 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 0.4 & -0.2 \\ -0.1 & 0.3 \end{pmatrix}.$$

The characteristic equation of $\mathbf{A}^{-1}$ is

$$\lambda^2 - 0.7\lambda + 0.1 = 0.$$

So the eigenvalues of $\mathbf{A}^{-1}$ are $\lambda = 0.5$ and $\lambda = 0.2$. These are the reciprocals of the eigenvalues of $\mathbf{A}$.

(iii) $$\mathbf{A} + 2\mathbf{I} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 1 & 6 \end{pmatrix}.$$

The characteristic equation of $\mathbf{A} + 2\mathbf{I}$ is

$$\lambda^2 - 11\lambda + 28 = 0.$$

So the eigenvalues of $\mathbf{A} + 2\mathbf{I}$ are $\lambda = 7$ and $\lambda = 4$. These can be obtained by adding 2 to the eigenvalues of $\mathbf{A}$.

(iv) $$(\mathbf{A} - 4\mathbf{I})^{-1} = \begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix}^{-1} = \frac{1}{(-2)}\begin{pmatrix} 0 & -2 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0.5 & 0.5 \end{pmatrix}.$$

The characteristic equation of $(\mathbf{A} - 4\mathbf{I})^{-1}$ is

$$\lambda^2 - 0.5\lambda - 0.5 = 0.$$

So the eigenvalues of $(\mathbf{A} - 4\mathbf{I})^{-1}$ are $\lambda = 1$ and $\lambda = -0.5$. These can be obtained by subtracting 4 from the eigenvalues of $\mathbf{A}$ and then finding the reciprocals.

(v) $$3\mathbf{A} = \begin{pmatrix} 9 & 6 \\ 3 & 12 \end{pmatrix}.$$

The characteristic equation of $3\mathbf{A}$ is

$$\lambda^2 - 21\lambda + 90 = 0.$$

So the eigenvalues of $3\mathbf{A}$ are $\lambda = 15$ and $\lambda = 6$, which are three times those of $\mathbf{A}$.

(b) (i) $$\mathbf{A}^2\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}11&14\\7&18\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}25\\25\end{pmatrix}=25\begin{pmatrix}1\\1\end{pmatrix},$$
$$\mathbf{A}^2\begin{pmatrix}-2\\1\end{pmatrix}=\begin{pmatrix}11&14\\7&18\end{pmatrix}\begin{pmatrix}-2\\1\end{pmatrix}=\begin{pmatrix}-8\\4\end{pmatrix}=4\begin{pmatrix}-2\\1\end{pmatrix},$$

so the eigenvectors of $\mathbf{A}$ are also eigenvectors of $\mathbf{A}^2$.

(ii) $$\mathbf{A}^{-1}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}0.4&-0.2\\-0.1&0.3\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}0.2\\0.2\end{pmatrix}=0.2\begin{pmatrix}1\\1\end{pmatrix},$$
$$\mathbf{A}^{-1}\begin{pmatrix}-2\\1\end{pmatrix}=\begin{pmatrix}0.4&-0.2\\-0.1&0.3\end{pmatrix}\begin{pmatrix}-2\\1\end{pmatrix}=\begin{pmatrix}-1\\0.5\end{pmatrix}=0.5\begin{pmatrix}-2\\1\end{pmatrix},$$

so the eigenvectors of $\mathbf{A}$ are also eigenvectors of $\mathbf{A}^{-1}$.

(iii) $$(\mathbf{A}+2\mathbf{I})\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}5&2\\1&6\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}7\\7\end{pmatrix}=7\begin{pmatrix}1\\1\end{pmatrix},$$
$$(\mathbf{A}+2\mathbf{I})\begin{pmatrix}-2\\1\end{pmatrix}=\begin{pmatrix}5&2\\1&6\end{pmatrix}\begin{pmatrix}-2\\1\end{pmatrix}=\begin{pmatrix}-8\\4\end{pmatrix}=4\begin{pmatrix}-2\\1\end{pmatrix},$$

so the eigenvectors of $\mathbf{A}$ are also eigenvectors of $\mathbf{A}+2\mathbf{I}$.

(iv) $$(\mathbf{A}-4\mathbf{I})^{-1}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}0&1\\0.5&0.5\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix},$$
$$\begin{aligned}(\mathbf{A}-4\mathbf{I})^{-1}\begin{pmatrix}-2\\1\end{pmatrix}&=\begin{pmatrix}0&1\\0.5&0.5\end{pmatrix}\begin{pmatrix}-2\\1\end{pmatrix}\\&=\begin{pmatrix}1\\-0.5\end{pmatrix}\\&=-0.5\begin{pmatrix}-2\\1\end{pmatrix},\end{aligned}$$

so the eigenvectors of $\mathbf{A}$ are also eigenvectors of $(\mathbf{A}-4\mathbf{I})^{-1}$.

(v) $$3\mathbf{A}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}9&6\\3&12\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}15\\15\end{pmatrix}=15\begin{pmatrix}1\\1\end{pmatrix},$$
$$3\mathbf{A}\begin{pmatrix}-2\\1\end{pmatrix}=\begin{pmatrix}9&6\\3&12\end{pmatrix}\begin{pmatrix}-2\\1\end{pmatrix}=\begin{pmatrix}-12\\6\end{pmatrix}=6\begin{pmatrix}-2\\1\end{pmatrix},$$

so the eigenvectors of $\mathbf{A}$ are also eigenvectors of $3\mathbf{A}$.

Solution to Exercise 19

(a) (i) The eigenvalues are 4^3 and $(-1)^3$, i.e. 64 and -1.

(ii) The eigenvalues are 4^{-1} and $(-1)^{-1}$, i.e. $\frac{1}{4}$ and -1.

(iii) The eigenvalues are $4-6$ and $(-1)-6$, i.e. -2 and -7.

(iv) The eigenvalues are $(4+3)^{-1}$ and $((-1)+3)^{-1}$, i.e. $\frac{1}{7}$ and $\frac{1}{2}$.

(b) The eigenvalues of $\mathbf{A}-4\mathbf{I}$ are $4-4=0$ and $-1-4=-5$.

The matrix $\mathbf{A}-4\mathbf{I}$ is non-invertible because one of the eigenvalues is 0, so the inverse does not exist.

Solution to Exercise 20

(a) Using Procedure 1, we solve the characteristic equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, which can be written as

$$\begin{vmatrix} 1-\lambda & 2 \\ 3 & -4-\lambda \end{vmatrix} = \lambda^2 + 3\lambda - 10 = 0.$$

The eigenvalues are $\lambda = -5$ and $\lambda = 2$. Solving the eigenvector equations for each eigenvalue, we obtain corresponding eigenvectors $(1 \quad -3)^T$ and $(2 \quad 1)^T$, respectively.

The eigenvalues of $\mathbf{A}^{10}$ are $(-5)^{10}$ and 2^{10}, corresponding to eigenvectors $(1 \quad -3)^T$ and $(2 \quad 1)^T$, respectively.

(b) $\begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix} = \mathbf{A} + 2\mathbf{I}$, where $\mathbf{A}$ is the matrix of part (a).

So the eigenvalues are $-5+2=-3$ and $2+2=4$, with corresponding eigenvectors $(1 \quad -3)^T$ and $(2 \quad 1)^T$, respectively.

Solution to Exercise 21

(a) The characteristic equation is

$$\lambda^2 - (2\cos\theta)\lambda + 1 = 0,$$

since $\sin^2\theta + \cos^2\theta = 1$. Using the quadratic equation formula, the eigenvalues are

$$\begin{aligned} \lambda &= \tfrac{1}{2}\left(2\cos\theta \pm \sqrt{4\cos^2\theta - 4}\right) \\ &= \tfrac{1}{2}\left(2\cos\theta \pm \sqrt{-4\sin^2\theta}\right) = \cos\theta \pm i\sin\theta. \end{aligned}$$

So the eigenvalues are $\cos\theta + i\sin\theta$ and $\cos\theta - i\sin\theta$.

The eigenvector equations are

$$\begin{aligned} (\cos\theta - \lambda)x - \quad (\sin\theta)y &= 0, \\ (\sin\theta)x + (\cos\theta - \lambda)y &= 0. \end{aligned}$$

- For $\lambda = \cos\theta + i\sin\theta$, the eigenvector equations become $-(i\sin\theta)x - (\sin\theta)y = 0$ and $(\sin\theta)x - (i\sin\theta)y = 0$, which reduce to the single equation $iy = x$ (since $\sin\theta \neq 0$ as θ is not an integer multiple of π), so a corresponding eigenvector is $(i \quad 1)^T$.
- For $\lambda = \cos\theta - i\sin\theta$, the eigenvector equations become $(i\sin\theta)x - (\sin\theta)y = 0$ and $(\sin\theta)x + (i\sin\theta)y = 0$, which reduce to the single equation $-iy = x$ (since $\sin\theta \neq 0$), so a corresponding eigenvector is $(-i \quad 1)^T$.

(b) If $\theta = n\pi$, where n is an integer, then $\sin\theta = 0$ and $\cos\theta = (-1)^n$, so $\mathbf{A}$ is the diagonal matrix

$$\mathbf{A} = \begin{pmatrix} (-1)^n & 0 \\ 0 & (-1)^n \end{pmatrix},$$

with equal diagonal elements. Thus $\mathbf{A}$ has a repeated eigenvalue given by $(-1)^n$. As was found in Exercise 13(a), any non-zero vector is an eigenvector and we are free to select any two linearly independent vectors as eigenvectors.

Solution to Exercise 22

$$\begin{pmatrix} 5 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

and the corresponding eigenvalue is 3.

$$\begin{pmatrix} 5 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},$$

and the corresponding eigenvalue is 1.

Solution to Exercise 23

The eigenvector equations are

$$\begin{aligned} (1-\lambda)x_1 \qquad\qquad\quad - \qquad\quad x_3 &= 0, \\ x_1 + (2-\lambda)x_2 + \qquad\quad x_3 &= 0, \\ 2x_1 + \qquad 2x_2 + (3-\lambda)x_3 &= 0. \end{aligned}$$

- For $\lambda = 1$, the eigenvector equations become

  $$-x_3 = 0, \quad x_1 + x_2 + x_3 = 0, \quad 2x_1 + 2x_2 + 2x_3 = 0,$$

  which reduce to the equations $x_3 = 0$ and $x_2 = -x_1$, so a corresponding eigenvector is $(1 \quad -1 \quad 0)^T$.

- For $\lambda = 2$, the eigenvector equations become

  $$-x_1 - x_3 = 0, \quad x_1 + x_3 = 0, \quad 2x_1 + 2x_2 + x_3 = 0,$$

  which reduce to the equations $-x_3 = x_1$ and $-2x_2 = x_1$, so a corresponding eigenvector is $(-2 \quad 1 \quad 2)^T$.

- For $\lambda = 3$, the eigenvector equations become

  $$-2x_1 - x_3 = 0, \quad x_1 - x_2 + x_3 = 0, \quad 2x_1 + 2x_2 = 0,$$

  which reduce to the equations $x_3 = -2x_1$ and $x_2 = -x_1$, so a corresponding eigenvector is $(1 \quad -1 \quad -2)^T$.

Solution to Exercise 24

The characteristic equation is

$$\begin{vmatrix} -\lambda & 0 & 6 \\ \frac{1}{2} & -\lambda & 0 \\ 0 & \frac{1}{3} & -\lambda \end{vmatrix} = 0.$$

Expanding the determinant by the top row gives

$$-\lambda \begin{vmatrix} -\lambda & 0 \\ \frac{1}{3} & -\lambda \end{vmatrix} + 6 \begin{vmatrix} \frac{1}{2} & -\lambda \\ 0 & \frac{1}{3} \end{vmatrix} = 0,$$

which simplifies to $\lambda^3 - 1 = 0$. Since $\lambda = 1$ satisfies this equation, it is an eigenvalue of $\mathbf{A}$. (The other two eigenvalues are complex numbers.)

For $\lambda = 1$, the eigenvector equations become

$$-x_1 + 6x_3 = 0, \quad \tfrac{1}{2}x_1 - x_2 = 0, \quad \tfrac{1}{3}x_2 - x_3 = 0,$$

which reduce to the equations $x_1 = 6x_3$ and $x_2 = 3x_3$, so an eigenvector is $(6 \quad 3 \quad 1)^T$.

Solution to Exercise 25

(a) The characteristic equation is

$$\begin{vmatrix} -\lambda & 0 & 1 \\ 0 & -\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix} = 0.$$

Expanding the determinant by the middle row (where there are two zeros) gives

$$-\lambda \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = 0,$$

which simplifies to $\lambda(\lambda^2 - 1) = 0$, so the eigenvalues are $\lambda = 0$, $\lambda = -1$ and $\lambda = 1$.

(b) The matrix is triangular, so from Subsection 2.3, the eigenvalues are $\lambda = 1$, $\lambda = 2$ and $\lambda = 3$.

Solution to Exercise 26

The characteristic equation is

$$\begin{vmatrix} 8-\lambda & 0 & -5 \\ 9 & 3-\lambda & -6 \\ 10 & 0 & -7-\lambda \end{vmatrix} = 0.$$

Expanding by the middle column gives

$$(3-\lambda) \begin{vmatrix} 8-\lambda & -5 \\ 10 & -7-\lambda \end{vmatrix} = 0,$$

so $\lambda = 3$ or $(8-\lambda)(-7-\lambda) + 50 = 0$. This quadratic equation simplifies to $\lambda^2 - \lambda - 6 = 0$, which has roots $\lambda = 3$ and $\lambda = -2$.

Thus the eigenvalues are $\lambda = 3$ (repeated) and $\lambda = -2$.

Solution to Exercise 27

The characteristic equation of the first matrix is

$$\begin{vmatrix} a-\lambda & 0 & 0 \\ d & b-\lambda & 0 \\ e & f & c-\lambda \end{vmatrix} = 0.$$

Expanding by the top row gives

$$(a-\lambda)\begin{vmatrix} b-\lambda & 0 \\ f & c-\lambda \end{vmatrix} = 0,$$

so $(a-\lambda)(b-\lambda)(c-\lambda) = 0$.

Thus the eigenvalues are $\lambda = a$, $\lambda = b$ and $\lambda = c$.

The second matrix is the transpose of the first, so it has the same eigenvalues.

Solution to Exercise 28

(a) Sum of eigenvalues $= 5+3+1 = 9$, and $\operatorname{tr}\mathbf{A} = 5+2+2 = 9$.

Also, $\det\mathbf{A} = 5\times(4-1) = 15$, and product of eigenvalues $= 5\times 3\times 1 = 15$.

(b) Sum of eigenvalues $= 6+7+3 = 16$, and $\operatorname{tr}\mathbf{A} = 5+5+6 = 16$.

Also, $\det\mathbf{A} = 6\times(25-4) = 126$, and product of eigenvalues $= 6\times 7\times 3 = 126$.

(c) Sum of eigenvalues $= 1+2+3 = 6$, and $\operatorname{tr}\mathbf{A} = 1+2+3 = 6$.

Also, $\det\mathbf{A} = 1\times(6-2)+(-1)\times(2-4) = 6$, and product of eigenvalues $= 1\times 2\times 3 = 6$.

(d) For part (a):

Sum of eigenvalues $= 0+(-1)+1 = 0$, and $\operatorname{tr}\mathbf{A} = 0+0+0 = 0$.

Also, $\det\mathbf{A} = 1\times(0-0) = 0$, and product of eigenvalues $= 0\times(-1)\times 1 = 0$.

For part (b):

Sum of eigenvalues $= 1+2+3 = 6$, and $\operatorname{tr}\mathbf{A} = 1+2+3 = 6$.

Also, $\det\mathbf{A} = 1\times(6-0) = 6$, and product of eigenvalues $= 1\times 2\times 3 = 6$.

(e) Sum of eigenvalues $= 3+3+(-2) = 4$, and $\operatorname{tr}\mathbf{A} = 8+3+(-7) = 4$.

Also, $\det\mathbf{A} = 8\times(-21-0)-5\times(0-30) = -18$, and product of eigenvalues $= 3\times 3\times(-2) = -18$.

Solution to Exercise 29

(a) The characteristic equation is

$$\begin{vmatrix} -\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix} = 0.$$

Expanding by the middle row gives

$$(1-\lambda)\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = 0,$$

so $\lambda = 1$ or $\lambda^2 - 1 = 0$.

Since $\lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$, the eigenvalues are $\lambda = 1$ (repeated) and $\lambda = -1$.

The eigenvector equations are

$$\begin{aligned} -\lambda x_1 & & + \quad x_3 &= 0, \\ & (1-\lambda)x_2 & &= 0, \\ x_1 & & - \lambda x_3 &= 0. \end{aligned}$$

- For $\lambda = 1$, the eigenvector equations become

  $$-x_1 + x_3 = 0, \quad 0 = 0, \quad x_1 - x_3 = 0,$$

  which reduce to the single equation $x_3 = x_1$.

  Since x_2 can take any value, two linearly independent eigenvectors are $(0 \quad 1 \quad 0)^T$ and $(1 \quad 0 \quad 1)^T$.

- For $\lambda = -1$, the eigenvector equations become

  $$x_1 + x_3 = 0, \quad 2x_2 = 0, \quad x_1 + x_3 = 0,$$

  which reduce to the equations $x_3 = -x_1$ and $x_2 = 0$, so an eigenvector is $(1 \quad 0 \quad -1)^T$.

(b) The characteristic equation is

$$\begin{vmatrix} \frac{1}{\sqrt{2}} - \lambda & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1-\lambda & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} - \lambda \end{vmatrix} = 0,$$

which, after expanding by the middle row, simplifies to

$$(1-\lambda)(\lambda^2 - \sqrt{2}\lambda + 1) = 0.$$

The quadratic equation $\lambda^2 - \sqrt{2}\lambda + 1 = 0$ has roots

$$\lambda = \tfrac{1}{\sqrt{2}}(1+i) \quad \text{and} \quad \lambda = \tfrac{1}{\sqrt{2}}(1-i).$$

Thus the eigenvalues are $\lambda = 1$, $\lambda = \frac{1}{\sqrt{2}}(1+i)$ and $\lambda = \frac{1}{\sqrt{2}}(1-i)$.

The eigenvector equations are

$$\begin{aligned} \left(\tfrac{1}{\sqrt{2}} - \lambda\right) x_1 & & + \qquad \tfrac{1}{\sqrt{2}} x_3 &= 0, \\ & (1-\lambda)x_2 & &= 0, \\ -\tfrac{1}{\sqrt{2}} x_1 & & + \left(\tfrac{1}{\sqrt{2}} - \lambda\right) x_3 &= 0. \end{aligned}$$

- For $\lambda = 1$, the eigenvector equations become

  $$\left(\tfrac{1}{\sqrt{2}} - 1\right)x_1 + \tfrac{1}{\sqrt{2}}x_3 = 0, \quad 0 = 0, \quad -\tfrac{1}{\sqrt{2}}x_1 + \left(\tfrac{1}{\sqrt{2}} - 1\right)x_3 = 0,$$

  which reduce to the equations $x_1 = 0$ and $x_3 = 0$, so a corresponding eigenvector is $(0 \quad 1 \quad 0)^T$.
- For $\lambda = \frac{1}{\sqrt{2}}(1+i)$, the eigenvector equations become

  $$-\tfrac{1}{\sqrt{2}}ix_1 + \tfrac{1}{\sqrt{2}}x_3 = 0, \quad \tfrac{1}{\sqrt{2}}(\sqrt{2} - 1 - i)x_2 = 0,$$
  $$-\tfrac{1}{\sqrt{2}}x_1 - \tfrac{1}{\sqrt{2}}ix_3 = 0,$$

  which reduce to the equations $x_3 = ix_1$ and $x_2 = 0$, so a corresponding eigenvector is $(1 \quad 0 \quad i)^T$.
- For $\lambda = \frac{1}{\sqrt{2}}(1-i)$, the eigenvector equations become

  $$\tfrac{1}{\sqrt{2}}ix_1 + \tfrac{1}{\sqrt{2}}x_3 = 0, \quad \tfrac{1}{\sqrt{2}}(\sqrt{2} - 1 + i)x_2 = 0,$$
  $$-\tfrac{1}{\sqrt{2}}x_1 + \tfrac{1}{\sqrt{2}}ix_3 = 0,$$

  which reduce to the equations $x_3 = -ix_1$ and $x_2 = 0$, so a corresponding eigenvector is $(1 \quad 0 \quad -i)^T$.

(The eigenvalue $\frac{1}{\sqrt{2}}(1-i)$ and eigenvector $(1 \quad 0 \quad -i)^T$ can be obtained from $\frac{1}{\sqrt{2}}(1+i)$ and $(1 \quad 0 \quad i)^T$, respectively, by replacing i by $-i$. That is, the second complex eigenvalue and corresponding eigenvector are the complex conjugates of the first complex eigenvalue and corresponding eigenvector.)

(c) The matrix is upper triangular, so the eigenvalues are 2, -3 and 4.

- For $\lambda = 2$, the eigenvector equations become

  $$x_2 - x_3 = 0, \quad -5x_2 + 2x_3 = 0, \quad 2x_3 = 0,$$

  which reduce to $x_2 = x_3 = 0$, so a corresponding eigenvector is $(1 \quad 0 \quad 0)^T$.
- For $\lambda = -3$, the eigenvector equations become

  $$5x_1 + x_2 - x_3 = 0, \quad 2x_3 = 0, \quad 7x_3 = 0,$$

  which reduce to $5x_1 + x_2 = 0$ and $x_3 = 0$, so a corresponding eigenvector is $(1 \quad -5 \quad 0)^T$.
- For $\lambda = 4$, the eigenvector equations become

  $$-2x_1 + x_2 - x_3 = 0, \quad -7x_2 + 2x_3 = 0, \quad 0 = 0.$$

  Choosing $x_3 = 14$ keeps the numbers simple, and a corresponding eigenvector is $(-5 \quad 4 \quad 14)^T$.

(d) The characteristic equation is

$$\begin{vmatrix} -\lambda & 2 & 0 \\ -2 & -\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = 0,$$

and (after expanding the determinant by the third row) this gives $(1-\lambda)(\lambda^2 + 4) = 0$, so the eigenvalues are 1, $2i$ and $-2i$.

- For $\lambda = 1$, the eigenvector equations become

  $$-x_1 + 2x_2 = 0, \quad -2x_1 - x_2 = 0, \quad 0 = 0,$$

  so a corresponding eigenvector is $(0 \quad 0 \quad 1)^T$.

- For $\lambda = 2i$, the eigenvector equations become

  $$-2ix_1 + 2x_2 = 0, \quad -2x_1 - 2ix_2 = 0, \quad (1 - 2i)x_3 = 0,$$

  which reduce to $x_2 = ix_1$ and $x_3 = 0$, so a corresponding eigenvector is $(1 \quad i \quad 0)^T$.

- For $\lambda = -2i$, similarly, an eigenvector corresponding to $\lambda = -2i$ is $(1 \quad -i \quad 0)^T$.

(e) The characteristic equation is

$$\begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 2 \\ 0 & -2 & 5-\lambda \end{vmatrix} = 0,$$

which gives $(1-\lambda)((1-\lambda)(5-\lambda) + 4) = 0$. This simplifies to $(1-\lambda)(\lambda-3)^2 = 0$, so the eigenvalues are $\lambda = 1$ and $\lambda = 3$ (repeated).

- For $\lambda = 1$, the eigenvector equations become

  $$0 = 0, \quad 2x_3 = 0, \quad -2x_2 + 4x_3 = 0,$$

  which reduce to $x_2 = x_3 = 0$, so $(1 \quad 0 \quad 0)^T$ is a corresponding eigenvector.

- For $\lambda = 3$, the eigenvector equations become

  $$-2x_1 = 0, \quad -2x_2 + 2x_3 = 0, \quad -2x_2 + 2x_3 = 0,$$

  which reduce to $x_1 = 0$ and $x_2 = x_3$, so a corresponding eigenvector is $(0 \quad 1 \quad 1)^T$. In this case, the repeated eigenvalue $\lambda = 3$ has only one linearly independent eigenvector.

(f) The matrix is lower triangular, so the eigenvalues are -2 and 1 (repeated).

- For $\lambda = -2$, the eigenvector equations become

  $$3x_1 = 0, \quad 3x_2 = 0, \quad x_1 + x_2 = 0,$$

  which reduce to $x_1 = x_2 = 0$, so $(0 \quad 0 \quad 1)^T$ is a corresponding eigenvector.

- For $\lambda = 1$, the eigenvector equations become

  $$0 = 0, \quad 0 = 0, \quad x_1 + x_2 - 3x_3 = 0.$$

  These equations are satisfied if $x_1 = 3x_3 - x_2$ (whatever values we choose for x_2 and x_3). Two linearly independent eigenvectors can be found. For example, setting $x_2 = 1$ and $x_3 = 0$ gives $(-1 \quad 1 \quad 0)^T$, and setting $x_2 = 0$ and $x_3 = 1$ gives $(3 \quad 0 \quad 1)^T$.

Solution to Exercise 30

(a) The characteristic equation is

$$\begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = 0.$$

Expanding gives $(1-\lambda)^2 - 4 = 0$, so $1-\lambda = \pm 2$. Hence the eigenvalues are $\lambda = 3$ and $\lambda = -1$, and these are both real. (Alternatively, you could have calculated $\operatorname{tr}\mathbf{A} = 2$ and $\det\mathbf{A} = -3$, and so obtained the equation $\lambda^2 - 2\lambda - 3 = (\lambda-3)(\lambda+1) = 0$.) The eigenvector equations are

$$\begin{aligned} (1-\lambda)x + \quad\;\; 2y &= 0, \\ 2x + (1-\lambda)y &= 0. \end{aligned}$$

- For $\lambda = 3$, the eigenvector equations become

  $$-2x + 2y = 0 \quad \text{and} \quad 2x - 2y = 0,$$

  which reduce to the single equation $y = x$. So an eigenvector corresponding to $\lambda = 3$ is $\mathbf{v}_1 = (1 \quad 1)^T$.
- For $\lambda = -1$, the eigenvector equations become

  $$2x + 2y = 0 \quad \text{and} \quad 2x + 2y = 0,$$

  which reduce to the single equation $y = -x$. So an eigenvector corresponding to $\lambda = -1$ is $\mathbf{v}_2 = (1 \quad -1)^T$.

The two eigenvectors are orthogonal since

$$\mathbf{v}_1^T\mathbf{v}_2 = (1)(1) + (1)(-1) = 0.$$

(b) The characteristic equation is

$$\begin{vmatrix} 4-\lambda & -2 \\ -2 & 1-\lambda \end{vmatrix} = 0.$$

Expanding this gives $(4-\lambda)(1-\lambda) - 4 = 0$, which simplifies to $\lambda^2 - 5\lambda = 0$. (Alternatively, note that $\operatorname{tr}\mathbf{A} = 5$ and $\det\mathbf{A} = 0$, from which it immediately follows that $\lambda = 0$ and $\lambda = 5$.) So the eigenvalues are $\lambda = 0$ and $\lambda = 5$, which are real.

- For $\lambda = 0$, the eigenvector equations become

  $$4x - 2y = 0 \quad \text{and} \quad -2x + y = 0,$$

  which reduce to the single equation $y = 2x$. So an eigenvector corresponding to $\lambda = 0$ is $\mathbf{v}_1 = (1 \quad 2)^T$.
- For $\lambda = 5$, the eigenvector equations become

  $$-x - 2y = 0 \quad \text{and} \quad -2x - 4y = 0,$$

  which reduce to the single equation $x = -2y$. So an eigenvector corresponding to $\lambda = 5$ is $\mathbf{v}_2 = (-2 \quad 1)^T$.

The two eigenvectors are orthogonal since

$$\mathbf{v}_1^T\mathbf{v}_2 = (1)(-2) + (2)(1) = 0.$$

(c) The characteristic equation is

$$\begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = 0,$$

and this gives $(2-\lambda)(\lambda^2-1)=0$, so the eigenvalues are $\lambda=2$, $\lambda=1$ and $\lambda=-1$, which are all real.

- For $\lambda = 2$, the eigenvector equations become

  $$0=0, \quad -2x_2+x_3=0, \quad x_2-2x_3=0,$$

  so a corresponding eigenvector is $\mathbf{v}_1 = (1 \quad 0 \quad 0)^T$.
- For $\lambda = 1$, the eigenvector equations become

  $$x_1=0, \quad -x_2+x_3=0, \quad x_2-x_3=0,$$

  so a corresponding eigenvector is $\mathbf{v}_2 = (0 \quad 1 \quad 1)^T$.
- For $\lambda = -1$, the eigenvector equations become

  $$3x_1=0, \quad x_2+x_3=0, \quad x_2+x_3=0,$$

  so a corresponding eigenvector is $\mathbf{v}_3 = (0 \quad 1 \quad -1)^T$.

Note that

$$\begin{aligned} \mathbf{v}_1^T\mathbf{v}_2 &= (1)(0)+(0)(1)+(0)(1)=0, \\ \mathbf{v}_1^T\mathbf{v}_3 &= (1)(0)+(0)(1)+(0)(-1)=0, \\ \mathbf{v}_2^T\mathbf{v}_3 &= (0)(0)+(1)(1)+(1)(-1)=0, \end{aligned}$$

so all the eigenvectors are mutually orthogonal.

Solution to Exercise 31

The characteristic equation is

$$\begin{vmatrix} 2-\lambda & 0 & 1 \\ 0 & 3-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{vmatrix} = 0,$$

and expanding the determinant by the middle row gives

$$(3-\lambda)((2-\lambda)^2-1)=0,$$

which simplifies to

$$(3-\lambda)^2(1-\lambda)=0,$$

so the eigenvalues are $\lambda=1$ and $\lambda=3$ (repeated), which are real.

- For $\lambda = 1$, the eigenvector equations become

  $$x_1+x_3=0, \quad 2x_2=0, \quad x_1+x_3=0,$$

  so a corresponding eigenvector is $\mathbf{v}_1 = (1 \quad 0 \quad -1)^T$.

- For $\lambda = 3$, the eigenvector equations become

  $$-x_1 + x_3 = 0, \quad 0 = 0, \quad x_1 - x_3 = 0,$$

  with general solution $x_1 = x_3 = k$ and $x_2 = l$, giving the vector

  $$\begin{pmatrix} k \\ l \\ k \end{pmatrix} = k \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + l \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

  viewed as a linear combination of two linearly independent eigenvectors $\mathbf{v}_2 = (1 \quad 0 \quad 1)^T$ and $\mathbf{v}_3 = (0 \quad 1 \quad 0)^T$, which also happen to be orthogonal since

  $$\mathbf{v}_2^T \mathbf{v}_3 = (1)(0) + (0)(1) + (1)(0) = 0.$$

Moreover, $\mathbf{v}_2$ and $\mathbf{v}_3$ are both orthogonal to $\mathbf{v}_1$ since

$$\mathbf{v}_1^T \mathbf{v}_2 = (1)(1) + (0)(0) + (-1)(1) = 0,$$
$$\mathbf{v}_1^T \mathbf{v}_3 = (1)(0) + (0)(1) + (-1)(0) = 0,$$

so it is possible to find three mutually orthogonal eigenvectors.

Solution to Exercise 32

$$\mathbf{e}_1 = \mathbf{A}\mathbf{e}_0 = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

$$\mathbf{e}_2 = \mathbf{A}\mathbf{e}_1 = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 11 \\ 7 \end{pmatrix},$$

$$\mathbf{e}_3 = \mathbf{A}\mathbf{e}_2 = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 11 \\ 7 \end{pmatrix} = \begin{pmatrix} 47 \\ 39 \end{pmatrix},$$

$$\mathbf{e}_4 = \mathbf{A}\mathbf{e}_3 = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 47 \\ 39 \end{pmatrix} = \begin{pmatrix} 219 \\ 203 \end{pmatrix}.$$

Solution to Exercise 33

(a) $$\mathbf{e}_1 = \mathbf{A}\mathbf{e}_0 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

$$\mathbf{e}_2 = \mathbf{A}\mathbf{e}_1 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix},$$

$$\mathbf{e}_3 = \mathbf{A}\mathbf{e}_2 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 14 \\ 13 \end{pmatrix}.$$

(b) $\mathbf{e}_{11} = \mathbf{A}\mathbf{e}_{10} = (88\,574 \quad 88\,573)^T$.

(c) The characteristic equation is $(2-\lambda)^2 - 1 = 0$, so $2 - \lambda = \pm 1$, thus $\lambda_1 = 3$ and $\lambda_2 = 1$. Corresponding eigenvectors are $\mathbf{v}_1 = (1 \quad 1)^T$ and $\mathbf{v}_2 = (1 \quad -1)^T$.

(d) We need to determine constants α and β so that $\mathbf{e}_0 = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$, that is,

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Solving this equation for α and β gives $\alpha = \frac{1}{2}$ and $\beta = \frac{1}{2}$. Thus we have

$$\mathbf{e}_0 = \tfrac{1}{2}\mathbf{v}_1 + \tfrac{1}{2}\mathbf{v}_2.$$

(e) $\mathbf{e}_1 = \mathbf{A}\mathbf{e}_0 = \mathbf{A}(\frac{1}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2) = \frac{1}{2}\mathbf{A}\mathbf{v}_1 + \frac{1}{2}\mathbf{A}\mathbf{v}_2 = \frac{3}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2,$

$\mathbf{e}_2 = \mathbf{A}\mathbf{e}_1 = \mathbf{A}(\frac{3}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2) = \frac{3}{2}\mathbf{A}\mathbf{v}_1 + \frac{1}{2}\mathbf{A}\mathbf{v}_2 = \frac{9}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2.$

Similarly,

$$\mathbf{e}_n = \tfrac{1}{2}(3^n\mathbf{v}_1 + \mathbf{v}_2).$$

(f) The coefficient of $\mathbf{v}_1$, namely $\frac{1}{2} \times 3^n$, dominates the expression for $\mathbf{e}_n$, so we obtain an approximation to $\mathbf{v}_1$ (an eigenvector corresponding to the eigenvalue of larger magnitude).

Solution to Exercise 34

(a) If $\mathbf{e}_0 = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$, then

$$\begin{pmatrix} 10\,000 \\ 8\,000 \end{pmatrix} = \alpha \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

so α and β satisfy the simultaneous equations

$$\begin{aligned} 2\alpha + \beta &= 10\,000, \\ \alpha - \beta &= 8000. \end{aligned}$$

Solving these equations gives $\alpha = 6000$, $\beta = -2000$, so

$$\mathbf{e}_0 = 6000\mathbf{v}_1 - 2000\mathbf{v}_2.$$

(b) Since $\mathbf{A}\mathbf{v}_1 = \mathbf{v}_1$ and $\mathbf{A}\mathbf{v}_2 = 0.7\mathbf{v}_2$, we have

$$\begin{aligned} \mathbf{e}_1 &= \mathbf{A}\mathbf{e}_0 = 6000\mathbf{A}\mathbf{v}_1 - 2000\mathbf{A}\mathbf{v}_2 = 6000\mathbf{v}_1 - 2000(0.7)\mathbf{v}_2, \\ \mathbf{e}_2 &= \mathbf{A}\mathbf{e}_1 = 6000\mathbf{A}\mathbf{v}_1 - 2000(0.7)\mathbf{A}\mathbf{v}_2 = 6000\mathbf{v}_1 - 2000(0.7)^2\mathbf{v}_2. \end{aligned}$$

Continuing in this way, we obtain

$$\mathbf{e}_n = 6000\mathbf{v}_1 - 2000(0.7)^n\mathbf{v}_2.$$

(c) As n becomes large, $(0.7)^n$ becomes small and the term involving $\mathbf{v}_2$ can be ignored in comparison with the term involving $\mathbf{v}_1$. Thus

$$\mathbf{e}_n \simeq 6000\mathbf{v}_1 = (12\,000 \quad 6000)^T,$$

which agrees with our observations in the Introduction.

Solution to Exercise 35

First iteration:

$$\mathbf{z}_1 = \begin{pmatrix} 3 & 6 \\ 4 & 8 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix},$$

$$\alpha_1 = 4,$$

$$\mathbf{e}_1 = \frac{\mathbf{z}_1}{\alpha_1} = \tfrac{1}{4} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ 1 \end{pmatrix}.$$

Second iteration:

$$\mathbf{z}_2 = \begin{pmatrix} 3 & 6 \\ 4 & 8 \end{pmatrix} \begin{pmatrix} \frac{3}{4} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{33}{4} \\ 11 \end{pmatrix},$$

$$\alpha_2 = 11,$$

$$\mathbf{e}_2 = \frac{\mathbf{z}_2}{\alpha_2} = \tfrac{1}{11} \begin{pmatrix} \frac{33}{4} \\ 11 \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ 1 \end{pmatrix}.$$

Since $\mathbf{e}_2 = \mathbf{e}_1$, the third iteration will be identical to the second, and $\mathbf{e}_3$ is also $(\frac{3}{4} \quad 1)^T$.

So $(\frac{3}{4} \quad 1)^T$ is an eigenvector. We have

$$\begin{pmatrix} 3 & 6 \\ 4 & 8 \end{pmatrix} \begin{pmatrix} \frac{3}{4} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{33}{4} \\ 11 \end{pmatrix} = 11 \begin{pmatrix} \frac{3}{4} \\ 1 \end{pmatrix},$$

so the corresponding eigenvalue is 11. Since $\operatorname{tr} \mathbf{A} = 11$, the other eigenvalue is 0, which explains why the iteration converges so rapidly.

(This is a very special case; generally, we would not expect $\mathbf{e}_n$ to be *equal* to an eigenvector for any value of n, unless we start with an eigenvector.)

Solution to Exercise 36

We have

$$\mathbf{A}^{-1} = \tfrac{1}{11} \begin{pmatrix} 5 & -3 \\ -8 & 7 \end{pmatrix}.$$

First iteration:

$$\mathbf{z}_1 = \tfrac{1}{11} \begin{pmatrix} 5 & -3 \\ -8 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{2}{11} \\ -\frac{1}{11} \end{pmatrix},$$

therefore $\alpha_1 = \frac{2}{11}$ and $\mathbf{e}_1 = (1 \quad -\frac{1}{2})^T$.

Second iteration:

$$\mathbf{z}_2 = \tfrac{1}{11} \begin{pmatrix} 5 & -3 \\ -8 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{13}{22} \\ -\frac{23}{22} \end{pmatrix},$$

therefore $\alpha_2 = -\frac{23}{22}$ and $\mathbf{e}_2 = (-\frac{13}{23} \quad 1)^T$.

Third iteration:

$$\mathbf{z}_3 = \tfrac{1}{11}\begin{pmatrix} 5 & -3 \\ -8 & 7 \end{pmatrix}\begin{pmatrix} -\frac{13}{23} \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{134}{253} \\ \frac{265}{253} \end{pmatrix},$$

therefore $\alpha_3 = \frac{265}{253}$ and $\mathbf{e}_3 = (-\frac{134}{265} \quad 1)^T$.

Taking $\mathbf{e}_3 = (-\frac{134}{265} \quad 1)^T \simeq (-0.506 \quad 1)^T$ as our approximation to the eigenvector, the estimate for the corresponding eigenvalue of $\mathbf{A}$ of smallest magnitude is $1/\alpha_3 = \frac{253}{265} \simeq 0.955$. (The exact eigenvector is $(-0.5 \quad 1)^T$, corresponding to the eigenvalue 1.)

Solution to Exercise 37

We follow Procedure 6(a).

Twenty-first iteration:

$$\begin{aligned}\mathbf{z}_{21} &= (\mathbf{A}+\mathbf{I})^{-1}\mathbf{e}_{20} \\ &= \begin{pmatrix} \frac{3}{2} & -\frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{4} & \frac{5}{8} & -\frac{1}{4} \\ -\frac{1}{2} & -\frac{1}{4} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 \\ -0.141\,05 \\ -0.379\,39 \end{pmatrix} = \begin{pmatrix} 1.724\,96 \\ -0.243\,31 \\ -0.654\,43 \end{pmatrix}, \\ \alpha_{21} &= 1.724\,96, \\ \mathbf{e}_{21} &= \frac{\mathbf{z}_{21}}{\alpha_{21}} \simeq (1 \quad -0.141\,05 \quad -0.379\,39)^T.\end{aligned}$$

The sequence α_n appears to have converged to 1.724 96, and a corresponding eigenvector is approximately $(1 \quad -0.141 \quad -0.379)^T$ (here given to three decimal places).

Since α_n converges to $1/(\lambda_1 - p) = 1/(\lambda_1 + 1) = 1.724\,96$, we have

$$\lambda_1 = \frac{1}{1.724\,96} - 1 \simeq -0.420.$$

We can be confident of results quoted to three decimal places since convergence of the iteration appears to have occurred at five decimal places.

Solution to Exercise 38

(a) We follow Procedure 4.

First iteration:

$$\begin{aligned}\mathbf{z}_1 &= \mathbf{A}\mathbf{e}_0 = \begin{pmatrix} 0.5 & 0.6 \\ 1.4 & -0.3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1.4 \end{pmatrix}, \\ \alpha_1 &= 1.4, \\ \mathbf{e}_1 &= \frac{\mathbf{z}_1}{\alpha_1} = \frac{1}{1.4}\begin{pmatrix} 0.5 \\ 1.4 \end{pmatrix} = \begin{pmatrix} 0.357\,143 \\ 1 \end{pmatrix}.\end{aligned}$$

Second iteration: $\mathbf{e}_2 = (1 \quad 0.256\,881)^T$.

Third iteration: $\mathbf{e}_3 = (0.494\,452 \quad 1)^T$.

(b) We follow Procedure 1.

The characteristic equation is

$$\lambda^2 - 0.2\lambda - 0.99 = 0,$$

so the eigenvalues are $\lambda_1 = -0.9$ and $\lambda_2 = 1.1$.

- For $\lambda_1 = -0.9$, the eigenvector equations both become

  $$1.4x + 0.6y = 0, \quad \text{i.e.} \quad 3y = -7x,$$

  so a corresponding eigenvector is $(3 \quad -7)^T$.
- For $\lambda_2 = 1.1$, the eigenvector equations become

  $$-0.6x + 0.6y = 0, \quad 1.4x - 1.4y = 0,$$

  which reduce to $x = y$, so a corresponding eigenvector is $(1 \quad 1)^T$.

(c) The sequence $\mathbf{e}_n$ will converge to an eigenvector corresponding to the eigenvalue of largest magnitude, that is, to $(1 \quad 1)^T$.

(d) $\mathbf{A}^8\mathbf{v} = \lambda^8\mathbf{v}$.

(e) We express $\mathbf{e}_0$ in terms of the eigenvectors as $\mathbf{e}_0 = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$, so

$$\mathbf{e}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \alpha \begin{pmatrix} 3 \\ -7 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

giving $\alpha = 0.1$ and $\beta = 0.7$.

In general, $\mathbf{e}_n = \mathbf{A}^n\mathbf{e}_0$, so we calculate

$$\begin{aligned} \mathbf{A}^8\mathbf{e}_0 &= \mathbf{A}^8(0.1\mathbf{v}_1 + 0.7\mathbf{v}_2) \\ &= 0.1(\mathbf{A}^8\mathbf{v}_1) + 0.7(\mathbf{A}^8\mathbf{v}_2) \\ &= (0.1)(-0.9)^8\mathbf{v}_1 + (0.7)(1.1)^8\mathbf{v}_2 \\ &\simeq 0.043\,047\mathbf{v}_1 + 1.500\,512\mathbf{v}_2 \\ &= (0.129\,140 \quad -0.301\,327)^T + (1.500\,512 \quad 1.500\,512)^T \\ &= (1.629\,652 \quad 1.199\,185)^T. \end{aligned}$$

Dividing by 1.629 652, we obtain $\mathbf{e}_8 = (1 \quad 0.735\,853)^T$.

(f) $\mathbf{A}^{-1} = \begin{pmatrix} 0.303\,030 & 0.606\,060 \\ 1.414\,141 & -0.505\,050 \end{pmatrix}$.

We follow Procedure 5(a).

First iteration:

$$\mathbf{z}_1 = \mathbf{A}^{-1}\mathbf{e}_0 = \begin{pmatrix} 0.484\,848 \\ -1.070\,707 \end{pmatrix},$$

$$\alpha_1 = -1.070\,707,$$

$$\mathbf{e}_1 = \frac{\mathbf{z}_1}{\alpha_1} = -\frac{1}{1.070\,707}\begin{pmatrix} 0.484\,848 \\ -1.070\,707 \end{pmatrix} = \begin{pmatrix} -0.452\,830 \\ 1 \end{pmatrix}.$$

Second iteration:

$$\alpha_2 = -1.145\,416, \quad \mathbf{e}_2 = (-0.409\,318 \quad 1)^T.$$

Third iteration:

$$\alpha_3 = -1.083\,884, \quad \mathbf{e}_3 = (-0.444\,720 \quad 1)^T.$$

This sequence of vectors is converging to

$$(-\tfrac{3}{7} \quad 1)^T \simeq (-0.428\,571 \quad 1)^T,$$

an eigenvector corresponding to the eigenvalue of smallest magnitude, $\lambda = -0.9$.

(g) Convergence is slow for direct iteration, since the two eigenvalues of $\mathbf{A}$ are -0.9 and 1.1, which are relatively close in magnitude. Similarly, inverse iteration is also slow to converge, since the eigenvalues of $\mathbf{A}^{-1}$ are -1.1111 and 0.9091, to four decimal places, and these are relatively close in magnitude.

Solution to Exercise 39

We could use Procedure 6 to find the third eigenvector, but without a computer this could be very hard work. Alternatively, we could find two of the eigenvalues from the given eigenvectors, and we could then find the third eigenvalue very easily because the sum of the eigenvalues is $\operatorname{tr}\mathbf{A}$.

We know that $\mathbf{v}_1$ is an eigenvector, so $\mathbf{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1$. Calculating just the third components, we have

$$(3 \times 0.477) + (4 \times 0.689) + (6 \times 1) = \lambda_1 \times 1,$$

so $\lambda_1 \simeq 10.187$.

Similarly, using the second components of $\mathbf{A}\mathbf{v}_2 = \lambda_2\mathbf{v}_2$, we have

$$(2 \times (-0.102)) + (3 \times 1) + (4 \times (-0.641)) = \lambda_2 \times 1,$$

so $\lambda_2 \simeq 0.232$.

Since $\operatorname{tr}\mathbf{A} = 1 + 3 + 6 = 10$,

$$\lambda_3 = 10 - \lambda_1 - \lambda_2 \simeq -0.419.$$