7 Eigenvalues and Eigenvectors

SECTION 7.1 Introduction to Eigenvalues and Eigenvectors

By the end of this section you will be able to

- determine eigenvalues and eigenvectors
- prove properties of eigenvalues and eigenvectors

Eigenvector/value problems crop up frequently in the physical sciences and engineering. They take the form $\mathbf{Av} = (\text{scalar}) \times \mathbf{v}$ where $\mathbf{v}$ is a non-zero vector and $\mathbf{A}$ is a square matrix. By knowing the eigenvalues and eigenvectors of a matrix we can easily find its determinant, decide whether the matrix has an inverse and determine the powers of the matrix. For an example of linear algebra at work, one needs to look no further than Google's search engine, which relies upon eigenvalues and eigenvectors to rank pages with respect to relevance.

7.1.1 Definition of eigenvalues and eigenvectors

Before we define what is meant by an eigenvalue and an eigenvector let's do an example which involves them.

Example 7.1

Let $\mathbf{A} = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}$ and $\mathbf{u} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ then evaluate $\mathbf{Au}$.

Solution

Multiplying the matrix $\mathbf{A}$ and vector $\mathbf{u}$ we have

$$\mathbf{Au} = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

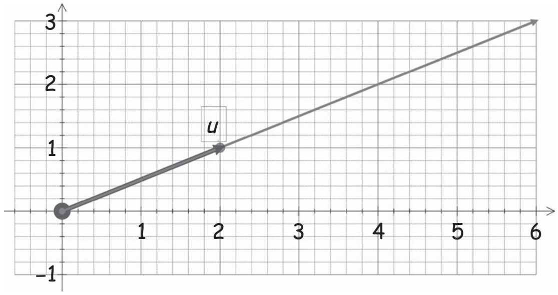
What do you notice about the result?

We have $\mathbf{Au} = 3\mathbf{u}$. The matrix $\mathbf{A}$ scalar multiplies the vector $\mathbf{u}$ by 3, as shown in Fig. 7.1.

In general terms, this can be written as

(7.1) $$\mathbf{Au} = \lambda\mathbf{u} \quad (\text{matrix } \mathbf{A} \text{ scalar multiplies vector } \mathbf{u})$$

where $\mathbf{A}$ is a square matrix, $\mathbf{u}$ is a vector and the Greek letter λ (lambda) is a scalar.



$$\mathbf{Au} = 3\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Figure 7.1

This is an important result which is used throughout this chapter and well worth becoming familiar with.

Why is formula (7.1) important?

Because the matrix $\mathbf{A}$ transforms the vector $\mathbf{u}$ by scalar multiplying it, which means that the transformation only changes the length of the vector $\mathbf{u}$ unless $\lambda = \pm 1$ (in which case the length remains unchanged). Note that (7.1) says that the matrix $\mathbf{A}$ applied to $\mathbf{u}$ gives a vector in the same or opposite (negative λ) direction of $\mathbf{u}$.

Can you think of a vector, $\boldsymbol{u}$, which satisfies equation (7.1)?

The zero vector $\mathbf{u} = \mathbf{O}$ because $\mathbf{AO} = \lambda\mathbf{O} = \mathbf{O}$. In this case we say that we have the trivial solution $\mathbf{u} = \mathbf{O}$. In this chapter we consider the non-trivial solutions, $\mathbf{u} \neq \mathbf{O}$ (not zero), and these solutions are powerful tools in linear algebra.

For a *non-zero* vector $\mathbf{u}$ the scalar λ is called an **eigenvalue** of the matrix $\mathbf{A}$ and the vector $\mathbf{u}$ is called an **eigenvector** belonging to or corresponding to λ, which satisfies $\mathbf{Au} = \lambda\mathbf{u}$.

In most linear algebra literature the Greek letter lambda, λ, is used for eigenvalues. These terms eigenvalue and eigenvector are derived from the German word 'Eigenwert' which means 'proper value'. The word eigen is pronounced 'i-gun'.

Eigenvalues were initially developed in the field of differential equations by Jean d' Alembert.

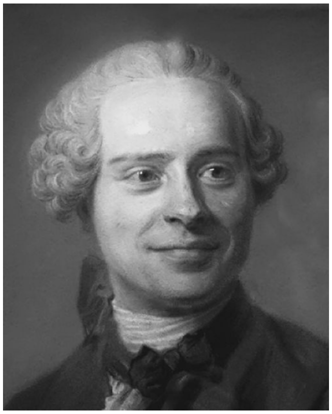


Figure 7.2 Jean d'Alembert 1717 to 1783.

Jean d'Alembert 1717–1783 (Fig. 7.2) was a French mathematician and the illegitimate son of Madam Tencin and an army officer, Louis Destouches. His mother left him on the steps of a local church and he was consequently sent to a home for orphans. His father recognised his son's difficulties and placed him under the care of Madam Rousseau, wife of a wealthy architect.

However, d'Alembert's father died when he was only nine years old and his father's family looked after his financial situation so that he could continue his education.

In 1735, Alembert graduated, and he thought that a career in law would suit him, but his real thirst and enthusiasm was for mathematics, and he studied this in his spare time. For most of his life he worked for the Paris Academy of Science and the French Academy.

Example 7.2

Let $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$. Verify the following:

(a) $\mathbf{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector of matrix $\mathbf{A}$ belonging to the eigenvalue $\lambda_1 = 2$.

(b) $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is an eigenvector of matrix $\mathbf{A}$ belonging to the eigenvalue $\lambda_2 = 3$.

Solution

(a) Multiplying the given matrix $\mathbf{A}$ and vector $\mathbf{u}$ we have

$$\mathbf{Au} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Thus $\mathbf{u} = (1 \quad 1)^T$ is an eigenvector of the matrix $\mathbf{A}$ belonging to $\lambda_1 = 2$ because $\mathbf{Au} = 2\mathbf{u}$. Matrix $\mathbf{A}$ doubles the vector $\mathbf{u}$.

(b) Similarly we have

$$\mathbf{Av} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Thus $\mathbf{v} = (1 \quad 2)^T$ is an eigenvector of the matrix $\mathbf{A}$ belonging to $\lambda_2 = 3$ because $\mathbf{Av} = 3\mathbf{v}$. This $\mathbf{Av} = 3\mathbf{v}$ means matrix $\mathbf{A}$ triples the vector $\mathbf{v}$.

What do you notice about your results?
A 2 by 2 matrix can have *more than* one eigenvalue and eigenvector.

We have eigenvalues λ and eigenvectors $\mathbf{u}$ for any *square* matrix $\mathbf{A}$ such that $\mathbf{Au} = \lambda\mathbf{u}$.

Example 7.3

Let $\mathbf{A} = \begin{pmatrix} 5 & 0 & 0 \\ -9 & 4 & -1 \\ -6 & 2 & 1 \end{pmatrix}$ and $\mathbf{u} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$. Show that the matrix $\mathbf{A}$ scalar multiplies the vector $\mathbf{u}$ and find the value of this scalar, λ, the eigenvalue.

(continued...)

Solution

Applying the matrix $\mathbf{A}$ to the vector $\mathbf{u}$ we have

$$\mathbf{Au} = \begin{pmatrix} 5 & 0 & 0 \\ -9 & 4 & -1 \\ -6 & 2 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix} = 2\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

We have $\mathbf{Au} = 2\mathbf{u}$ so $\lambda = 2$. Hence $\lambda = 2$ is an eigenvalue of the matrix $\mathbf{A}$ with an eigenvector $\mathbf{u}$. Matrix $\mathbf{A}$ transforms the vector $\mathbf{u}$ by a scalar multiple of 2 because $\mathbf{Au} = 2\mathbf{u}$.

7.1.2 Characteristic equation

From the above formula (7.1) $\mathbf{Au} = \lambda\mathbf{u}$ we have

$$\mathbf{Au} = \lambda\mathbf{Iu}$$

[$\lambda\mathbf{Iu} = \lambda\mathbf{u}$ – multiplying by the identity keeps it the same]
where $\mathbf{I}$ is the identity matrix. We can rewrite this as

$$\mathbf{Au} - \lambda\mathbf{Iu} = \mathbf{O}$$
$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = \mathbf{O}$$

Under what condition is the non-zero vector $\mathbf{u}$ a solution of this equation?
By question 26 of Exercises 6.3:

$$\mathbf{Ax} = \mathbf{O} \text{ has an infinite number of solutions} \Leftrightarrow \det(\mathbf{A}) = 0.$$

Applying this result to $(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = \mathbf{O}$ means that we must have a non-zero vector $\mathbf{u}$ (because there are an infinite number of solutions which satisfy this equation) $\Leftrightarrow$

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

This is an important equation because we use this to find the eigenvalues and it is called the **characteristic equation**:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0 \tag{7.2}$$

The procedure for determining eigenvalues and eigenvectors is:

1. Solve the characteristic equation (7.2) for the scalar λ.
2. For the eigenvalue λ determine the corresponding eigenvector $\mathbf{u}$ by solving the system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = \mathbf{O}$.

Let's follow this procedure for the next example.

Note that eigenvalues and eigenvectors come in pairs. You *cannot* have one without the other. It is a relationship like mother and child because eigenvalues give birth to eigenvectors.

Example 7.4

Determine the eigenvalues and corresponding eigenvectors of $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}$. Also sketch the effect of multiplying the eigenvectors by matrix $\mathbf{A}$.

Solution

What do we find first, the eigenvalues or eigenvectors?

Eigenvalues, because they produce eigenvectors.

We carry out the above procedure:

Step 1.

We need to find the values of λ which satisfy $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. First we obtain $\mathbf{A} - \lambda\mathbf{I}$:

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

Substituting this into $\det(\mathbf{A} - \lambda\mathbf{I})$ gives

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det \begin{pmatrix} 2-\lambda & 0 \\ 1 & 3-\lambda \end{pmatrix}$$

To find the determinant, we use formula (6.1), $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$, thus:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det \begin{pmatrix} 2-\lambda & 0 \\ 1 & 3-\lambda \end{pmatrix} = (2-\lambda)\,(3-\lambda) - 0$$

For eigenvalues we equate this determinant to zero:

$$(2-\lambda)\,(3-\lambda) = 0 \quad \text{implies } \lambda_1 = 2 \text{ or } \lambda_2 = 3$$

Step 2.

For each eigenvalue, λ determine the corresponding eigenvector $\mathbf{u}$ by solving the system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = \mathbf{O}$.

Let $\mathbf{u}$ be the eigenvector corresponding to $\lambda_1 = 2$. Substituting $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}$ and $\lambda_1 = \lambda = 2$ into $(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = \mathbf{O}$ gives

$$\begin{aligned} (\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = \left[\begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \right] \mathbf{u} &= \mathbf{O} \\ \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \mathbf{u} &= \mathbf{O} \end{aligned}$$

Remember, $\mathbf{O} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and let $\mathbf{u} = \begin{pmatrix} x \\ y \end{pmatrix}$, so we have

$$\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Multiplying out gives

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

(continued...)

Remember, the eigenvector *cannot* be the zero vector, therefore at least one of the values, x or y, must be non-zero. From the bottom equation we have $x = -y$.

The simplest solution is $x = 1$, $y = -1$ but we could have $\begin{pmatrix} 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -3 \\ 3 \end{pmatrix}, \begin{pmatrix} \pi \\ -\pi \end{pmatrix}, \begin{pmatrix} 5 \\ -5 \end{pmatrix}, \ldots$

Hence we have an infinite number of eigenvectors belonging to $\lambda = 2$. We can write down the general eigenvector $\mathbf{u}$.

How?

Let $x = s$ then $y = -s$ where $s \neq 0$ and is a real number. Thus the eigenvectors belonging to $\lambda = 2$ are

$$\mathbf{u} = \begin{pmatrix} s \\ -s \end{pmatrix} = s\begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ where } s \neq 0 \qquad \left[\begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ is one of the simplest eigenvectors}\right]$$

Similarly, we find the general eigenvector $\mathbf{v}$ belonging to the other eigenvalue $\lambda_2 = 3$. Putting $\lambda_2 = \lambda = 3$ into $[\mathbf{A} - \lambda\mathbf{I}]\,\mathbf{v} = \mathbf{O}$ gives

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \left[\begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix} - \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}\right]\mathbf{v} = \mathbf{O} \text{ simplifies to } \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix}\mathbf{v} = \mathbf{O}$$

By writing $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ [different x and y from those above] and $\mathbf{O} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ we obtain

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

Multiplying out:

$$-x + 0 = 0, \quad x + 0 = 0$$

From these equations we must have $x = 0$.

What is y equal to?

We can choose y to be any real number apart from zero because the eigenvector cannot be zero. Thus

$$y = s \text{ where } s \neq 0$$

The general eigenvector belonging to $\lambda_2 = 3$ is

$$\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ s \end{pmatrix} = s\begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ where } s \neq 0 \qquad \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ is one of the simplest eigenvectors}\right]$$

Summarizing the above we have:

Eigenvector $\mathbf{u} = s\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ belonging to $\lambda_1 = 2$ and eigenvector $\mathbf{v} = s\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ belonging to $\lambda_2 = 3$.

What does all this mean?

The given matrix $\mathbf{A}$ scalar multiplies the eigenvector $\mathbf{u}$ by 2 and $\mathbf{v}$ by 3 because

$$\mathbf{Au} = 2\mathbf{u} \text{ and } \mathbf{Av} = 3\mathbf{v}$$

Plotting these eigenvectors, the effect of multiplying by the matrix $\mathbf{A}$ is shown in Fig. 7.3.

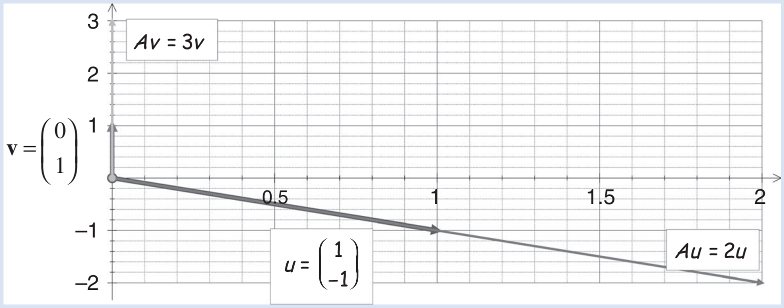


Figure 7.3

Matrix $\mathbf{A}$ doubles ($\lambda_1 = 2$) the eigenvector $\mathbf{u}$ and triples ($\lambda_2 = 3$) the eigenvector $\mathbf{v}$ as you can see in Fig. 7.3. Matrix $\mathbf{A}$ does *not* change the direction of the eigenvectors.

Eigenvectors are non-zero vectors which are transformed by the matrix $\mathbf{A}$ to a scalar multiple λ of itself.

Next, we find the eigenvalues and eigenvectors of a 3 by 3 matrix. Follow the algebra carefully because you will have to expand brackets like $(1-\lambda)(-3-\lambda)$.

To expand this it is usually easier to take out two minus signs and then expand, that is:

$$(1-\lambda)(-3-\lambda) = --(-1+\lambda)(3+\lambda) = (\lambda-1)(3+\lambda) \quad [\text{Because } -- = +]$$

Example 7.5

Determine the eigenvalues of $\mathbf{A} = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 4 & 0 \\ 3 & 5 & -3 \end{pmatrix}$

Solution

We have

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 4 & 0 \\ 3 & 5 & -3 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} 1-\lambda & 0 & 4 \\ 0 & 4-\lambda & 0 \\ 3 & 5 & -3-\lambda \end{pmatrix}$$

It is easier to remember that $\mathbf{A} - \lambda\mathbf{I}$ is actually matrix $\mathbf{A}$ with $-\lambda$ along the leading diagonal (from top left to bottom right). We need to evaluate $\det(\mathbf{A} - \lambda\mathbf{I})$.

What is the simplest way to find $\det(\mathbf{A} - \lambda\mathbf{I})$*?*

From the properties of determinants of the last chapter, we know that it will be easier to evaluate the determinant along the middle row, containing the elements 0, $4-\lambda$ and 0.

Why?

Because it has two zeros we do *not* have to evaluate the 2 by 2 determinants associated with these zeros. [Spending a second or two in choosing an easy way forward can really help save on the arithmetic later on.] From above we have

(continued...)

$$
\begin{aligned}
\det(\mathbf{A}-\lambda\mathbf{I}) &= \det\begin{pmatrix} 1-\lambda & 0 & 4 \\ 0 & 4-\lambda & 0 \\ 3 & 5 & -3-\lambda \end{pmatrix} \longleftarrow \text{middle row} \\
&= (4-\lambda)\left[\det\begin{pmatrix} 1-\lambda & 4 \\ 3 & -3-\lambda \end{pmatrix}\right] && \left[\begin{matrix}\text{expanding the}\\ \text{middle Row}\end{matrix}\right] \\
&= (4-\lambda)\left[(1-\lambda)(-3-\lambda)-(3\times 4)\right] && [\text{by determinant of 2 by 2}] \\
&= (4-\lambda)\left[(\lambda-1)(3+\lambda)-12\right] && [\text{taking out minus signs}] \\
&= (4-\lambda)\left[3\lambda+\lambda^2-3-\lambda-12\right] && [\text{opening brackets}] \\
&= (4-\lambda)\left[\lambda^2+2\lambda-15\right] && [\text{simplifying}] \\
&= (4-\lambda)(\lambda+5)(\lambda-3) && [\text{factorizing}]
\end{aligned}
$$

By the characteristic equation (7.2), $\det(\mathbf{A}-\lambda\mathbf{I})=0$, we equate all the above to zero:

$$(4-\lambda)(\lambda+5)(\lambda-3)=0$$

Solving this equation gives the eigenvalues $\lambda_1=4$, $\lambda_2=-5$ and $\lambda_3=3$.

Example 7.6

Determine the eigenvectors associated with $\lambda_3=3$ for the matrix $\mathbf{A}$ given in Example 7.5.

Solution

Substituting the eigenvalue $\lambda_3=\lambda=3$ and the matrix $\mathbf{A}=\begin{pmatrix} 1 & 0 & 4 \\ 0 & 4 & 0 \\ 3 & 5 & -3 \end{pmatrix}$ into $(\mathbf{A}-\lambda\mathbf{I})\mathbf{u}=\mathbf{O}$ (subtract 3 from the leading diagonal) gives:

$$(\mathbf{A}-3\mathbf{I})\,\mathbf{u}=\begin{pmatrix} 1-3 & 0 & 4 \\ 0 & 4-3 & 0 \\ 3 & 5 & -3-3 \end{pmatrix}\mathbf{u}=\mathbf{O}$$

where $\mathbf{u}$ is the eigenvector corresponding to $\lambda_3=3$.
What is the zero vector, $\mathbf{O}$, equal to?

Remember, this zero vector is $\mathbf{O}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Let $\mathbf{u}=\begin{pmatrix} x \\ y \\ z \end{pmatrix}$. Substituting these into the above and simplifying gives

$$\begin{pmatrix} -2 & 0 & 4 \\ 0 & 1 & 0 \\ 3 & 5 & -6 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Expanding this yields the linear system

$$
\begin{aligned}
-2x + 0 + 4z &= 0 \qquad (1) \\
0 + y + 0 &= 0 \qquad (2) \\
3x + 5y - 6z &= 0 \qquad (3)
\end{aligned}
$$

From the middle equation (2) we have $y = 0$. From the top equation (1) we have

$$2x = 4z \text{ which gives } x = 2z$$

If $z = 1$ then $x = 2$; or more generally if $z = s$ then $x = 2s$ where $s \neq 0$ [not zero].

The general eigenvector $\mathbf{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2s \\ 0 \\ s \end{pmatrix} = s\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ where $s \neq 0$ and corresponds to $\lambda_3 = 3$.

The given matrix $\mathbf{A}$ triples the eigenvector $\mathbf{u}$ because $\mathbf{Au} = 3\mathbf{u}$.

You are asked to find the eigenvectors belonging to $\lambda_1 = 4$ and $\lambda_2 = -5$ in Exercises 7.1.

7.1.3 Eigenspace

Note that for $\lambda_3 = 3$ in the above Example 7.6 we have an infinite number of eigenvectors by substituting various non-zero values of s:

$$\mathbf{u} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \text{ or } \mathbf{u} = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix} \text{ or } \mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 1/2 \end{pmatrix} \text{ or } \mathbf{u} = \begin{pmatrix} -4 \\ 0 \\ -2 \end{pmatrix} \ldots$$

Check that the matrix $\mathbf{A}$ triples each of these eigenvectors by verifying $\mathbf{Au} = 3\mathbf{u}$. The above solutions $\mathbf{u}$ are given by *all* the points (apart from $x = y = z = 0$) on the line shown in Fig. 7.4:

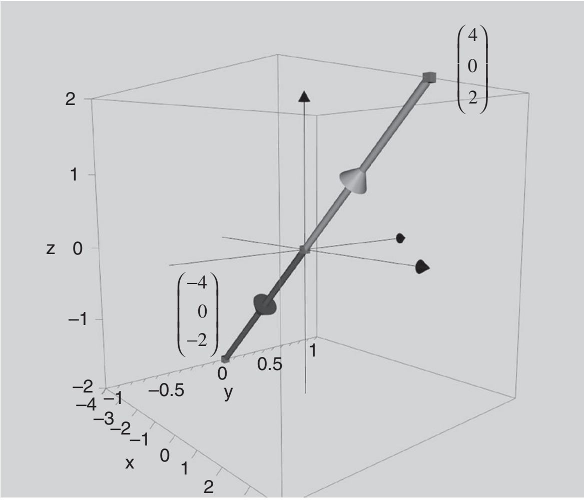


Figure 7.4

In general, if $\mathbf{A}$ is a square matrix and λ is an eigenvalue of $\mathbf{A}$ with an eigenvector $\mathbf{u}$ then every scalar multiple (apart from 0) of the vector $\mathbf{u}$ is also an eigenvector belonging to the eigenvalue λ. For example, if we have $\mathbf{Au} = 3\mathbf{u}$ then $666\mathbf{u}$ is also an eigenvector because

$$\mathbf{A}(666\mathbf{u}) = 666(\mathbf{Au}) = 666 \underbrace{(3\mathbf{u})}_{\text{because } \mathbf{Au}=3\mathbf{u}} = 3(666\mathbf{u})$$

Since the matrix $\mathbf{A}$ *triples* the vector $666\mathbf{u}$ so $666\mathbf{u}$ is an eigenvector with eigenvalue 3. Thus we have the general proposition:

Proposition (7.3). If λ is an eigenvalue of a square matrix $\mathbf{A}$ with an eigenvector $\mathbf{u}$ then *every* non-zero scalar multiplication of $\mathbf{u}$, such as $k\mathbf{u}$, is also an eigenvector belonging to λ.

This means that if $\mathbf{u}$ is an eigenvector belonging to λ then so is $2\mathbf{u}$, $0.53\mathbf{u}$, $-666\mathbf{u}$, ...

Proof.
Consider an arbitrary non-zero scalar k, then

$$\begin{aligned} \mathbf{A}(k\mathbf{u}) &= k(\mathbf{Au}) && \left[\text{by rules of matrices}\right] \\ &= k(\lambda\mathbf{u}) && \left[\text{by (7.1) } \mathbf{Au} = \lambda\mathbf{u}\right] \\ &= \lambda(k\mathbf{u}) \end{aligned}$$

Thus we have $\mathbf{A}(k\mathbf{u}) = \lambda(k\mathbf{u})$, which means that the matrix $\mathbf{A}$ acting on the vector $k\mathbf{u}$ produces a scalar multiple λ of $k\mathbf{u}$. Hence $k\mathbf{u}$ is an eigenvector belonging to the eigenvalue λ. Since k was arbitrary, every non-zero scalar multiple of $\mathbf{u}$ is an eigenvector of the matrix $\mathbf{A}$ belonging to the eigenvalue λ.

Hence the scalar λ produces an infinite number of eigenvectors.

■

Proposition (7.4). If $\mathbf{A}$ is an n by n matrix with an eigenvalue of λ, then the set S of *all* eigenvectors of $\mathbf{A}$ belonging to λ together with the zero vector, $\mathbf{O}$, is a subspace of $\mathbb{R}^n$:

$$S = \{\mathbf{O}\} \cup \left\{\mathbf{u} \,\middle|\, \mathbf{u} \text{ is an eigenvector belonging to } \lambda\right\}$$

How do we prove the given set is a subspace of $\mathbb{R}^n$?
We can use result (3.7) of chapter 3:

Proposition (3.7). A non-empty subset S containing vectors $\mathbf{u}$ and $\mathbf{v}$ is a subspace of a vector space $V \Leftrightarrow$ any linear combination $k\mathbf{u} + c\mathbf{v}$ is also in S (k and c are scalars).

This means that we need to show:

If vectors $\mathbf{u}$ and $\mathbf{v}$ are in S then for any scalars k and c we have $k\mathbf{u} + c\mathbf{v}$ is also in S. This means that S is a subspace $\Leftrightarrow$ S is closed so the vector $k\mathbf{u} + c\mathbf{v}$ cannot escape from S, as shown in Fig. 7.5.

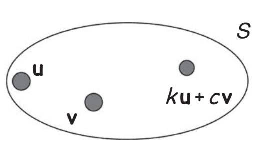


Figure 7.5

Proof.

What do we need to prove?

Required to prove that if $\mathbf{u}$ and $\mathbf{v}$ are eigenvectors belonging to the eigenvalue λ then $k\mathbf{u} + c\mathbf{v}$ is also an eigenvector belonging to λ.

Let $\mathbf{u}$ and $\mathbf{v}$ be eigenvectors belonging to the same eigenvalue λ, and k and c be any non-zero scalars. Then by the above Proposition (7.3):

If λ is an eigenvalue of a square matrix $\mathbf{A}$ with an eigenvector $\mathbf{u}$ then $k\mathbf{u}$ is also an eigenvector belonging to λ.

We have $k\mathbf{u}$ and $c\mathbf{v}$ are eigenvectors belonging to λ, therefore by (7.1):

$$\mathbf{A}(k\mathbf{u}) = \lambda(k\mathbf{u}) \text{ and } \mathbf{A}(c\mathbf{v}) = \lambda(c\mathbf{v}) \qquad (*)$$

We need to show that $\mathbf{A}(k\mathbf{u} + c\mathbf{v}) = \lambda(k\mathbf{u} + c\mathbf{v})$:

$$\begin{aligned}
\mathbf{A}(k\mathbf{u} + c\mathbf{v}) &= \mathbf{A}(k\mathbf{u}) + \mathbf{A}(c\mathbf{v}) && [\text{applying the rules of matrices}] \\
&= \lambda(k\mathbf{u}) + \lambda(c\mathbf{v}) && [\text{by the above } (*)] \\
&= \lambda(k\mathbf{u} + c\mathbf{v}) && [\text{factorizing}]
\end{aligned}$$

Since $\mathbf{A}(k\mathbf{u} + c\mathbf{v}) = \lambda(k\mathbf{u} + c\mathbf{v})$ so the matrix $\mathbf{A}$ scalar (λ) multiplies the vector $k\mathbf{u} + c\mathbf{v}$ so it is an eigenvector belonging to λ which means it is a member of the set S. By Proposition (3.7) we conclude that the set S is a subspace of $\mathbb{R}^n$.

■

This subspace S of Proposition (7.4)

$$S = \{\mathbf{O}\} \cup \left\{\mathbf{u} \,\middle|\, \mathbf{u} \text{ is an eigenvector belonging to } \lambda\right\}$$

is called an **eigenspace** of λ and is denoted by E_λ, that is $E_\lambda = S$.

For example, the eigenspace associated with Example 7.4 for the eigenvalue $\lambda_1 = 2$ is the eigenvector $\mathbf{u} = s\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and for $\lambda_2 = 3$ the eigenvector $\mathbf{v} = s\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ which are shown in Fig. 7.6.

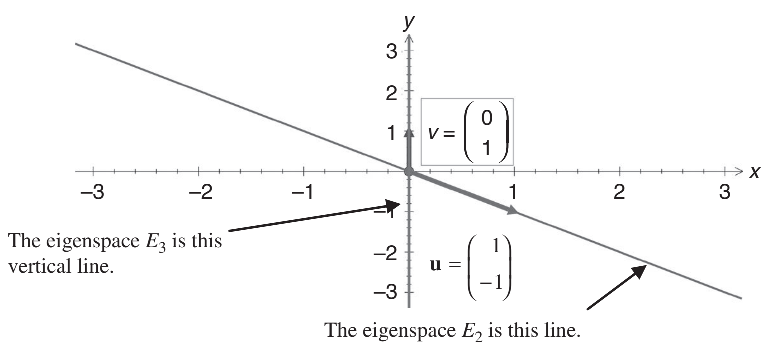


Figure 7.6

E_2 and E_3 denote the eigenspaces given by $\lambda = 2$ and $\lambda = 3$ respectively.

Note that the vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is a basis (axis) for the eigenspace E_3 and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is a basis (axis) for the eigenspace E_2. These *eigenvectors* are a basis (axis) for each eigenspace.

We can also use the numerical software MATLAB to find eigenvalues and eigenvectors: As an example, we'll find consider the matrix in Example 7.6: $\mathbf{A} = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 4 & 0 \\ 3 & 5 & -3 \end{pmatrix}$.

In MATLAB, we enter the matrix $\mathbf{A}$ by typing: `A=[1 0 4 ; 0 4 0 ; 3 5 -3]` where the semicolon denotes the start of the new row. We'll let the matrix containing the eigenvectors be called V and the matrix containing the eigenvalues as d. We then use the following MATLAB command. `[V,d]=eig(A,'nobalance')`. The 'nobalance' prevents MATLAB from normalizing the vector. The result of this command is:

```
V =
1.0000     0.6667   -1.0000
   0          0     -0.4500
0.5000    -1.0000   -0.7500
```

Reading down each column gives the eigenvectors.

```
d =
3    0   0
0   -5   0
0    0   4
```

The leading diagonal entries give the eigenvalues.

By reading the above MATLAB output, the eigenvectors are

$$\begin{pmatrix} 1 \\ 0 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 2/3 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 9/20 \\ 3/4 \end{pmatrix}$$

So the general eigenvectors are

$$\mathbf{u} = r\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \mathbf{v} = s\begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}, \mathbf{w} = t\begin{pmatrix} 20 \\ 9 \\ 15 \end{pmatrix}$$

where $r, s, t \neq 0$. Note that the eigenvector $\mathbf{u}$ is the eigenvector found in Example 7.6. You are asked to verify the other two by hand in Exercises 7.1.

Summary

The eigenvector $\mathbf{u}$ belonging to eigenvalue λ satisfies:

(7.1) $$\mathbf{Au} = \lambda\mathbf{u} \quad \text{(matrix } \mathbf{A} \text{ scalar multiplies eigenvector } \mathbf{u} \text{ by } \lambda)$$

The following equation is used to find the eigenvalues:

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

Eigenvectors $\mathbf{u}$ are found using $(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = \mathbf{O}$.

EXERCISES 7.1

(Brief solutions at end of book. Full solutions available at <http://www.oup.co.uk/companion/singh>.)

In this exercise you may check your numerical answers using MATLAB.

1. Find the eigenvalues and particular eigenvectors of the following matrices:

   (a) $\mathbf{A} = \begin{pmatrix} 7 & 3 \\ 0 & -4 \end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix} 5 & -2 \\ 4 & -1 \end{pmatrix}$ (c) $\mathbf{A} = \begin{pmatrix} -1 & 4 \\ 2 & 1 \end{pmatrix}$

2. Obtain the general eigenvectors of $\lambda_1 = 4$ and $\lambda_2 = -5$ for the matrix in Example 7.5.

3. Find the eigenvalues and eigenvectors of $\mathbf{A} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$. Plot the eigenspaces E_λ.

4. Find the eigenvalues and eigenvectors of $\mathbf{A} = \begin{pmatrix} 5 & -2 \\ 7 & -4 \end{pmatrix}$. State the effect of multiplying the eigenvector by the matrix $\mathbf{A}$. Plot the eigenspaces E_λ and write down a basis vector for each of the eigenspaces.

5. Let $\mathbf{A} = \begin{pmatrix} -2 & 8 \\ 5 & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} -4 & 16 \\ 10 & 2 \end{pmatrix}$.

   (a) Determine the eigenvalues of $\mathbf{A}$.

   (b) Determine the eigenvalues of $\mathbf{B}$.

   (c) State a relationship between the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$ and predict a general relationship.

6. Let $\mathbf{A}$ be a square matrix and $\mathbf{B} = r\mathbf{A}$, where r is a real number. Prove that if λ is the eigenvalue of matrix $\mathbf{A}$ then the eigenvalue of $\mathbf{B}$ is $r\lambda$.

7. Prove that the zero $n \times n$ matrix, $\mathbf{O} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$, only has *zero* eigenvalues.

   [Hint: Use the Proposition of chapter 6 which gives the determinant of a diagonal matrix.]

8. Determine the eigenvalues and eigenvectors of $\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. By using appropriate software plot the eigenspaces and write down a basis for each eigenspace.

SECTION 7.2 Properties of Eigenvalues and Eigenvectors

By the end of this section you will be able to

- determine eigenvalues and eigenvectors of particular matrices
- prove some properties of eigenvalues and eigenvectors
- apply the Cayley–Hamilton theorem

As we discovered in chapter 6, finding the inverse of a matrix can be a lengthy process. In this section we will show an easier way to find the inverse. We will also show that the Cayley–Hamilton theorem significantly reduces the workload in finding the powers of matrices.

It is worth reminding ourselves that eigenvalues and eigenvectors always come in pairs. You *cannot* have an eigenvalue on its own; it must have an associated eigenvector. But they are like chalk and cheese because an eigenvalue is a scalar and an eigenvector is a vector.

First we look at multiple eigenvalues and their corresponding eigenvectors.

7.2.1 Multiple eigenvalues

What is the characteristic equation of a square matrix $\mathbf{A}$?

The characteristic equation was defined in the last section 7.1 as:

(7.2) $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ where $\mathbf{A}$ is a square matrix, $\mathbf{I}$ is the identity matrix and λ is the eigenvalue which is a scalar. By expanding this we have

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det\left[\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} - \begin{pmatrix} \lambda & 0 & \cdots \\ 0 & \ddots & 0 \\ 0 & \cdots & \lambda \end{pmatrix}\right]$$

$$= \det\begin{pmatrix} a_{11} - \lambda & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} - \lambda \end{pmatrix} = 0 \quad \begin{bmatrix}\text{taking away } \lambda \text{ along} \\ \text{the leading diagonal}\end{bmatrix}$$

Evaluating the determinant of this matrix results in a polynomial equation of degree n. An example of a polynomial $p(\lambda)$ of degree 3 is

$$5\lambda^3 - 2\lambda^2 + 3\lambda + 69$$

An example of a polynomial $p(\lambda)$ of degree 4 is

$$7\lambda^4 - 5\lambda^3 - 0.1\lambda^2 + 34\lambda + 5$$

An example of a polynomial $p(\lambda)$ of degree n is

$$333\lambda^n + 2\lambda^{n-1} + \cdots + 6\lambda^2 + \lambda + 2.7\,1828$$

In general terms, we can write a polynomial $p(\lambda)$ of degree n as

$$p(\lambda) = c_n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$$

where the c's are the coefficients.

What is the difference between a characteristic polynomial and an equation?

A characteristic polynomial is an expression $p(\lambda)$ while an equation is $p\,(\lambda) = 0$.

How many roots does the equation $p(\lambda) = 0$ have?
The impressively titled 'Fundamental Theorem of Algebra' tells us that the polynomial equation $p(\lambda) = 0$ of degree n has *exactly* n roots. Don't worry if you have *not* heard of the 'Fundamental Theorem of Algebra'. It is a known theorem in algebra which claims that a polynomial equation of degree n has *exactly* n roots. For example, $x^2 - 2x + 1 = 0$ has exactly *two* roots because it is a polynomial equation of degree 2 (quadratic equation).

Other examples are:

Equation	Number of roots
$x^5 - 1 = 0$	5
$2x^{12} - 2x^3 + 1 = 0$	12
$-5x^{101} - x^{100} - \cdots - 1 = 0$	101

We will *not* prove the fundamental theorem of algebra but assume it is true. Thus by the fundamental theorem of algebra we conclude that

$$p(\lambda) = c_n\lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + \cdots + c_2\lambda^2 + c_1\lambda + c_0 = 0$$

has n eigenvalues (roots). There might be n distinct eigenvalues or they might be *repeated* such as

$$(\lambda - 1)^3 (\lambda - 2) = 0 \text{ which gives } \lambda_1 = 1, \lambda_2 = 1, \lambda_3 = 1 \text{ and } \lambda_4 = 2$$

Normally we write the first three roots in compact form as $\lambda_{1,2,3} = 1$ rather than as above.

We distinguish between simple and multiple eigenvalues in the next definition.

Definition (7.5). Let $\mathbf{A}$ be an n by n matrix and have the eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots$ and λ_n. If λ occurs *only once* then we say λ is a **simple eigenvalue**, otherwise it is called a **multiple eigenvalue**. If λ occurs m times where $m > 1$ then we say λ is an eigenvalue with **multiplicity** of m or λ has *multiplicity* m.

In the above equation $(\lambda - 1)^3 (\lambda - 2) = 0$ we have $\lambda_{1,2,3} = 1$ is an eigenvalue of multiplicity 3 and $\lambda_4 = 2$ is a simple eigenvalue.

Example 7.7

Determine the eigenvalues and eigenspaces of $\mathbf{A} = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Solution

What do we determine first?
The eigenvalues, because they produce the eigenvectors. Using the characteristic equation (7.2) $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ we have

(continued...)

$$\det(\mathbf{A}-\lambda\mathbf{I}) = \det\begin{pmatrix} \boxed{\begin{matrix} 2-\lambda \\ 0 \\ 0 \end{matrix}} & \begin{matrix} 1 \\ 2-\lambda \\ 0 \end{matrix} & \begin{matrix} 3 \\ 0 \\ 2-\lambda \end{matrix} \end{pmatrix}$$

$$= (2-\lambda)\det\begin{pmatrix} 2-\lambda & 0 \\ 0 & 2-\lambda \end{pmatrix} \qquad \begin{bmatrix}\text{expanding along} \\ \text{the first column}\end{bmatrix}$$

$$= (2-\lambda)\Big[(2-\lambda)(2-\lambda)-0\Big] = (2-\lambda)^3 = 0$$

Thus we *only* have one repeated eigenvalue $\lambda = \lambda_{1,2,3} = 2$. We say that $\lambda = 2$ has multiplicity 3.
How do we find the corresponding eigenvector?
By substituting this, $\lambda = 2$, into $(\mathbf{A}-\lambda\mathbf{I})\mathbf{u} = \mathbf{O}$ where $\mathbf{u}$ is the eigenvector:

$$(\mathbf{A}-2\mathbf{I})\,\mathbf{u} = \begin{pmatrix} 2-2 & 1 & 3 \\ 0 & 2-2 & 0 \\ 0 & 0 & 2-2 \end{pmatrix}\mathbf{u} = \mathbf{O}$$

Simplifying this and substituting unknowns x, y, z for the eigenvector $\mathbf{u}$ and zeros into the zero vector, $\mathbf{O}$, gives

$$\begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Note that this matrix is already in reduced row echelon form (rref). There is only one non-zero equation and three unknowns, therefore there are $3-1=2$ free variables (x and z, because none of the equations begin with these). The term 'free variables' was defined in chapter 1.

By expanding the first row we have $y+3z=0$, which gives $y=-3z$. Let $z=s$ where $s \neq 0$ then $y=-3s$. Clearly x can be any real number, that is $x=t$. Hence, we write the eigenvector ($x=t$, $y=-3s$ and $z=s$) in terms of two separate vectors which are a basis for the eigenspace E_2:

$$\mathbf{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t \\ -3s \\ s \end{pmatrix} = \begin{pmatrix} t \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -3s \\ s \end{pmatrix} = t\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 0 \\ -3 \\ 1 \end{pmatrix}$$

where s and t are *not both zero*. If s and t were *both* zero then $\mathbf{u}$ would be the zero vector but $\mathbf{u}$ is an eigenvector so $\neq \mathbf{O}$. This $\mathbf{u}$ is our general eigenvector and we can write our eigenspace as

$E_2 = \left\{ t\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 0 \\ -3 \\ 1 \end{pmatrix} \right\}$ and plot this in $\mathbb{R}^3$ as shown in Fig. 7.7.

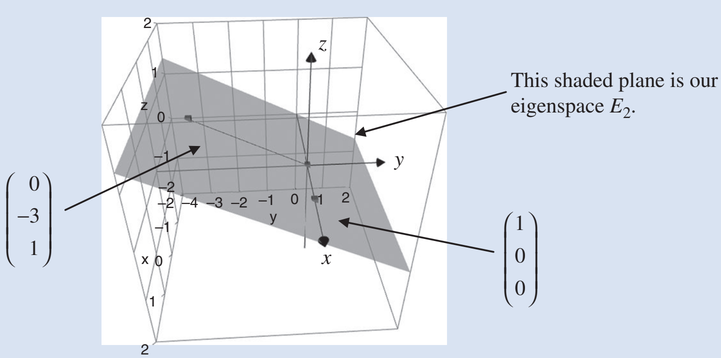


Figure 7.7

Note that we have a plane rather than just a line because we have two basis vectors (as shown above) which span a plane. A set of basis (axes) vectors B of the eigenspace E_2 are given by

$$B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 1 \end{pmatrix} \right\}$$

The matrix given in the above example is a type of matrix called a triangular matrix which was defined in the previous chapter. A smarter way to evaluate the eigenvalues of such matrices is described next.

7.2.2 Eigenvalues of diagonal and triangular matrices

In Exercises 7.1 you proved that the eigenvalues of the zero matrix, $\mathbf{O}$, are 0.

What sort matrix is the zero matrix?
A diagonal matrix.

What is a diagonal or triangular matrix?
By definition (6.17), a *triangular* matrix is an n by n matrix where *all* the entries to *one* side of the leading diagonal are zero.

By definition (6.18) we have a *diagonal* matrix is an n by n matrix where *all* the entries to *both* sides of the leading diagonal are zero.

The following are examples:

$$\mathbf{A} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}}_{\text{diagonal matrix}}, \mathbf{B} = \underbrace{\begin{pmatrix} 8 & -1 & 5 & 9 \\ 0 & -8 & 3 & -6 \\ 0 & 0 & 5 & 7 \\ 0 & 0 & 0 & 4 \end{pmatrix}}_{\text{triangular matrix}} \text{ and } \mathbf{C} = \underbrace{\begin{pmatrix} -9 & 0 & 0 & 0 \\ -8 & 3 & 0 & 0 \\ -2 & 4 & -1 & 0 \\ 1 & 0 & 5 & 10 \end{pmatrix}}_{\text{triangular matrix}}$$

$\mathbf{B}$ is actually called an *upper* triangular matrix and $\mathbf{C}$ a *lower* triangular matrix.

We prove that for a diagonal or triangular matrix the eigenvalues are given by the entries along the leading diagonal which goes from top left to bottom right of the matrix. For example, the above diagonal matrix $\mathbf{A}$ has eigenvalues 1, 2, 5 and 4.

What are the eigenvalues of the triangular matrices $\mathbf{B}$ and $\mathbf{C}$?
$\mathbf{B}$ has eigenvalues 8, $-$ 8, 5 and 4. $\mathbf{C}$ has eigenvalues -9, 3, -1 and 10.

Proposition (7.6). If an n by n matrix $\mathbf{A}$ is a diagonal or triangular matrix then the eigenvalues of $\mathbf{A}$ are the entries along the leading diagonal.

How do we prove this?
We apply Proposition (6.19) to the matrix $\mathbf{A}$: The determinant of a triangular or diagonal matrix is the product of the entries along the leading diagonal.

The proofs for the three different cases of upper, lower triangular and diagonal are very similar, so we only proof this for one case. In mathematical proof, we say ‘without loss of generality (WLOG)’ meaning that we prove it for one case, and the proof for the other case is identical.

Proof of (7.6).

Without loss of generality (WLOG), let $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & 0 & \ddots & \vdots \\ 0 & \vdots & 0 & a_{nn} \end{pmatrix}$ be an upper triangular matrix. The characteristic equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ is given by

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & 0 & \ddots & \vdots \\ 0 & \vdots & 0 & a_{nn} - \lambda \end{pmatrix}$$

$$\underbrace{=}_{\substack{\text{multiplying the leading} \\ \text{diagonal entries.} \\ \text{(by Proposition (6.19))}}} (a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda) = 0 \text{ implies } \lambda_1 = a_{11}, \lambda_2 = a_{22}, \ldots, \lambda_n = a_{nn}$$

The roots or eigenvalues of this equation are a_{11}, a_{22}, $a_{33}, \ldots$ and a_{nn} which are the leading diagonal entries of the given matrix $\mathbf{A}$. This completes our proof.

■

Example 7.8

Determine the eigenvalues of $\mathbf{A} = \begin{pmatrix} 5 & 3 & 5 & 8 \\ 0 & 9 & 7 & -9 \\ 0 & 0 & 6 & -5 \\ 0 & 0 & 0 & -17 \end{pmatrix}$.

Solution

What do you notice about matrix $\boldsymbol{A}$?

Matrix $\boldsymbol{A}$ is an (upper) triangular matrix, therefore we can apply the above Proposition (7.6) to find the eigenvalues of matrix $\boldsymbol{A}$.

What are the eigenvalues of $\boldsymbol{A}$?

The eigenvalues are the entries on the leading diagonal which runs from top left to bottom right. Thus the eigenvalues are

$$\lambda_1 = 5,\ \lambda_2 = 9,\ \lambda_3 = 6 \text{ and } \lambda_4 = -17$$

Note that the eigenvalues of the matrix $\mathbf{A}$ given in Example 7.7 are

$$\lambda_{1,\,2,\,3} = 2$$

because *all* the entries along the leading diagonal of the triangular matrix $\mathbf{A}$ are 2.

7.2.3 Properties of eigenvalues and eigenvectors

Proposition (7.7). A square matrix $\mathbf{A}$ is invertible (has an inverse) $\Leftrightarrow \lambda = 0$ is *not* an eigenvalue of the matrix $\mathbf{A}$.

Proof – Exercises 7.2.

■

This proposition means that if $\lambda = 0$ is an eigenvalue of matrix $\mathbf{A}$ then $\mathbf{A}$ has no inverse.

Proposition (7.8). Let $\mathbf{A}$ be a square matrix with eigenvector $\mathbf{u}$ belonging to eigenvalue λ.

(a) If m is a natural number then λ^m is an eigenvalue of the matrix $\mathbf{A}^m$ with the *same* eigenvector $\mathbf{u}$.

(b) If the matrix $\mathbf{A}$ is invertible (has an inverse) then the eigenvalue of the inverse matrix $\mathbf{A}^{-1}$ is $\dfrac{1}{\lambda} = \lambda^{-1}$ with the *same* eigenvector $\mathbf{u}$.

What does this proposition mean?

Matrix	Eigenvector	Eigenvalue
$\mathbf{A}$	$\mathbf{u}$	λ
(a) $\mathbf{A}^m$ (power)	$\mathbf{u}$	λ^m
(b) $\mathbf{A}^{-1}$ (inverse)	$\mathbf{u}$	λ^{-1}

We illustrate this for $\lambda = 2$ (the matrix $\mathbf{A}$ doubles the eigenvector $\mathbf{u}$) as shown in Fig. 7.8.

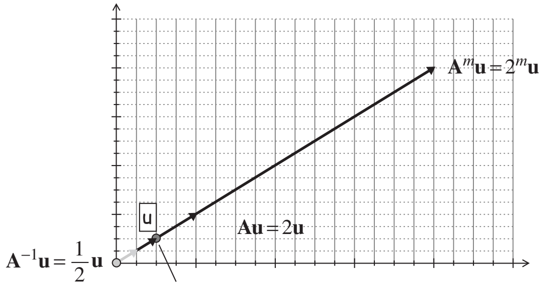


Figure 7.8

How do we prove proposition (a)?

By using mathematical induction. The three steps of mathematical induction are:

Step 1: Check the result for some base case $m = m_0$.

Step 2: Assume that the result is true for $m = k$.

Step 3: Prove the result for $m = k + 1$.

Proof.
Step 1: Using the definition of eigenvalues and eigenvectors (7.1) we have $\mathbf{Au} = \lambda\mathbf{u}$ which means the result holds for $m = 1$:

$$\mathbf{Au} = \lambda\mathbf{u} \tag{*}$$

Step 2: Assume that the result is true for $m = k$:

$$\mathbf{A}^k\mathbf{u} = \lambda^k\mathbf{u} \tag{†}$$

Step 3: Required to prove the case $m = k + 1$; that is, we need to prove

$$\mathbf{A}^{k+1}\mathbf{u} = \lambda^{k+1}\mathbf{u}$$

Expanding the left hand side:

$$\begin{aligned}\mathbf{A}^{k+1}\mathbf{u} &= \mathbf{A}\left(\mathbf{A}^k\mathbf{u}\right) \qquad \left[\text{writing } \mathbf{A}^{k+1} = \mathbf{A}\mathbf{A}^k\right] \\ &= \mathbf{A}\underbrace{\left(\lambda^k\mathbf{u}\right)}_{\text{by } (\dagger)} = \lambda^k(\mathbf{Au}) = \lambda^k\underbrace{(\lambda\mathbf{u})}_{\text{by } (*)} = \underbrace{\lambda^k\lambda}_{=\lambda^{k+1}}\mathbf{u} = \lambda^{k+1}\mathbf{u}\end{aligned}$$

Thus $\mathbf{A}^{k+1}\mathbf{u} = \lambda^{k+1}\mathbf{u}$, therefore by mathematical induction we have our result that λ^m is an eigenvalue of the matrix $\mathbf{A}^m$ with the eigenvector $\mathbf{u}$.

■

Proof of (b).
Using formula (7.1) we have $\mathbf{Au} = \lambda\mathbf{u}$. Left multiplying both sides of this by $\mathbf{A}^{-1}$ gives

$$\underbrace{\left(\mathbf{A}^{-1}\mathbf{A}\right)}_{=\mathbf{I}}\mathbf{u} = \mathbf{A}^{-1}(\lambda\mathbf{u}) = \lambda\mathbf{A}^{-1}\mathbf{u}$$

Remember, multiplying by the identity $\mathbf{I}$ keeps it the same, that is $\mathbf{Iu} = \mathbf{u}$, so we have

$$\mathbf{u} = \lambda\mathbf{A}^{-1}\mathbf{u}$$

Dividing both sides by λ gives

$$\frac{1}{\lambda}\mathbf{u} = \mathbf{A}^{-1}\mathbf{u} \text{ or writing this the other way, we have } \mathbf{A}^{-1}\mathbf{u} = \frac{1}{\lambda}\mathbf{u}$$

Since $\mathbf{A}^{-1}$ scalar multiplies $\mathbf{u}$ by $\frac{1}{\lambda}$ so the eigenvalue of $\mathbf{A}^{-1}$ is $\frac{1}{\lambda}$ with the eigenvector $\mathbf{u}$.

■

Example 7.9

Find the eigenvalues of $\mathbf{A}^7$ where $\mathbf{A} = \begin{pmatrix} 1 & 4 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix}$.

Solution

Because $\mathbf{A}$ is an upper triangular matrix, the eigenvalues are the entries on the leading diagonal, that is $\lambda_1 = 1, \lambda_2 = 2$ and $\lambda_3 = 3$. By the above Proposition (7.8) (a) we can see that the eigenvalues of $\mathbf{A}^7$ are

$$\lambda_1^7 = 1^7 = 1, \quad \lambda_2^7 = 2^7 = 128 \quad \text{and} \quad \lambda_3^7 = 3^7 = 2187$$

Proposition (7.9). Let $\mathbf{A}$ be any n by n matrix with eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_n$. We have:

(a) The determinant of the matrix $\mathbf{A}$ is given by $\det(\mathbf{A}) = \lambda_1 \times \lambda_2 \times \lambda_3 \times \cdots \times \lambda_n$.

(b) The trace of the matrix $\mathbf{A}$ is given by $tr(\mathbf{A}) = \lambda_1 + \lambda_2 + \lambda_3 + \cdots + \lambda_n$.

What is the trace of a matrix?

The trace (tr) of a matrix is the addition of all the entries on the leading diagonal:

$$tr\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = a_{11} + a_{22} + a_{33} + \cdots + a_{nn}$$

What does part (b) mean?

Adding *all* the eigenvalues of a matrix is equal to the *trace* of the matrix.

Can you see any use for part (b)?

We can use part (b) as a rough check to see if we have the correct eigenvalues.

Why is (b) a rough rather than an exact check for eigenvalues?

Because there are so many different ways of adding to the same value.

What does the first part (a) mean?

It means that we can find the determinant of a matrix by *multiplying* all the eigenvalues of that matrix.

How do we prove part (a)?

By using the characteristic equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$.

Proof of (a).

We are given that matrix $\mathbf{A}$ has eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots$ and λ_n.

What does this mean?

It means that these, $\lambda_1, \lambda_2, \ldots$ and λ_n, are roots of the characteristic equation, $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, which implies that we have

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda)\cdots(\lambda_n - \lambda) \quad (\dagger)$$

[For example, if the eigenvalues of a matrix $\mathbf{B}$ are $\lambda_1 = 2$, $\lambda_2 = 4$ and $\lambda_3 = 5$ then the characteristic equation would be given by $\det(\mathbf{B} - \lambda\mathbf{I}) = (2 - \lambda)(4 - \lambda)(5 - \lambda)$.]

If we substitute $\lambda = 0$ into $(\dagger)$ we get

$$\begin{aligned}\det(\mathbf{A}) &= (\lambda_1 - 0)(\lambda_2 - 0)(\lambda_3 - 0)\cdots(\lambda_n - 0)\\ &= (\lambda_1)(\lambda_2)(\lambda_3)\cdots(\lambda_n) = \lambda_1 \times \lambda_2 \times \lambda_3 \times \cdots \times \lambda_n\end{aligned}$$

This is our required result.

■

Proof of (b) – See website.

■

Summarizing this proposition we have:

Matrix A	$\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n$	Eigenvalues
Determinant of $\mathbf{A}$	$\lambda_1 \times \lambda_2 \times \lambda_3 \times \cdots \times \lambda_n$	Product of eigenvalues
Trace of $\mathbf{A}$	$\lambda_1 + \lambda_2 + \lambda_3 + \cdots + \lambda_n$	Addition of eigenvalues

The above proposition states that we can evaluate the determinant (or trace) of any n by n matrix by finding the eigenvalues and multiplying (or adding) them.

Example 7.10

Find the determinant and trace of $\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ given that the eigenvalues of $\mathbf{A}$ are 1, 4 and -1.

Solution

By the above Proposition (7.9) we have

$$\det(\mathbf{A}) = 1 \times 4 \times (-1) = -4 \quad \text{[multiplying the eigenvalues]}$$

$$Tr(\mathbf{A}) = 1 + 4 - 1 = 4 \quad \text{[adding \textit{all} eigenvalues]}$$

Remember, the trace of the matrix is found by adding all the entries in the leading diagonal

$$Tr(\mathbf{A}) = 1 + 1 + 2 = 4$$

Hence we have illustrated for matrix $\mathbf{A}$ that the trace of the matrix is the same as the sum of *the* eigenvalues.

Proposition (7.10). Let $\mathbf{A}$ be an n by n matrix with *distinct* eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_m$ and corresponding eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots, \mathbf{u}_m$ where $1 \leq m \leq n$. Then these eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots$ and $\mathbf{u}_m$ are *linearly independent*.

We need this proposition to prove the Cayley–Hamilton theorem which is given in the next subsection.

Proof – Exercises 7.2.

■

7.2.4 Cayley–Hamilton theorem

The biography of Arthur Cayley was given in section 1.4.1. Here we give a brief profile of Sir William Rowan Hamilton.

Figure 7.9

William Hamilton (Fig. 7.9) was born in Dublin, Ireland in 1805 and became one of the greatest Irish mathematicians. Initially he took an interest in languages, but during his early school days he found an affection for mathematics. At the age of 18 he entered Trinity College, Dublin and spent the rest of his life there. In 1827, Hamilton was appointed Professor of Astronomy at Trinity College, but he did not take much interest in astronomy and devoted all his time to mathematics. Hamilton is best known for his work on *quaternions* which is a vector space of four dimensions. In fact, in 1843 while he was walking along the local canal in Dublin with his wife he had a flash of inspiration and discovered the formula for quaternion multiplication.

At present there is plaque at the bridge stating this formula. He also invented the dot and cross product of vectors.

Throughout his life he had problems with alcohol and love. He fell in love with Catherine but due to unfortunate circumstances he ended up marrying Helen which he regretted for the rest of his life.

In general, we have *not* evaluated powers of matrices such as $\mathbf{A}^3, \mathbf{A}^4, \ldots$. This is because trying to determine $\mathbf{A}^n$ is a tedious task for almost all matrices. Finding $\mathbf{A}^2$ by $\mathbf{A} \times \mathbf{A}$ is simple enough for small size matrices but $\mathbf{A}^3$ is *not* so elementary. The Cayley–Hamilton theorem simplifies evaluation of $\mathbf{A}^n$ for a given matrix $\mathbf{A}$.

The statement of the Cayley–Hamilton theorem is straightforward.

Cayley–Hamilton (7.11). Every square matrix $\mathbf{A}$ is a root of the characteristic equation, that is $p(\mathbf{A}) = \mathbf{O}$ where p represents the characteristic polynomial.

Proof – This is a difficult proof and is on the book's website.

■

Example 7.11

Find the characteristic polynomial $p(\lambda)$ of the matrix $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 3 & 4 \end{pmatrix}$ and illustrate the Cayley–Hamilton theorem for this matrix $\mathbf{A}$.

Solution

We use $p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$ and substitute the given matrix $\mathbf{A}$ into this:

$$\begin{aligned} p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) &= \det\begin{pmatrix} 3-\lambda & 2 \\ 3 & 4-\lambda \end{pmatrix} && \left[\begin{array}{l}\text{taking away } \lambda \text{ along} \\ \text{the leading diagonal}\end{array}\right] \\ &= (3-\lambda)(4-\lambda) - 6 \\ &= 12 - 7\lambda + \lambda^2 - 6 = \lambda^2 - 7\lambda + 6 \end{aligned}$$

Thus the characteristic polynomial is $p(\lambda) = \lambda^2 - 7\lambda + 6$.

What is $p(\mathbf{A})$ equal to?

Substituting matrix $\mathbf{A}$ into this equation yields:

$$\begin{aligned} p(\mathbf{A}) &= \mathbf{A}^2 - 7\mathbf{A} + 6\mathbf{I} \\ &= \begin{pmatrix} 3 & 2 \\ 3 & 4 \end{pmatrix}^2 - 7\begin{pmatrix} 3 & 2 \\ 3 & 4 \end{pmatrix} + 6\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} && [\text{substituting } \mathbf{A} \text{ and } \mathbf{I}] \\ &= \begin{pmatrix} 3 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ 3 & 4 \end{pmatrix} - \begin{pmatrix} 21 & 14 \\ 21 & 28 \end{pmatrix} + \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix} && \left[\begin{array}{l}\text{carrying out scalar} \\ \text{multiplication}\end{array}\right] \\ &= \begin{pmatrix} 15 & 14 \\ 21 & 22 \end{pmatrix} - \begin{pmatrix} 21 & 14 \\ 21 & 28 \end{pmatrix} + \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 15-21+6 & 14-14+0 \\ 21-21+0 & 22-28+6 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \mathbf{O} \end{aligned}$$

Thus $p(\mathbf{A}) = \mathbf{O}$.

What does this mean?

This means that matrix $\mathbf{A}$ satisfies its characteristic equation $p(\mathbf{A}) = \mathbf{A}^2 - 7\mathbf{A} + 6\mathbf{I} = \mathbf{O}$ which is what the Cayley–Hamilton theorem states.

We can use this result $\mathbf{A}^2 - 7\mathbf{A} + 6\mathbf{I} = \mathbf{O}$ to find higher powers of matrix $\mathbf{A}$.

We can apply the Cayley–Hamilton theorem to find powers of matrices. For example, if

$$\mathbf{A} = \begin{pmatrix} 0.9 & 0.2 \\ 0.3 & 0.6 \end{pmatrix} \text{ then } \mathbf{A}^{100} = \begin{pmatrix} 29.4 & 13.5 \\ 20.2 & 9.2 \end{pmatrix} \text{ (1dp)}$$

$\mathbf{A}^{100}$ is *not* evaluated by multiplying 100 copies of matrix $\mathbf{A}$. The evaluation of $\mathbf{A}^{100}$ is found by using the eigenvalues of matrix $\mathbf{A}$. Calculating $\mathbf{A}^{100}$ is a very tedious task which can be significantly reduced by applying the Cayley–Hamilton theorem. The powers of a matrix can be written in terms of a polynomial of lower degree in matrix $\mathbf{A}$. This is illustrated in the next example.

Example 7.12

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

Solution

Working out the determinant of $\mathbf{A} - \lambda\mathbf{I}$ gives

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}) = \det\begin{pmatrix} -2-\lambda & -4 \\ 1 & 3-\lambda \end{pmatrix} &= (-2-\lambda)(3-\lambda) + 4 \\ &= (\lambda+2)(\lambda-3) + 4 \\ &= \lambda^2 - \lambda - 2 = p(\lambda) \end{aligned}$$

By the Cayley–Hamilton theorem we have $p(\mathbf{A}) = \mathbf{A}^2 - \mathbf{A} - 2\mathbf{I} = \mathbf{O}$. Transposing this gives

$$\begin{aligned} \mathbf{A}^2 &= \mathbf{A} + 2\mathbf{I} && (\dagger) \\ \mathbf{A}^3 &= \mathbf{A}^2 + 2\mathbf{A} && [\text{multiplying by } \mathbf{A}] \\ &= \underbrace{\mathbf{A} + 2\mathbf{I}}_{=\mathbf{A}^2 \text{ by above } (\dagger)} + 2\mathbf{A} = 3\mathbf{A} + 2\mathbf{I} \end{aligned}$$

Multiplying the last equation $\mathbf{A}^3 = 3\mathbf{A} + 2\mathbf{I}$ by $\mathbf{A}$ gives

$$\begin{aligned} \mathbf{A}^4 &= 3\mathbf{A}^2 + 2\mathbf{A} \\ &= 3\underbrace{\left[\mathbf{A} + 2\mathbf{I}\right]}_{=\mathbf{A}^2 \text{ by above } (\dagger)} + 2\mathbf{A} = 5\mathbf{A} + 6\mathbf{I} \end{aligned}$$

Thus $\mathbf{A}^4 = 5\mathbf{A} + 6\mathbf{I} = 5\begin{pmatrix} -2 & -4 \\ 1 & 3 \end{pmatrix} + 6\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -4 & -20 \\ 5 & 21 \end{pmatrix}$.

Note that we can evaluate $\mathbf{A}^4$ without working out $\mathbf{A}^3$ and $\mathbf{A}^2$ because $\mathbf{A}^4 = 5\mathbf{A} + 6\mathbf{I}$ is a polynomial of degree 1 in matrix $\mathbf{A}$.

In MATLAB we can find the characteristic polynomial of a matrix $\mathbf{A}$ by using the command poly(A). We can also use the Cayley–Hamilton theorem to find the *inverse* of a matrix as the next example demonstrates. Remember, determining the inverse of a 3 by 3 or larger matrix is a laborious job. However, by applying the Cayley–Hamilton theorem we find it becomes much simpler.

Example 7.13

Determine $\mathbf{A}^{-1}$ where $\mathbf{A} = \begin{pmatrix} 10 & 15 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 6 \end{pmatrix}$ given that the characteristic polynomial of this matrix is:

$$p(\lambda) = \lambda^3 - 20\lambda^2 + 94\lambda - 60$$

Solution

By the Cayley–Hamilton theorem (7.11) we have

$$p(\mathbf{A}) = \mathbf{A}^3 - 20\mathbf{A}^2 + 94\mathbf{A} - 60\mathbf{I} = \mathbf{O} \qquad \left[\text{because } p(\lambda) = \lambda^3 - 20\lambda^2 + 94\lambda - 60\right]$$

(continued...)

Adding $60\mathbf{I}$ to both sides gives

$$\mathbf{A}^3 - 20\mathbf{A}^2 + 94\mathbf{A} = 60\mathbf{I}$$
$$\mathbf{A}\left(\mathbf{A}^2 - 20\mathbf{A} + 94\mathbf{I}\right) = 60\mathbf{I} \qquad \left[\text{factorizing out the matrix } \mathbf{A}\right]$$

By the definition of the inverse matrix we have $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$, which means that dividing both sides by 60 in the above gives

$$\mathbf{A}\underbrace{\frac{1}{60}\left(\mathbf{A}^2 - 20\mathbf{A} + 94\mathbf{I}\right)}_{=\mathbf{A}^{-1}} = \mathbf{I}$$

Hence $\mathbf{A}^{-1} = \frac{1}{60}\left(\mathbf{A}^2 - 20\mathbf{A} + 94\mathbf{I}\right)$. Evaluating the components of this yields

$$\mathbf{A}^2 = \begin{pmatrix} 10 & 15 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 6 \end{pmatrix}\begin{pmatrix} 10 & 15 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 6 \end{pmatrix} = \begin{pmatrix} 130 & 210 & 0 \\ 28 & 46 & 0 \\ 60 & 105 & 36 \end{pmatrix}$$

$$20\mathbf{A} = 20\begin{pmatrix} 10 & 15 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 6 \end{pmatrix} = \begin{pmatrix} 200 & 300 & 0 \\ 40 & 80 & 0 \\ 60 & 120 & 120 \end{pmatrix} \quad \text{and} \quad 94\mathbf{I} = \begin{pmatrix} 94 & 0 & 0 \\ 0 & 94 & 0 \\ 0 & 0 & 94 \end{pmatrix}$$

Putting these into $\mathbf{A}^{-1} = \frac{1}{60}\left(\mathbf{A}^2 - 20\mathbf{A} + 94\mathbf{I}\right)$ gives

$$\mathbf{A}^{-1} = \frac{1}{60}\left(\mathbf{A}^2 - 20\mathbf{A} + 94\mathbf{I}\right) = \frac{1}{60}\left[\begin{pmatrix} 130 & 210 & 0 \\ 28 & 46 & 0 \\ 60 & 105 & 36 \end{pmatrix} - \begin{pmatrix} 200 & 300 & 0 \\ 40 & 80 & 0 \\ 60 & 120 & 120 \end{pmatrix} + \begin{pmatrix} 94 & 0 & 0 \\ 0 & 94 & 0 \\ 0 & 0 & 94 \end{pmatrix}\right]$$

$$= \frac{1}{60}\begin{pmatrix} 130 - 200 + 94 & 210 - 300 + 0 & 0 - 0 + 0 \\ 28 - 40 + 0 & 46 - 80 + 94 & 0 - 0 + 0 \\ 60 - 60 + 0 & 105 - 120 + 0 & 36 - 120 + 94 \end{pmatrix}$$

$$= \frac{1}{60}\begin{pmatrix} 24 & -90 & 0 \\ -12 & 60 & 0 \\ 0 & -15 & 10 \end{pmatrix}$$

Finding $\mathbf{A}^{-1}$ still involves a fair amount of calculation but is generally easier than finding the cofactors of each entry.

Summary

Proposition (7.6). If $\mathbf{A}$ is a diagonal or triangular matrix then the eigenvalues of $\mathbf{A}$ are the entries along the leading diagonal.

Let $\mathbf{A}$ be a square matrix with an eigenvalue λ and eigenvector $\mathbf{u}$ belonging to λ.

Then we have the following:

Matrix		Eigenvector	Eigenvalue
(7.8) (a)	$\mathbf{A}^m$	$\mathbf{u}$	λ^m
(7.8) (b)	$\mathbf{A}^{-1}$	$\mathbf{u}$	λ^{-1}

Cayley–Hamilton theorem (7.11). Every square matrix $\mathbf{A}$ is a root of the characteristic equation, that is $p(\mathbf{A}) = \mathbf{O}$.

EXERCISES 7.2

(Brief solutions at end of book. Full solutions available at <http://www.oup.co.uk/companion/singh>.)

In this exercise check your numerical answers using any appropriate software.

1. Determine the eigenvalues and eigenvectors of $\mathbf{A} = \begin{pmatrix} 5 & 2 \\ -2 & 1 \end{pmatrix}$ and plot the eigenspace E_λ and write down a set of basis vectors for E_λ.
2. Determine the eigenvalues and eigenvectors of $\mathbf{A} = \begin{pmatrix} -12 & 7 \\ -7 & 2 \end{pmatrix}$ and plot the eigenspace E_λ and write down a set of basis vectors for E_λ.
3. Let $\mathbf{A}$ be a 2 by 2 matrix. Show that the characteristic polynomial $p(\lambda)$ is given by
$$p(\lambda) = \lambda^2 - tr(\mathbf{A})\lambda + \det(\mathbf{A})$$
where $tr(\mathbf{A})$ is the trace of the matrix $\mathbf{A}$.
4. Let $\mathbf{A}$ be a 2 by 2 matrix with $tr(\mathbf{A}) = 2a$ and $\det(\mathbf{A}) = a^2$ where a is a real number. Show that this matrix has an eigenvalue of a with *multiplicity* of 2.
5. Determine the eigenvalues and eigenvectors of $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 7 & 9 & 1 \end{pmatrix}$. By using appropriate software or otherwise, plot the eigenspace E_λ and write down a set of basis vectors for this eigenspace.
6. Determine the eigenvalues and eigenvectors of the following matrices and write down a set of basis vectors for the eigenspace E_λ.
(a) $\mathbf{A} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ (b) $\mathbf{B} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 1 \\ 0 & 0 & 9 \end{pmatrix}$ (c) $\mathbf{C} = \begin{pmatrix} -2 & 0 & 0 \\ 2 & -2 & 0 \\ 4 & 10 & -2 \end{pmatrix}$
7. Determine the eigenvalues and eigenvectors of the following matrices:
(a) $\mathbf{A} = \begin{pmatrix} 7 & 0 & 2 & 3 \\ 0 & 7 & 4 & 6 \\ 0 & 0 & 5 & -3 \\ 0 & 0 & 0 & 5 \end{pmatrix}$ (b) $\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$ (c) $\mathbf{C} = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$

8. Prove Proposition (7.7), which says a matrix is invertible $\Leftrightarrow \lambda = 0$ is *not* an eigenvalue.
9. Let $\mathbf{A}$ be a square matrix and λ be an eigenvalue with the corresponding eigenvector $\mathbf{u}$. Prove that the eigenvalue λ is unique for the eigenvector $\mathbf{u}$.
10. Let $\mathbf{A}$ be a 2 by 2 matrix. By using the result of question 3, the characteristic polynomial $p(\lambda)$ is given by

$$p(\lambda) = \lambda^2 - tr(\mathbf{A})\lambda + \det(\mathbf{A})$$

Show that if $\left[tr(\mathbf{A})\right]^2 > 4\det(\mathbf{A})$ then $\mathbf{A}$ has distinct eigenvalues.
State under what conditions we have equal and complex eigenvalues.

11. For each of the following matrices

(a) $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 5 & 2 & 3 & 0 \\ 9 & 8 & 1 & 4 \end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix} -1 & 3 & 4 & 7 \\ 0 & 6 & -3 & 5 \\ 0 & 0 & -8 & 9 \\ 0 & 0 & 0 & 3 \end{pmatrix}$ (c) $\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

Determine
(i) eigenvalues of $\mathbf{A}$ (ii) eigenvalues of $\mathbf{A}^5$ (iii) eigenvalues of $\mathbf{A}^{-1}$
(iv) $\det(\mathbf{A})$ (v) $tr(\mathbf{A})$

12. Find the characteristic polynomial $p(\lambda)$ of the matrix $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}$ and determine the inverse of this matrix by using the Cayley–Hamilton theorem.
13. Let $\mathbf{A} = \begin{pmatrix} 6 & 5 \\ 3 & 4 \end{pmatrix}$. By using the Cayley–Hamilton theorem determine $\mathbf{A}^2$ and $\mathbf{A}^3$.
14. Let $\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ and the characteristic polynomial of this matrix be given by

$$p(\lambda) = \lambda^3 - 4\lambda^2 - \lambda + 4$$

Determine expressions for $\mathbf{A}^{-1}$ and $\mathbf{A}^4$ in terms of the matrices $\mathbf{A}$, $\mathbf{I}$ and $\mathbf{A}^2$.

15. Prove Proposition (7.10).
16. Prove that the eigenvalues of the transposed matrix, $\mathbf{A}^T$, are exactly the eigenvalues of the matrix $\mathbf{A}$.

SECTION 7.3 Diagonalization

By the end of this section you will be able to

- understand what is meant by similar matrices
- diagonalize a matrix
- find powers of matrices

If $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ *then how do we find* $\mathbf{A}^{100}$*?*

We could apply the Cayley–Hamilton theorem of the last section. However, using Cayley–Hamilton requires us to find a formula for $\mathbf{A}^2, \mathbf{A}^3, \ldots, \mathbf{A}^{99}$ and then use this to determine $\mathbf{A}^{100}$. Clearly this is a very laborious way of finding $\mathbf{A}^{100}$. In this section, we examine an easier method to find powers of matrices such as $\mathbf{A}^{100}$. We factorize a given matrix $\mathbf{A}$ into three matrices, one of which is a diagonal matrix. It is a lot simpler to deal with a diagonal matrix because *any* matrix calculation is easier with diagonal matrices. For example, it is easy to find the inverse of a diagonal matrix. Also, proving results about diagonal matrices is simpler than proving results about general matrices.

In this section we aim to answer the following question:

For a square matrix $\mathbf{A}$*, is there an invertible matrix* $\mathbf{P}$ *(has an inverse) such that* $\mathbf{P}^{-1}\mathbf{AP}$ *produces a diagonal matrix?*

The goal of this section is to convert an n by n matrix into a diagonal matrix. The process of converting any n by n matrix into a diagonal matrix is called **diagonalization** (Fig. 7.10).

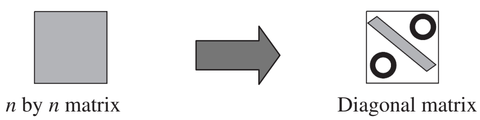


Figure 7.10

What has this got to do with eigenvalues and eigenvectors?

This section explores how to diagonalize a matrix, for which we need to find the eigenvalues and the corresponding eigenvectors. These eigenvalues are the leading diagonal entries in the diagonal matrix.

First, we define **similar** matrices and some of their properties.

7.3.1 Similar matrices

Example 7.14

Let $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$ and $\mathbf{P} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$. Determine $\mathbf{P}^{-1}\mathbf{AP}$.

Solution

We first find the inverse of the matrix $\mathbf{P}$, denoted $\mathbf{P}^{-1}$:

$$\mathbf{P}^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}^{-1} = \frac{1}{0-(-1)}\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \qquad \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\right]$$

Carrying out the matrix multiplication

$$\begin{aligned} \mathbf{P}^{-1}\mathbf{AP} &= \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 2 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \qquad \left[\begin{array}{l}\text{first multiplying the} \\ \text{left and centre matrices.}\end{array}\right] \end{aligned}$$

Definition (7.12). A square matrix $\mathbf{B}$ is **similar** to a matrix $\mathbf{A}$ if there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{B}$.

In Example 7.14, the final matrix $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ is *similar* to matrix $\mathbf{A}$ because $\mathbf{P}^{-1}\mathbf{AP} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$.

Similar matrices have the following properties (equivalence relation):

Proposition (7.13). Let $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ be square matrices. Then

(a) Matrix $\mathbf{A}$ is similar to matrix $\mathbf{A}$.

(b) If matrix $\mathbf{B}$ is similar to matrix $\mathbf{A}$ then the other way round is also true, that is matrix $\mathbf{A}$ is similar to matrix $\mathbf{B}$.

(c) If matrix $\mathbf{A}$ is similar to $\mathbf{B}$ and $\mathbf{B}$ is similar to matrix $\mathbf{C}$ then matrix $\mathbf{A}$ is similar to matrix $\mathbf{C}$.

Proof – Exercises 7.3.

■

By property (b) we can say matrices $\mathbf{A}$ and $\mathbf{B}$ are similar. The following proposition gives another important property of similar matrices.

Proposition (7.14). Let $\mathbf{A}$ and $\mathbf{B}$ be similar matrices. The eigenvalues of these matrices are identical.

Proof.

We are given that matrices $\mathbf{A}$ and $\mathbf{B}$ are similar.

What does this mean?

By (7.12) there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{B}$. Let $\det(\mathbf{B} - \lambda\mathbf{I})$ be the characteristic polynomial for the matrix $\mathbf{B}$ and $\det(\mathbf{A} - \lambda\mathbf{I})$ be the characteristic polynomial for the matrix $\mathbf{A}$.

Required to prove that the polynomials given by these determinants are equal:

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

We have

$$\begin{aligned}
\det(\mathbf{B} - \lambda\mathbf{I}) &= \det\left(\mathbf{P}^{-1}\mathbf{AP} - \lambda\mathbf{I}\right) && \left[\text{replacing } \mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}\right] \\
&= \det\left(\mathbf{P}^{-1}\mathbf{AP} - \lambda\mathbf{P}^{-1}\mathbf{P}\right) && \left[\text{substituting } \mathbf{I} = \mathbf{P}^{-1}\mathbf{P}\right] \\
&= \det\left(\mathbf{P}^{-1}\mathbf{AP} - \mathbf{P}^{-1}\lambda\mathbf{P}\right) && \left[\text{moving the scalar } \lambda\right] \\
&= \det\left(\mathbf{P}^{-1}\mathbf{AP} - \mathbf{P}^{-1}\lambda\mathbf{IP}\right) && \left[\text{rewriting } \lambda\mathbf{P} = \lambda\mathbf{IP}\right] \\
&= \det\left[\mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{P}\right] && \left[\text{factorizing out } \mathbf{P}^{-1} \text{ and } \mathbf{P}\right] \\
&= \det\left(\mathbf{P}^{-1}\right)\det(\mathbf{A} - \lambda\mathbf{I})\det(\mathbf{P}) && \left[\text{using } \det(\mathbf{ABC}) = \det(\mathbf{A})\det(\mathbf{B})\det(\mathbf{C})\right] \\
&= \det(\mathbf{A} - \lambda\mathbf{I}) && \left[\text{because } \det\left(\mathbf{P}^{-1}\right)\det(\mathbf{P}) = 1\right]
\end{aligned}$$

Similar matrices $\mathbf{A}$ and $\mathbf{B}$ have the *same* eigenvalues λ because λ satisfies the equation:

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

■

Proposition (7.14) means that similar matrices carry out an identical transformation – they scalar multiply each eigenvector by the same scalar λ.

7.3.2 Introduction to diagonalization

What do we mean when we say that a matrix is diagonalizable?

Definition (7.15). An n by n matrix $\mathbf{A}$ is diagonalizable if it is *similar* to a diagonal matrix $\mathbf{D}$.

The matrix $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$ in the above Example 7.14 is diagonalizable because the matrix $\mathbf{P} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$ gives $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$ where $\mathbf{D} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ is a *diagonal* matrix.

What does this definition (7.15) mean?

$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$ which means that we can convert a matrix $\mathbf{A}$ into a diagonal matrix by left multiplying by $\mathbf{P}^{-1}$ and right multiplying by $\mathbf{P}$. We say the matrix $\mathbf{P}$ **diagonalizes** the matrix $\mathbf{A}$.

Why diagonalize a matrix?

In general, a diagonal matrix is easier to work with because if you are multiplying, solving a system of equations or finding eigenvalues, it is always preferable to have a diagonal matrix. The diagonal matrix significantly reduces the amount of numerical calculations needed to find powers of a matrix, for example. We will show later in this section that it is easy to evaluate the 100th power of a diagonal matrix. Remember, we aim to convert a given matrix into a diagonal matrix.

Next, we give a trivial example of a matrix which is diagonalizable.

Example 7.15

Show that a diagonal matrix $\mathbf{A} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$ is diagonalizable.

Solution

We need to find a matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$ where $\mathbf{D}$ is a diagonal matrix.

What matrix $\mathbf{P}$ should we consider?

The identity matrix $\mathbf{P} = \mathbf{I}$ because the given matrix $\mathbf{A}$ is already a diagonal matrix:

$$\mathbf{I}^{-1}\mathbf{AI} = \mathbf{A}$$

Thus the given matrix $\mathbf{A}$ is diagonalizable.

Example 7.16

Show that any diagonal matrix is diagonalizable.

Solution

Let $\mathbf{D}$ be any diagonal matrix. The diagonalizing matrix is the identity matrix $\mathbf{P} = \mathbf{I}$ because $\mathbf{I}^{-1}\mathbf{D}\mathbf{I} = \mathbf{D}$.

How do we know which matrices are diagonalizable?

The next theorem states a test for establishing whether a matrix is diagonalizable or not.

Theorem (7.16). An n by n matrix $\mathbf{A}$ is diagonalizable $\Leftrightarrow$ it has n *linearly independent* eigenvectors.

Proof – Exercises 7.3.

■

We must have n linearly *independent* eigenvectors for the matrix to be *diagonalizable*. The procedure for diagonalizing an n by n matrix $\mathbf{A}$ can be derived from the proof of (7.16). In a nutshell it is the following procedure:

1. Using $\det(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = \mathbf{O}$ for each eigenvalue $\lambda_1, \lambda_2, \ldots, \lambda_n$, we find the eigenvectors belonging to these $\lambda_1, \lambda_2, \ldots, \lambda_n$. Call these eigenvectors $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \ldots$ and $\mathbf{p}_n$. If matrix $\mathbf{A}$ does *not* have n linearly independent eigenvectors then it is *not* diagonalizable.
2. Form the matrix $\mathbf{P}$ by having these eigenvectors $\mathbf{p}_1, \mathbf{p}_2, \ldots$ and $\mathbf{p}_n$, as its columns. That is matrix $\mathbf{P}$ contains the eigenvectors:

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 & \cdots & \mathbf{p}_n \end{pmatrix}$$

3. The diagonal matrix $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ will have the eigenvalues $\lambda_1, \lambda_2, \ldots$ and λ_n of $\mathbf{A}$ along its leading diagonal, that is

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & & O \\ & \ddots & \\ O & & \lambda_n \end{pmatrix} \quad \text{eigenvalues of } \mathbf{A}$$

The eigenvector $\mathbf{p}_j$ belongs to the eigenvalue λ_j.

4. It is good practice to check that matrices $\mathbf{P}$ and $\mathbf{D}$ actually work. For matrices of size *greater than* 2 by 2, the evaluation of the inverse matrix $\mathbf{P}^{-1}$ can be lengthy so prone to calculation errors. To bypass this evaluation of the inverse, simply check that $\mathbf{PD} = \mathbf{AP}$.

Why?
Because left multiplying the above $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$ by $\mathbf{P}$ gives

$$\begin{aligned}\mathbf{PD} &= \mathbf{P}\left(\mathbf{P}^{-1}\mathbf{AP}\right) \\ &= \left(\mathbf{PP}^{-1}\right)\mathbf{AP} = \mathbf{IAP} = \mathbf{AP} \qquad \left[\text{remember } \mathbf{PP}^{-1} = \mathbf{I}\right]\end{aligned}$$

Hence it is enough to check that $\mathbf{PD} = \mathbf{AP}$.

Note that since matrices $\mathbf{A}$ and $\mathbf{D}$ are similar, they have the same eigenvalues.

The matrix $\mathbf{P}$ is called the **eigenvector matrix** and the diagonal matrix $\mathbf{D}$ the **eigenvalue matrix**.

Example 7.17

Determine the eigenvector matrix $\mathbf{P}$ which diagonalizes the matrix $\mathbf{A} = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ given that the eigenvalues of this matrix are $\lambda_1 = -1$ and $\lambda_2 = 5$ with corresponding eigenvectors $\mathbf{u} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ respectively.

Solution

Steps 1 and 2.
We have been given the eigenvectors, $\mathbf{u}$ and $\mathbf{v}$, of matrix $\mathbf{A}$ so the eigenvector matrix $\mathbf{P}$ is

$$\mathbf{P} = (\mathbf{u} \quad \mathbf{v}) = \begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix}$$

Step 3.
Since matrices $\mathbf{A}$ and $\mathbf{D}$ are similar, so our diagonal matrix $\mathbf{D}$ contains the eigenvalues of $\mathbf{A}$, $\lambda_1 = -1$ and $\lambda_2 = 5$:

$$\mathbf{D} = \begin{pmatrix} -1 & 0 \\ 0 & 5 \end{pmatrix} \longleftarrow \text{eigenvalues of } \mathbf{A}$$

Step 4.
We need to confirm that this matrix $\mathbf{P}$ does indeed diagonalize the given matrix $\mathbf{A}$.
How?
By checking that $\mathbf{PD} = \mathbf{AP}$:

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

$$\mathbf{AP} = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ -1 & 5 \end{pmatrix}$$

Thus the eigenvector matrix $\mathbf{P}$ does indeed diagonalize the given matrix $\mathbf{A}$.

Notice that the eigenvalues, $\lambda_1 = -1$ and $\lambda_2 = 5$, of the given matrix $\mathbf{A}$ (and $\mathbf{D}$) are entries along the leading diagonal in the matrix $\mathbf{D}$. They occur in the order λ_1 and then λ_2

because the matrix $\mathbf{P}$ is created by $\mathbf{P} = (\mathbf{u} \quad \mathbf{v})$ where $\mathbf{u}$ is the eigenvector belonging to λ_1 and $\mathbf{v}$ is the eigenvector belonging to the other eigenvalue λ_2.

What would be the diagonal matrix if the matrix $\boldsymbol{P}$ was created by swapping $\boldsymbol{u}$ and $\boldsymbol{v}$, that is $\mathbf{P} = (\mathbf{v} \;\; \mathbf{u})$?

Our diagonal matrix would be $\mathbf{D}' = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$. [See Exercises 7.3].

Note that the diagonal eigenvalue matrix $\mathbf{D}$ contains the eigenvalues, the eigenvector matrix $\mathbf{P}$ contains the corresponding eigenvectors.

Example 7.18

Show that the matrix $\mathbf{A} = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 2 & 5 \\ 0 & 0 & 2 \end{pmatrix}$ with eigenvalues and eigenvectors given by

$$\lambda_1 = 1, \mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}; \lambda_2 = 2, \mathbf{v} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \text{ and } \lambda_3 = 2, \mathbf{w} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$$

is *not* diagonalizable.

Solution

Why can't we diagonalize the given matrix $\boldsymbol{A}$?

According to the procedure outlined above we *cannot* diagonlize matrix $\mathbf{A}$ if the three eigenvectors are linearly dependent. (Linear dependency occurs when we can write one vector in terms of the others.)

Note that vectors $\mathbf{v}$ and $\mathbf{w}$ are linearly *dependent* because $\mathbf{v} = -\mathbf{w}$ or $\mathbf{v} + \mathbf{w} = \mathbf{O}$. Thus the matrix $\mathbf{A}$ is *not* diagonalizable.

7.3.3 Distinct eigenvalues

Next, we state a proposition that if an n by n matrix has n *distinct* eigenvalues then the eigenvectors belonging to these are linearly *independent*.

Proposition (7.17). Let $\mathbf{A}$ be an n by n matrix with n *distinct* eigenvalues, $\lambda_1, \lambda_2, \ldots$ and λ_n with the corresponding eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots$ and $\mathbf{u}_n$. These eigenvectors are *linearly independent*.

Proof.

By Proposition (7.10):

Let $\mathbf{A}$ have *distinct* eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ with corresponding eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ where $1 \le m \le n$. Then these eigenvectors are *linearly independent*.

Using (7.10) with $n = m$ gives us our required result.

■

Proposition (7.18). If an n by n matrix $\mathbf{A}$ has n *distinct* eigenvalues then the matrix $\mathbf{A}$ is diagonalizable.

Proof.
The n by n matrix $\mathbf{A}$ has n distinct eigenvalues, therefore by the above proposition (7.17) the corresponding n eigenvectors are linearly independent. Thus by the above result (7.16):

An n by n matrix $\mathbf{A}$ is diagonalizable $\Leftrightarrow$ it has n *independent* eigenvectors.

We conclude that the matrix $\mathbf{A}$ is diagonalizable.

■

Example 7.19

Determine whether the matrix $\mathbf{A} = \begin{pmatrix} 1 & -6 & 2 \\ 0 & 4 & 25 \\ 0 & 0 & 9 \end{pmatrix}$ is diagonalizable.

Solution

What type of matrix is $\mathbf{A}$*?*

Matrix $\mathbf{A}$ is an upper triangular matrix, therefore by Proposition (7.6): If an n by n matrix $\mathbf{A}$ is a diagonal or triangular matrix then the eigenvalues of $\mathbf{A}$ are the entries along the leading diagonal.

The eigenvalues are the entries along the leading diagonal; $\lambda_1 = 1$, $\lambda_2 = 4$ and $\lambda_3 = 9$.

How do we determine whether the given matrix $\mathbf{A}$ *is diagonalizable or not?*

$\mathbf{A}$ is a 3 by 3 matrix and it has three distinct eigenvalues; $\lambda_1 = 1$, $\lambda_2 = 4$ and $\lambda_3 = 9$,

Therefore by Proposition (7.18) the matrix $\mathbf{A}$ is diagonalizable.

An n by n matrix might be diagonalizable even if it does *not* have n distinct eigenvalues.

For example, the identity matrix $\mathbf{I}$ is diagonalizable even though it has n copies of the same eigenvalue 1. (If we have n distinct eigenvalues for an n by n matrix then we are guaranteed that the matrix is diagonalizable.)

The process of finding the eigenvalues, eigenvectors and the matrix $\mathbf{P}$ which diagonalizes a given matrix can be a tedious task if you are not given the eigenvalues and eigenvectors; you have to go through the whole process.

Example 7.20

For the matrix $\mathbf{A} = \begin{pmatrix} 1 & -6 & 2 \\ 0 & 4 & 25 \\ 0 & 0 & 9 \end{pmatrix}$, determine the eigenvector matrix $\mathbf{P}$ which diagonalizes matrix $\mathbf{A}$ given that the eigenvalues of $\mathbf{A}$ are $\lambda_1 = 1$, $\lambda_2 = 4$ and $\lambda_3 = 9$ with corresponding eigenvectors $\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} -7 \\ 10 \\ 2 \end{pmatrix}$ respectively.

Check that $\mathbf{P}$ does indeed diagonalize the given matrix $\mathbf{A}$.

Solution

Step 1 and Step 2:

We have been given the eigenvalues and corresponding eigenvectors of the matrix $\mathbf{A}$.

(continued...)

What is our eigenvector matrix $\mathbf{P}$ equal to?
Eigenvector matrix $\mathbf{P}$ contains the eigenvectors:

$$\mathbf{P} = \begin{pmatrix} \mathbf{u} & \mathbf{v} & \mathbf{w} \end{pmatrix} = \begin{pmatrix} 1 & -2 & -7 \\ 0 & 1 & 10 \\ 0 & 0 & 2 \end{pmatrix} \left[\text{given } \mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \mathbf{w} = \begin{pmatrix} -7 \\ 10 \\ 2 \end{pmatrix} \right]$$

Step 3:
The diagonal eigenvalue matrix $\mathbf{D}$ has the eigenvalues $\lambda_1 = 1$, $\lambda_2 = 4$ and $\lambda_3 = 9$ along the leading diagonal, that is

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

eigenvalues of matrix $\mathbf{A}$.

Step 4:
Checking that we have the correct $\mathbf{P}$ and $\mathbf{D}$ matrices by showing $\mathbf{PD} = \mathbf{AP}$:

$$\mathbf{PD} = \begin{pmatrix} 1 & -2 & -7 \\ 0 & 1 & 10 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix} = \begin{pmatrix} 1 & -8 & -63 \\ 0 & 4 & 90 \\ 0 & 0 & 18 \end{pmatrix}$$

$$\mathbf{AP} = \begin{pmatrix} 1 & -6 & 2 \\ 0 & 4 & 25 \\ 0 & 0 & 9 \end{pmatrix} \begin{pmatrix} 1 & -2 & -7 \\ 0 & 1 & 10 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -8 & -63 \\ 0 & 4 & 90 \\ 0 & 0 & 18 \end{pmatrix}$$

Hence, this confirms that the matrix $\mathbf{P}$ does indeed diagonalize the given matrix $\mathbf{A}$.

7.3.4 Powers of matrices

We discussed the above diagonalization process so that we can find powers of matrices. For example, to find $\mathbf{A}^{100}$ is a difficult task.

What does diagonalization have to do with powers of matrices?
If $\mathbf{A}$ is a square matrix which is diagonalizable, so that there exists a matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$ where $\mathbf{D}$ is a diagonal matrix then

$$\mathbf{A}^m = \mathbf{P}\mathbf{D}^m\mathbf{P}^{-1} \text{ where } m \text{ is any real number}$$

We will show this result in the next proposition.

In the meantime, using this formula, we can find the inverse of matrix $\mathbf{A}$ by substituting $m = -1$. To find $\mathbf{A}^{100}$ it is much easier to use this formula, $\mathbf{A}^m = \mathbf{P}\mathbf{D}^m\mathbf{P}^{-1}$, rather than multiplying a 100 copies of the matrix $\mathbf{A}$. We can use this formula to find $\mathbf{A}^m$ if we first determine $\mathbf{D}^m$.

How?
The matrix $\mathbf{D}^m$ is simply a diagonal matrix with its leading diagonal entries raised to the power m, that is

$$\text{If } \mathbf{D} = \begin{pmatrix} d_1 & & O \\ & \ddots & \\ O & & d_n \end{pmatrix} \quad \text{then } \mathbf{D}^m = \begin{pmatrix} {d_1}^m & & O \\ & \ddots & \\ O & & d_n^m \end{pmatrix}$$

You are asked to show this result in Exercises 7.3. $\mathbf{A}^m$ might be hard to calculate but because $\mathbf{D}$ is a diagonal matrix, $\mathbf{D}^m$ is simply $\mathbf{D}$ with the entries on the leading diagonal raised to the power m.

The eigenvalue matrix $\mathbf{D}$ consists of the eigenvalues on the leading diagonal, therefore $\mathbf{D}^m$ has these *eigenvalues* to the *power m* on the leading diagonal.

Proposition (7.19). If an n by n matrix $\mathbf{A}$ is diagonalizable with $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$ where $\mathbf{D}$ is a diagonal matrix then

$$\mathbf{A}^m = \mathbf{PD}^m\mathbf{P}^{-1}$$

How do we prove this result?

By using mathematical induction.

What is the mathematical induction procedure?

Step 1: Check for $m = 1$.

Step 2: Assume that the result is true for $m = k$.

Step 3: Prove it for $m = k + 1$.

Proof.

Step 1: Check for $m = 1$, that is we need to show $\mathbf{PDP}^{-1} = \mathbf{A}$.

We have $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$. Left multiplying this by matrix $\mathbf{P}$ and right multiplying it by $\mathbf{P}^{-1}$:

$$\begin{aligned} \mathbf{PDP}^{-1} &= \mathbf{P}\underbrace{\left(\mathbf{P}^{-1}\mathbf{AP}\right)}_{=\mathbf{D}}\mathbf{P}^{-1} \\ &= \left(\mathbf{PP}^{-1}\right)\mathbf{A}\left(\mathbf{PP}^{-1}\right) = \mathbf{IAI} = \mathbf{A} \end{aligned}$$

Thus we have our result for $m = 1$ which is $\mathbf{A} = \mathbf{PDP}^{-1}$. This means that we can factorize matrix $\mathbf{A}$ into three matrices $\mathbf{P}$, $\mathbf{D}$ and $\mathbf{P}^{-1}$.

Step 2: Assume that the result is true for $m = k$, that is

$$\mathbf{A}^k = \mathbf{PD}^k\mathbf{P}^{-1} \qquad (*)$$

Step 3: We need to prove the result for $m = k + 1$, that is we need to prove

$$\mathbf{A}^{k+1} = \mathbf{PD}^{k+1}\mathbf{P}^{-1}$$

Starting with the left hand side of this we have

$$\begin{aligned}
\mathbf{A}^{k+1} &= \mathbf{A}^k\mathbf{A} && \left[\text{applying the rules of indices}\right]\\
&= \underbrace{\left(\mathbf{PD}^k\mathbf{P}^{-1}\right)}_{=\mathbf{A}^k \text{ by } (*)}\underbrace{\left(\mathbf{PDP}^{-1}\right)}_{=\mathbf{A} \text{ by Step 1}}\\
&= \mathbf{PD}^k\left(\mathbf{P}^{-1}\mathbf{P}\right)\mathbf{DP}^{-1} && \left[\text{using the rules of matrices } (\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})\right]\\
&= \mathbf{P}\underbrace{\mathbf{D}^k(\mathbf{I})\,\mathbf{D}}_{=\mathbf{D}^k\mathbf{D}}\mathbf{P}^{-1} && \left[\text{because } \mathbf{P}^{-1}\mathbf{P} = \mathbf{I}\right]\\
&= \mathbf{PD}^k\mathbf{DP}^{-1} = \mathbf{PD}^{k+1}\mathbf{P}^{-1} && \left[\text{because } \mathbf{D}^k\mathbf{D} = \mathbf{D}^{k+1}\right]
\end{aligned}$$

This is our required result. Hence by mathematical induction we have $\mathbf{A}^m = \mathbf{PD}^m\mathbf{P}^{-1}$.

■

In general, to find $\mathbf{A}^m$ we have to multiply m copies of the matrix $\mathbf{A}$ which is a laborious task. It is much easier if we factorize $\mathbf{A}^m$ into $\mathbf{A}^m = \mathbf{PD}^m\mathbf{P}^{-1}$ (even though we need to find $\mathbf{P}$, $\mathbf{P}^{-1}$ and $\mathbf{D}$, which is no easy task in itself). This formula means that if you want to evaluate $\mathbf{A}^m$ without working out lower powers then using the diagonal matrix is more efficient than the Cayley–Hamilton method discussed in the previous section.

Note that in the above subsection when we diagonalized a matrix we could avoid the calculation of $\mathbf{P}^{-1}\mathbf{AP}$.

Why?

Because we check that $\mathbf{PD} = \mathbf{AP}$. This means that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$ is the diagonal eigenvalue matrix. However, in evaluating $\mathbf{A}^m$ we need to find $\mathbf{P}^{-1}$ because $\mathbf{A}^m = \mathbf{PD}^m\mathbf{P}^{-1}$.

Example 7.21

Let $\mathbf{A} = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 2 & 5 \\ 0 & 0 & 3 \end{pmatrix}$. Find $\mathbf{A}^5$ given that $\mathbf{P} = \begin{pmatrix} 1 & -2 & -7 \\ 0 & 1 & 10 \\ 0 & 0 & 2 \end{pmatrix}$ and $\mathbf{P}^{-1} = \frac{1}{2}\begin{pmatrix} 2 & 4 & -13 \\ 0 & 2 & -10 \\ 0 & 0 & 1 \end{pmatrix}$.

Solution

The diagonal matrix $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Because $\mathbf{A}$ is an upper triangular matrix, its eigenvalues are the entries on the leading diagonal of $\mathbf{A}$, that is $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = 3$.

How do we find $\mathbf{A}^5$?

By applying the above result (7.19), $\mathbf{A}^m$ factorizes into $\mathbf{A}^m = \mathbf{PD}^m\mathbf{P}^{-1}$ with $m = 5$:

$$\mathbf{A}^5 = \mathbf{PD}^5\mathbf{P}^{-1}$$

Substituting matrices $\mathbf{P}$, $\mathbf{D}$ and $\mathbf{P}^{-1}$ into this $\mathbf{A}^5 = \mathbf{P}\mathbf{D}^5\mathbf{P}^{-1}$ gives

$$\mathbf{A}^5 = \mathbf{P}\mathbf{D}^5\mathbf{P}^{-1}$$

$$= \begin{pmatrix} 1 & -2 & -7 \\ 0 & 1 & 10 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}^5 \frac{1}{2}\begin{pmatrix} 2 & 4 & -13 \\ 0 & 2 & -10 \\ 0 & 0 & 1 \end{pmatrix}$$

eigenvalues of $\mathbf{A}$.

$$= \frac{1}{2}\begin{pmatrix} 1 & -2 & -7 \\ 0 & 1 & 10 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1^5 & 0 & 0 \\ 0 & 2^5 & 0 \\ 0 & 0 & 3^5 \end{pmatrix} \begin{pmatrix} 2 & 4 & -13 \\ 0 & 2 & -10 \\ 0 & 0 & 1 \end{pmatrix} \quad \left[\text{taking } \tfrac{1}{2} \text{ to the front and e.values to the power 5}\right]$$

$$= \frac{1}{2}\begin{pmatrix} 1 & -2 & -7 \\ 0 & 1 & 10 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 32 & 0 \\ 0 & 0 & 243 \end{pmatrix} \begin{pmatrix} 2 & 4 & -13 \\ 0 & 2 & -10 \\ 0 & 0 & 1 \end{pmatrix} \quad \left[\text{replacing the e.values } 1^5 = 1, 2^5 = 32 \text{ and } 3^5 = 243\right]$$

$$= \frac{1}{2}\begin{pmatrix} 1 & -64 & -1701 \\ 0 & 32 & 2430 \\ 0 & 0 & 486 \end{pmatrix} \begin{pmatrix} 2 & 4 & -13 \\ 0 & 2 & -10 \\ 0 & 0 & 1 \end{pmatrix} \quad \left[\text{multiplying the first two matrices on the left}\right]$$

$$= \frac{1}{2}\begin{pmatrix} 2 & -124 & -1074 \\ 0 & 64 & 2110 \\ 0 & 0 & 486 \end{pmatrix} = \begin{pmatrix} 1 & -62 & -537 \\ 0 & 32 & 1055 \\ 0 & 0 & 243 \end{pmatrix} \quad \left[\text{multiplying by the scalar } 1/2\right]$$

You may wish to check this final result by using appropriate software.

Note that in the above example the diagonal matrix $\mathbf{D}$ has eigenvalues 1, 2 and 3 on the leading diagonal, and $\mathbf{D}^5$ has 1^5, 2^5 and 3^5 on the leading diagonal. 1^5, 2^5 and 3^5 are the eigenvalues of $\mathbf{A}^5$.

7.3.5 Application of powers of matrices

Matrix powers are particularly useful in **Markov chains** – these are based on matrices whose entries are probabilities. Many real life systems have an element of uncertainty which develops over time, and this can be explained through Markov chains.

Example 7.22

The transition matrix $\mathbf{T}$ below gives the percentage of people involved in accidents who were either injured (I) or were killed (K) on urban (U) and rural (R) roads. The entries in the first column of matrix $\mathbf{T}$ indicate that 60% of road injuries on urban roads and 40% on rural roads. The second column of $\mathbf{T}$ represents 50% of road accident deaths occured on urban roads and 50% on rural roads. Out of a sample of 100 accidents this year the number of accidents on urban roads was 90 and rural roads 10 and this is represented by the vector $\mathbf{x}$.

$$\mathbf{T} = \begin{array}{c} \begin{array}{cc} \mathrm{I} & \mathrm{K} \end{array} \\ \begin{pmatrix} 0.6 & 0.5 \\ 0.4 & 0.5 \end{pmatrix} \end{array} \begin{array}{c} \mathrm{U} \\ \mathrm{R} \end{array} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} 90 \\ 10 \end{pmatrix}$$

The vector $\mathbf{x}_n$ given by $\mathbf{x}_n = \mathbf{T}^n\mathbf{x}$ gives us the number of accidents on urban and rural roads out of a sample of 100 accidents after n number of years. Determine to 2sf
(i) $\mathbf{x}_n$ for $n = 10$. (ii) $\mathbf{x}_n$ as $n \to \infty$

(continued...)

(This gives us our long term number of injuries and deaths on urban and rural roads out of a sample of 100 accidents.) For a Markov chain, we are interested in the long-term behaviour of $\mathbf{x}_n$.

Solution

(i) This means that we need to find $\mathbf{x}_n = \mathbf{T}^n\mathbf{x}$ when $n = 10$, that is $\mathbf{x}_{10} = \mathbf{T}^{10}\mathbf{x}$. To evaluate $\mathbf{T}^{10}$, we diagonalize the matrix $\mathbf{T}$ by finding the eigenvalue $\mathbf{D}$ and eigenvector $\mathbf{P}$ matrices. Verify that the eigenvalues and the corresponding eigenvectors are given by

$$\lambda_1 = 1,\ \mathbf{u} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \text{ and } \lambda_2 = 0.1, \mathbf{v} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Our eigenvector matrix $\mathbf{P} = (\mathbf{u}\ \ \mathbf{v}) = \begin{pmatrix} 5 & -1 \\ 4 & 1 \end{pmatrix}$ and eigenvalue matrix $\mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 0.1 \end{pmatrix}$.
By result (7.19) $\mathbf{A}^m = \mathbf{PD}^m\mathbf{P}^{-1}$ with $m = 10$ and $\mathbf{A} = \mathbf{T}$ we have

$$\mathbf{T}^{10} = \mathbf{PD}^{10}\mathbf{P}^{-1} \qquad (\dagger)$$

To evaluate $\mathbf{T}^{10}$ we need to find $\mathbf{P}^{-1}$, which is given by

$$\mathbf{P}^{-1} = \frac{1}{9}\begin{pmatrix} 1 & 1 \\ -4 & 5 \end{pmatrix}$$

Substituting $\mathbf{P} = \begin{pmatrix} 5 & -1 \\ 4 & 1 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 0.1 \end{pmatrix}$ and $\mathbf{P}^{-1} = \frac{1}{9}\begin{pmatrix} 1 & 1 \\ -4 & 5 \end{pmatrix}$ into ($\dagger$) gives

$$\begin{aligned}
\mathbf{T}^{10} = \mathbf{PD}^{10}\mathbf{P}^{-1} &= \begin{pmatrix} 5 & -1 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0.1 \end{pmatrix}^{10}\frac{1}{9}\begin{pmatrix} 1 & 1 \\ -4 & 5 \end{pmatrix} \\
&= \frac{1}{9}\begin{pmatrix} 5 & -1 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -4 & 5 \end{pmatrix} \quad \left[\begin{array}{l}\text{because } 1^{10} = 1 \text{ and} \\ 0.1^{10} = 1 \times 10^{-10} = 0\text{(3dp)}\end{array}\right] \\
&= \frac{1}{9}\begin{pmatrix} 5 & 0 \\ 4 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -4 & 5 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 5 & 5 \\ 4 & 4 \end{pmatrix}
\end{aligned}$$

Substituting $\mathbf{T}^{10} = \frac{1}{9}\begin{pmatrix} 5 & 5 \\ 4 & 4 \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} 90 \\ 10 \end{pmatrix}$ into $\mathbf{x}_{10} = \mathbf{T}^{10}\mathbf{x}$ gives

$$\mathbf{x}_{10} = \frac{1}{9}\begin{pmatrix} 5 & 5 \\ 4 & 4 \end{pmatrix}\begin{pmatrix} 90 \\ 10 \end{pmatrix} \underbrace{=}_{\substack{\text{taking out a} \\ \text{factor of 10}}} \frac{10}{9}\begin{pmatrix} 5 & 5 \\ 4 & 4 \end{pmatrix}\begin{pmatrix} 9 \\ 1 \end{pmatrix} = \frac{10}{9}\begin{pmatrix} 50 \\ 40 \end{pmatrix} = \begin{pmatrix} 55.5 \\ 44.4 \end{pmatrix} \begin{array}{l} \text{U} \\ \text{R} \end{array}$$

This means that after 10 years the number of people likely to be injured or killed on an urban road is 56 (2sf) and on a rural road is 44 (2sf) out of a sample of 100 accidents.

(ii) We have $\mathbf{D}^n = \begin{pmatrix} 1 & 0 \\ 0 & 0.1 \end{pmatrix}^n = \begin{pmatrix} 1^n & 0 \\ 0 & 0.1^n \end{pmatrix}$.

How does this change as $n \to \infty$?

As $n \to \infty$ we have $1^n \to 1$ and $0.1^n \to 0$. This means that $\mathbf{D}^n = \mathbf{D}^{10}$ correct to 2sf which gives the same results as part (i).

Summary

If $\mathbf{A}$ is diagonalizable then we can convert $\mathbf{A}$ into a diagonal matrix $\mathbf{D}$.
If an n by n matrix $\mathbf{A}$ is diagonalizable with $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$ where $\mathbf{D}$ is a diagonal matrix then

$$\mathbf{A}^m = \mathbf{P}\mathbf{D}^m\mathbf{P}^{-1}$$

EXERCISES 7.3

(Brief solutions at end of book. Full solutions available at <http://www.oup.co.uk/companion/singh>.)
In this exercise check your numerical answers using MATLAB.

1. For the following matrices find:

   (i) The eigenvalues and corresponding eigenvectors.

   (ii) Eigenvector matrix $\mathbf{P}$ and eigenvalue matrix $\mathbf{D}$.

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

2. (i) For the matrices in question 1 find $\mathbf{A}^5$.

   (ii) For the matrix in question 1 part (c) find $\mathbf{A}^{-1/2}$.

3. For the following matrices find:

   (i) The eigenvalues and corresponding eigenvectors.

   (ii) Matrices $\mathbf{P}$ and $\mathbf{D}$ where $\mathbf{P}$ is the invertible (non-singular) matrix and $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is the diagonal matrix. To find $\mathbf{P}^{-1}$ you may use MATLAB.

   (iii) Determine $\mathbf{A}^4$ in each case by using the results of parts (i) and (ii).

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

4. For the following matrices determine whether they are diagonalizable.

   (a) $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix} -1 & 2 & 3 \\ 0 & 2 & 5 \\ 0 & 0 & 8 \end{pmatrix}$ (c) $\mathbf{A} = \begin{pmatrix} \sqrt{2} & 0 & 0 & 0 \\ 1 & \sqrt{3} & 0 & 0 \\ 6 & 7 & 1/2 & 0 \\ 2 & 9 & 7 & -5 \end{pmatrix}$

5. In Example 7.17 take $\mathbf{P} = (\mathbf{v} \quad \mathbf{u})$ and determine the diagonal matrix $\mathbf{D}$ given by $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ where $\mathbf{A} = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$.

What do you notice about your diagonal matrix D?

6. Let $\mathbf{A}$ be a 3 by 3 matrix with the following eigenvalues and eigenvectors:

$$\lambda_1 = -2,\ \mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \lambda_2 = -5,\ \mathbf{v} = \begin{pmatrix} 5 \\ 4 \\ 0 \end{pmatrix} \text{ and } \lambda_3 = -1,\ \mathbf{w} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Is the matrix $\mathbf{A}$ diagonalizable? If it is then find the diagonal eigenvalue matrix $\mathbf{D}$ which is similar to the matrix $\mathbf{A}$ and also determine the invertible matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$.

Find $\mathbf{A}^3$. [Note that you do *not* need to know the elements of matrix $\mathbf{A}$.]

7. Show that the following matrices are *not* diagonalizable:

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

8. Let $\mathbf{A} = \begin{pmatrix} -4 & 2 \\ -9 & 5 \end{pmatrix}$. Determine (i) $\mathbf{A}^{11}$ (ii) $\mathbf{A}^{-1}$

9. Prove that if $\mathbf{D}$ is a diagonal matrix then the matrix $\mathbf{D}^m$ is simply a diagonal matrix with its leading diagonal entries raised to the power m.

10. Prove Proposition (7.13).

11. Prove that if $\mathbf{A}$ is diagonalizable then the transpose of $\mathbf{A}$, that is $\mathbf{A}^T$, is also diagonalizable.

12. In a differential equations course, the matrix $\exp(\mathbf{A}t)$ is defined as

$$\exp(\mathbf{A}t) = \mathbf{I} + \mathbf{A}t + \mathbf{A}^2\frac{t^2}{2!} + \mathbf{A}^3\frac{t^3}{3!} + \mathbf{A}^4\frac{t^4}{4!} + \cdots$$

Let $\mathbf{A} = \begin{pmatrix} 3 & 5 \\ 0 & 2 \end{pmatrix}$ and find an expression for $\exp(\mathbf{A}t)$ up to and including the term t^4 by diagonalizing matrix $\mathbf{A}$.

13. Let $\mathbf{F} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ [$\mathbf{F}$ is known as the **Fibonacci matrix**]. Evaluate the matrices $\mathbf{P}$, $\mathbf{P}^{-1}$ and $\mathbf{D}$, where $\mathbf{P}$ is an invertible matrix and $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is a diagonal matrix.

14. Let $\mathbf{A}$ be a 2 by 2 matrix with $\left[tr(\mathbf{A})\right]^2 > 4\det(\mathbf{A})$ where tr represents the trace of the matrix. Show that $\mathbf{A}$ is diagonalizable.

15. Let $\mathbf{A}$ be an invertible matrix which is diagonalizable. Prove that $\mathbf{A}^{-1}$ is also diagonalizable.

16. Prove that if $\mathbf{A}$ is diagonalizable then $\mathbf{A}^m$ (where $m \in \mathbb{N}$) is diagonalizable.

17. Let $\mathbf{A}$ and $\mathbf{B}$ be invertible matrices. Prove that $\mathbf{AB}$ is similar to $\mathbf{BA}$.

18. If matrices $\mathbf{A}$ and $\mathbf{B}$ are similar, prove that

(i) $tr(\mathbf{A}) = tr(\mathbf{B})$ where tr is trace.

(ii) $\det(\mathbf{A}) = \det(\mathbf{B})$.

19. Let **A** be a diagonal matrix such that the modulus of each eigenvalue is less than 1. Evaluate the matrix $\mathbf{A}^m$ as $m \to \infty$.

[You may assume that if $|x| < 1$ then $\lim_{m \to \infty} (x^m) = 0$.]

20. Let **A** be a diagonalizable matrix with eigenvalues $\lambda_1, \lambda_2, \ldots$ and λ_n. Prove that the eigenvalues of $\mathbf{A}^m$ are $(\lambda_1)^m, (\lambda_2)^m, \ldots$ and $(\lambda_n)^m$.

21. Prove Theorem (7.16).

SECTION 7.4 Diagonalization of Symmetric Matrices

By the end of this section you will be able to

- prove properties of symmetric matrices
- orthogonally diagonalize symmetric matrices

In this section we continue the diagonalization process. Diagonalization was described in the previous section – we found a matrix **P** which diagonalized a given matrix; this allowed us to find the matrix **D**:

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} \text{ where } \mathbf{D} \text{ is a diagonal matrix.}$$

Left multiplying this by **P** and right multiplying by $\mathbf{P}^{-1}$ gives the factorization of matrix **A**:

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$$

Eigenvector matrix **P** contains the eigenvectors of **A**, and **D** contains the eigenvalues of **A**.

From this we deduced (result (7.19)) that the powers of matrix **A** can be found by factorizing $\mathbf{A}^m$ into three matrices:

$$\mathbf{A}^m = \mathbf{P}\mathbf{D}^m\mathbf{P}^{-1}$$

If you want to find $\mathbf{A}^{10}$ then $\mathbf{A}^{10} = \mathbf{P}\mathbf{D}^{10}\mathbf{P}^{-1}$. This can still be a tedious task even for a small size matrix such as 3 by 3.

Why?

Because we need to find the inverse matrix $\mathbf{P}^{-1}$, which will involve putting the matrix **P** into reduced row echelon form or using cofactors. Either way, a cumbersome task.

Is there a type of matrix for which we can easily find the inverse?
Yes, the orthogonal matrix $\mathbf{Q}$, described in chapter 4, for which $\mathbf{Q}^{-1} = \mathbf{Q}^T$.

What is an orthogonal matrix?
It's a square matrix whose columns are orthonormal (perpendicular unit) vectors.

In this section we aim to find a diagonalizing matrix $\mathbf{Q}$ which is an *orthogonal* matrix. Eigenvector matrix $\mathbf{Q}$ is the diagonalizing matrix which is made up by writing its columns as the eigenvectors of the given matrix $\mathbf{A}$.

However, when we find eigenvectors, they are usually *not* perpendicular (orthogonal) to each other. In this section we obtain eigenvectors which are perpendicular and normalized. We aim to get *orthonormal* (perpendicular unit) eigenvectors as columns of the diagonalizing matrix. Once we have achieved unit perpendicular eigenvectors as columns of the diagonalizing matrix then we will find working with the diagonal matrix even easier than the previous section. The columns of $\mathbf{Q}$ are also an orthonormal *basis* for the eigenspace E_λ. Remember, orthonormal bases (axes) are one of the simplest bases to work with.

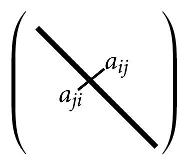
We cannot *guarantee* that the diagonalizing matrix will be an orthogonal matrix. However, in this section we will show that if the given matrix is *symmetric* then we can always find an orthogonal diagonalizing matrix.

7.4.1 Symmetric matrices

Can you recall what a symmetric matrix is?
A square matrix $\mathbf{A}$ is a symmetric matrix if $\mathbf{A}^T = \mathbf{A}$ ($\mathbf{A}$ transpose equals $\mathbf{A}$).

Examples are

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 3 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & -1 & \sqrt{5} \\ -1 & 2 & \pi \\ \sqrt{5} & \pi & 3 \end{pmatrix}, \quad \mathbf{D} =$$



What do you notice?
We get a reflection of the entries by placing a mirror on the leading diagonal as highlighted.

Why are symmetric matrices important?
We will show later in this section that *all* symmetric matrices are *diagonalizable* by an orthogonal matrix. This is *not* the case for non-symmetric matrices.

Example 7.23

Let $\mathbf{A} = \begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}$. Diagonalize matrix $\mathbf{A}$.

Solution

The characteristic equation is given by

$$\det(\mathbf{A}-\lambda\mathbf{I}) = \det\begin{pmatrix} -3-\lambda & 4 \\ 4 & 3-\lambda \end{pmatrix} = \lambda^2 - 25 = 0 \text{ which gives } \lambda_1 = 5 \text{ and } \lambda_2 = -5$$

The corresponding eigenvectors are $\mathbf{u} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ for $\lambda_1 = 5$ and $\mathbf{v} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ for $\lambda_2 = -5$.

The eigenvectors $\mathbf{u}$ and $\mathbf{v}$ are linearly independent (*not* multiples of each other) therefore

$$\mathbf{P} = (\mathbf{u} \quad \mathbf{v}) = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \text{ and } \mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix}$$

The eigenvector matrix $\mathbf{P} = (\mathbf{u} \quad \mathbf{v})$ contains the eigenvectors of $\mathbf{A}$ and the eigenvalue matrix $\mathbf{D}$ contains the eigenvalues of $\mathbf{A}$.

Remember, matrices $\mathbf{A}$ and $\mathbf{D}$ are similar, so they have the same eigenvalues.

What do you notice about the eigenvectors $\mathbf{u} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ *and* $\mathbf{v} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$*?*

The inner (dot) product of eigenvectors $\mathbf{u}$ and $\mathbf{v}$ is zero:

$$\mathbf{u}\cdot\mathbf{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix} = (1\times 2) + (2\times(-1)) = 0$$

What does $\mathbf{u}\cdot\mathbf{v} = 0$ *mean?*

Eigenvectors $\mathbf{u}$ and $\mathbf{v}$ are *orthogonal* which means that they are perpendicular to each other. (See Fig. 7.11 overleaf.) We can normalize these eigenvectors (that is make their length 1).

How?

By dividing by its length : (2.16) $\widehat{\mathbf{u}} = \frac{1}{\|\mathbf{u}\|}\mathbf{u}$ where $\widehat{\mathbf{u}}$ is the normalized (unit) vector and $\|\mathbf{u}\|$ is the norm (length) of $\mathbf{u}$.

For the eigenvector $\boldsymbol{u}$ *what is length* $\|\mathbf{u}\|$ *equal to?*

$$\|\mathbf{u}\|^2 = \mathbf{u}\cdot\mathbf{u} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 1^2 + 2^2 = 5 \quad \text{[by Pythagoras]}$$

We have $\|\mathbf{u}\|^2 = 5$, therefore taking the square root gives $\|\mathbf{u}\| = \sqrt{5}$. Thus the normalized eigenvector $\widehat{\mathbf{u}} = \dfrac{1}{\|\mathbf{u}\|}\mathbf{u} = \dfrac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Similarly, the other normalized eigenvector $\widehat{\mathbf{v}} = \dfrac{1}{\|\mathbf{v}\|}\mathbf{v} = \dfrac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

Note that $\widehat{\mathbf{u}}$ and $\widehat{\mathbf{v}}$ are unit eigenvectors, which means that they have norm (length) of 1. Plotting these in $\mathbb{R}^2$ is shown in Fig. 7.11.

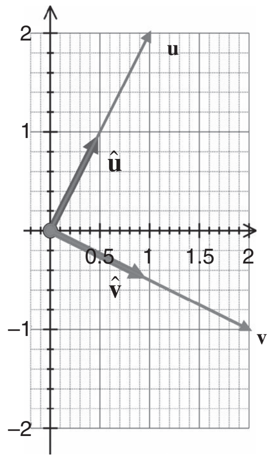


Figure 7.11

These vectors $\widehat{\mathbf{u}}$ and $\widehat{\mathbf{v}}$ form an orthonormal (perpendicular unit) basis for $\mathbb{R}^2$.

In the above Example 7.23 the diagonalizing matrix for $\mathbf{A}$ was $\mathbf{P} = (\mathbf{u} \quad \mathbf{v})$. However, we can show that the matrix $\mathbf{Q} = (\widehat{\mathbf{u}} \quad \widehat{\mathbf{v}})$ also diagonalizes the matrix $\mathbf{A}$ because $\widehat{\mathbf{u}}$ and $\widehat{\mathbf{v}}$ are the same vectors $\mathbf{u}$ and $\mathbf{v}$ but normalized, which means that they have the same direction as $\mathbf{u}$ and $\mathbf{v}$ but have length 1-see Fig. 7.11.

Example 7.24

Show that $\mathbf{Q} = (\widehat{\mathbf{u}} \quad \widehat{\mathbf{v}})$ diagonalizes the matrix $\mathbf{A}$ of Example 7.23.

Solution

What is the matrix $\mathbf{Q}$ *equal to?*

$$\mathbf{Q} = (\widehat{\mathbf{u}} \quad \widehat{\mathbf{v}}) = \left(\frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \end{pmatrix} \right) \qquad \left[\text{because } \widehat{\mathbf{u}} = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix} \text{ and } \widehat{\mathbf{v}} = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \end{pmatrix}\right]$$

$$= \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} = \frac{1}{\sqrt{5}}\mathbf{P} \qquad \left[\text{because } \mathbf{P} = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}\right]$$

How do we show that matrix $\mathbf{Q}$ *diagonalizes matrix* $\mathbf{A}$*?*

We need to verify that $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{D}$, where $\mathbf{D}$ is the diagonal eigenvalue matrix. From the above we have $\mathbf{Q} = \frac{1}{\sqrt{5}}\mathbf{P}$, and taking the inverse of this gives

$$\mathbf{Q}^{-1} = \left(\frac{1}{\sqrt{5}}\mathbf{P}\right)^{-1} = \left(\frac{1}{\sqrt{5}}\right)^{-1}\mathbf{P}^{-1} = \sqrt{5}\,\mathbf{P}^{-1} \qquad \left[\text{because } (k\mathbf{A})^{-1} = k^{-1}\mathbf{A}^{-1}\right]$$

Substituting these into $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ yields

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \left(\sqrt{5}\,\mathbf{P}^{-1}\right)\mathbf{A}\left(\frac{1}{\sqrt{5}}\mathbf{P}\right) \underbrace{=}_{\text{cancelling } \sqrt{5}} \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \underbrace{=}_{\text{by Example 7.23}} \mathbf{D}$$

Thus matrix $\mathbf{Q}$ diagonalizes the matrix $\mathbf{A}$ because $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{D}$.

7.4.2 Orthogonal matrices

The matrix $\mathbf{Q}$ which diagonalizes $\mathbf{A}$ in the above Example 7.24 is an *orthogonal* matrix.

Can you remember what an orthogonal matrix is?

By chapter 4 Definition (4.18): A square matrix $\mathbf{Q} = (\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_n)$ the columns of which $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are *orthonormal* (perpendicular unit) vectors is called an *orthogonal matrix*.

Why is $\mathbf{Q}$ an orthogonal matrix in the above Example 7.24?

That's because the columns of $\mathbf{Q} = (\hat{\mathbf{u}}\quad \hat{\mathbf{v}})$ are perpendicular unit vectors, $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$, as illustrated in Fig. 7.11 above.

We use an orthogonal matrix $\mathbf{Q}$ to diagonalize a given matrix $\mathbf{A}$. One critical application of diagonalization is the evaluation of powers of a matrix, which we found in the previous section and was given by formula (7.19): $\mathbf{A}^m = \mathbf{P}\mathbf{D}^m\mathbf{P}^{-1}$

From chapter 4, Proposition (4.20) we have: If $\mathbf{Q}$ is an orthogonal matrix then $\mathbf{Q}^{-1} = \mathbf{Q}^T$.

In this case, the diagonalizing matrix is the orthogonal matrix $\mathbf{Q}$ therefore

$$\mathbf{A}^m = \mathbf{Q}\mathbf{D}^m\mathbf{Q}^{-1} = \mathbf{Q}\mathbf{D}^m\mathbf{Q}^T \qquad \left[\text{because for an orthogonal matrix } \mathbf{Q}^{-1} = \mathbf{Q}^T\right]$$

This means that calculating the power of a matrix is even simpler because we don't have to evaluate the inverse of matrix $\mathbf{Q}$ by some tedious method but just transpose matrix $\mathbf{Q}$. This is the great advantage of using an orthogonal matrix to diagonalize a matrix because

$$\mathbf{A}^m = \mathbf{Q}\mathbf{D}^m\mathbf{Q}^T \tag{7.20}$$

Evaluation of all the matrices on the right hand side; $\mathbf{Q}$, $\mathbf{D}^m$ and $\mathbf{Q}^T$, is straightforward.

Example 7.25

Find $\mathbf{A}^6$ for the matrix $\mathbf{A}$ given in Example 7.23.

Solution

To find $\mathbf{A}^6$, we use the above formula $\mathbf{A}^m = \mathbf{Q}\mathbf{D}^m\mathbf{Q}^T$ with $m = 6$:

$$\mathbf{A}^6 = \mathbf{Q}\mathbf{D}^6\mathbf{Q}^T \qquad (*)$$

By Examples 7.23 and 7.24 we have

$$\mathbf{D} = \begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix}, \quad \mathbf{Q} = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \text{ and taking the transpose } \mathbf{Q}^T = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} = \mathbf{Q}$$

(continued...)

Substituting these into (*) yields

$$\begin{aligned}
\mathbf{A}^6 = \mathbf{QD}^6\mathbf{Q}^T &= \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix}^6 \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \\
&= \frac{1}{\sqrt{5}}\frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 5^6 & 0 \\ 0 & (-5)^6 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \\
&= \frac{5^6}{5}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \quad \left[\text{because } (-5)^6 = 5^6\right] \\
&= 5^5\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} = 5^5\begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = 5^6\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 5^6\mathbf{I}
\end{aligned}$$

7.4.3 Properties of symmetric matrices

Proposition (7.21). Let $\mathbf{A}$ be a real symmetric matrix. Then *all* the eigenvalues of $\mathbf{A}$ are real.

Proof.
We omit the proof because we need to use complex numbers, which are not covered.

■

Proposition (7.22). Let $\mathbf{A}$ be a symmetric matrix. If λ_1 and λ_2 are *distinct* eigenvalues of matrix $\mathbf{A}$ then their corresponding eigenvectors $\mathbf{u}$ and $\mathbf{v}$ respectively are orthogonal (perpendicular).

How do we prove this result?
We use the dot product result of chapter 2:

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v}, \text{ and show that } \mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} = 0.$$

Proof.
Let $\mathbf{u}$ and $\mathbf{v}$ be eigenvectors belonging to distinct eigenvalues λ_1 and λ_2 respectively.

What do we need to prove?
Required to prove that the eigenvectors are orthogonal, which means that we need to show $\mathbf{u}\cdot\mathbf{v} = 0$. Since $\mathbf{u}$ and $\mathbf{v}$ are eigenvectors belonging to distinct eigenvalues λ_1 and λ_2, matrix $\mathbf{A}$ scalar multiplies each of the eigenvectors by λ_1 and λ_2 respectively:

$$\mathbf{Au} = \lambda_1\mathbf{u} \quad \text{and} \quad \mathbf{Av} = \lambda_2\mathbf{v} \qquad (*)$$

Taking the transpose of both sides of $\lambda_1\mathbf{u} = \mathbf{Au}$ gives

$$\begin{aligned}
(\lambda_1\mathbf{u})^T &= (\mathbf{Au})^T \\
\lambda_1\mathbf{u}^T &= \mathbf{u}^T\mathbf{A}^T \qquad \left[\text{by (1.19) (b) } (k\,\mathbf{C})^T = k\,\mathbf{C}^T \text{ and (d) } (\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T\right] \\
&= \mathbf{u}^T\mathbf{A} \qquad \left[\text{because } \mathbf{A} \text{ is symmetric therefore } \mathbf{A}^T = \mathbf{A}\right]
\end{aligned}$$

Right-multiplying the last line $\lambda_1 \mathbf{u}^T = \mathbf{u}^T \mathbf{A}$ by the eigenvector $\mathbf{v}$ gives

$$\begin{aligned} \lambda_1 \mathbf{u}^T \mathbf{v} &= \mathbf{u}^T \mathbf{A} \mathbf{v} \\ &= \mathbf{u}^T \lambda_2 \mathbf{v} \qquad [\text{by } (*)] \\ \lambda_1 \mathbf{u}^T \mathbf{v} - \lambda_2 \mathbf{u}^T \mathbf{v} &= 0 \\ (\lambda_1 - \lambda_2)\, \mathbf{u}^T \mathbf{v} &= 0 \qquad [\text{factorizing}] \end{aligned}$$

λ_1 and λ_2 are distinct eigenvalues therefore $\lambda_1 - \lambda_2 \neq 0$ [not zero] so we have

$$\mathbf{u}^T \mathbf{v} = 0 \text{ which means that } \mathbf{u} \cdot \mathbf{v} = 0$$

The dot product of the eigenvectors $\mathbf{u}$ and $\mathbf{v}$ is *zero*, therefore they are *orthogonal*.

■

We can extend Proposition (7.22) to n distinct eigenvalues:

Proposition (7.23). Let $\mathbf{A}$ be a symmetric matrix with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots$ and $\mathbf{v}_n$. Then these eigenvectors are orthogonal.

Proof – Exercises 7.3.

■

The next two examples show applications of this Proposition (7.23).

Example 7.26

Show that the eigenvectors of matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ are orthogonal.

Solution

If you have reached this point then you should be able to find the eigenvalues and eigenvectors of the given matrix. Verify that the characteristic equation $p(\lambda)$ is given by

$$p(\lambda) = \lambda^3 - 4\lambda^2 - \lambda + 4 = 0$$
$$(\lambda + 1)(\lambda - 1)(\lambda - 4) = 0 \text{ yields } \lambda_1 = -1, \lambda_2 = 1 \text{ and } \lambda_3 = 4$$

Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ be the eigenvectors belonging to $\lambda_1 = -1$, $\lambda_2 = 1$ and $\lambda_3 = 4$ respectively. We have (verify)

$$\mathbf{u} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \text{ and } \mathbf{w} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

(continued...)

How do we check that the eigenvectors $\mathbf{u}$*,* $\mathbf{v}$ *and* $\mathbf{w}$ *are orthogonal to each other?*
We need to confirm that the dot product is zero: $\mathbf{u}\cdot\mathbf{v}=0$, $\mathbf{u}\cdot\mathbf{w}=0$ and $\mathbf{v}\cdot\mathbf{w}=0$.

$$\mathbf{u}\cdot\mathbf{v}=\begin{pmatrix}1\\-1\\0\end{pmatrix}\cdot\begin{pmatrix}1\\1\\-2\end{pmatrix}=(1\times 1)+((-1)\times 1)+(0\times(-2))=0$$

$$\mathbf{u}\cdot\mathbf{w}=\begin{pmatrix}1\\-1\\0\end{pmatrix}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}=0\text{ and }\mathbf{v}\cdot\mathbf{w}=\begin{pmatrix}1\\1\\-2\end{pmatrix}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}=0$$

Thus the eigenvectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are orthogonal (perpendicular) to each other.

Example 7.27

Show that the eigenvectors belonging to *distinct* eigenvalues of $\mathbf{A}=\begin{pmatrix}2&2&-2\\2&-1&4\\-2&4&-1\end{pmatrix}$ are orthogonal.

Solution

By solving the determinant you can verify that the characteristic equation $p(\lambda)$ is

$$p(\lambda)=\lambda^3-27\lambda+54=0$$
$$(\lambda-3)^2(\lambda+6)=0\text{ gives }\lambda_1=3,\lambda_2=3\text{ and }\lambda_3=-6$$

Let $\mathbf{u}$, $\mathbf{v}$ be the eigenvectors belonging to $\lambda_1=3$, $\lambda_2=3$ and $\mathbf{w}$ be the eigenvector belonging to $\lambda_3=-6$. You can verify the following in your own time:

$$\left\{\mathbf{u}=\begin{pmatrix}2\\1\\0\end{pmatrix},\mathbf{v}=\begin{pmatrix}-2\\0\\1\end{pmatrix}\right\}\text{ belong to the } \textit{same} \text{ e.value }\lambda_1=3,\ \lambda_2=3\text{ and }\mathbf{w}=\begin{pmatrix}-1\\2\\-2\end{pmatrix}\text{ to }\lambda_3=-6$$

We need to check that eigenvectors $\mathbf{u}$, $\mathbf{w}$ and $\mathbf{v}$, $\mathbf{w}$ are orthogonal (dot product is zero):

$$\mathbf{u}\cdot\mathbf{w}=\begin{pmatrix}2\\1\\0\end{pmatrix}\cdot\begin{pmatrix}-1\\2\\-2\end{pmatrix}=0\text{ and }\mathbf{v}\cdot\mathbf{w}=\begin{pmatrix}-2\\0\\1\end{pmatrix}\cdot\begin{pmatrix}-1\\2\\-2\end{pmatrix}=0$$

However, note that the eigenvectors $\mathbf{u}$ and $\mathbf{v}$ belonging to the *same* eigenvalue $\lambda_1=3$ and $\lambda_2=3$ need *not* be orthogonal to each other. Actually

$$\mathbf{u}\cdot\mathbf{v}=\begin{pmatrix}2\\1\\0\end{pmatrix}\cdot\begin{pmatrix}-2\\0\\1\end{pmatrix}=(2\times(-2))+(1\times 0)+(0\times 1)=-4\quad(\dagger)$$

Thus the eigenvectors $\mathbf{u}$ and $\mathbf{v}$ belonging to the same eigenvalue $\lambda_1=3$ and $\lambda_2=3$ are *not* orthogonal because $\mathbf{u}\cdot\mathbf{v}=-4\neq 0$ [not zero].

7.4.4 Orthogonal diagonalization

Definition (7.24). In general, a matrix $\mathbf{A}$ is **orthogonally diagonalizable** if there is an orthogonal matrix $\mathbf{Q}$ such that

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{Q}^T\mathbf{A}\mathbf{Q} = \mathbf{D} \text{ where } \mathbf{D} \text{ is a diagonal matrix.}$$

Eigenvector matrix $\mathbf{Q}$ contains the eigenvectors of $\mathbf{A}$ and eigenvalue matrix $\mathbf{D}$ contains the eigenvalues of $\mathbf{A}$.

In the above Examples 7.23 and 7.24, the matrix $\mathbf{A} = \begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}$ is orthogonally diagonalizable because with $\mathbf{Q} = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}$ we have $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix}$ which is a diagonal matrix.

Theorem (7.25). Let $\mathbf{A}$ be a square matrix. If the matrix $\mathbf{A}$ is orthogonally diagonalizable then $\mathbf{A}$ is a symmetric matrix.

How do we prove this result?
We assume that $\mathbf{A}$ is orthogonally diagonalizable and deduce that $\mathbf{A}$ is symmetric, which means that we need to show that $\mathbf{A}^T = \mathbf{A}$.

Why?
Because $\mathbf{A}^T = \mathbf{A}$ means that the matrix $\mathbf{A}$ is symmetric.

Proof.
Assume that the matrix $\mathbf{A}$ is orthogonally diagonalizable. This means that there is an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{D}$, where $\mathbf{D}$ is a diagonal matrix. Left multiplying this $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{D}$ by $\mathbf{Q}$ and right multiplying by $\mathbf{Q}^{-1}$ gives

$$\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^{-1} \tag{†}$$

Taking the transpose of both sides gives

$$\begin{aligned}
\mathbf{A}^T = \left(\mathbf{Q}\mathbf{D}\mathbf{Q}^{-1}\right)^T &= \left(\mathbf{Q}^{-1}\right)^T\mathbf{D}^T\mathbf{Q}^T && \left[\text{by } (\mathbf{A}\mathbf{B}\mathbf{C})^T = \mathbf{C}^T\mathbf{B}^T\mathbf{A}^T\right] \\
&= \left(\mathbf{Q}^T\right)^{-1}\mathbf{D}\mathbf{Q}^T && \left[\text{using } \left(\mathbf{A}^{-1}\right)^T = \left(\mathbf{A}^T\right)^{-1} \text{ and } \mathbf{D}^T = \mathbf{D}\right] \\
&= \left(\mathbf{Q}^{-1}\right)^{-1}\mathbf{D}\mathbf{Q}^{-1} && \left[\text{because } \mathbf{Q} \text{ is orthogonal, } \mathbf{Q}^T = \mathbf{Q}^{-1}\right] \\
&= \mathbf{Q}\mathbf{D}\mathbf{Q}^{-1} && \left[\text{because } \left(\mathbf{Q}^{-1}\right)^{-1} = \mathbf{Q}\right]
\end{aligned}$$

We have $\mathbf{A}^T = \mathbf{Q}\mathbf{D}\mathbf{Q}^{-1}$.

What do you notice?

By (†) we can see that this is equal to matrix $\mathbf{A}$. Thus

$$\mathbf{A}^T = \mathbf{QDQ}^{-1} = \mathbf{A} \text{ which means we have } \mathbf{A}^T = \mathbf{A}$$

Hence $\mathbf{A}$ is a symmetric matrix because $\mathbf{A}^T = \mathbf{A}$, which is our required result.

■

Now we will show that the other way round is also true.

What does this mean?

If $\mathbf{A}$ is a symmetric matrix then $\mathbf{A}$ is orthogonally diagonalizable. This means that there is at least one type of matrix, *symmetric matrices*, which can be *diagonalized orthogonally*.

Lemma (7.26). If $\mathbf{A}$ is a real symmetric matrix with an eigenvalue λ of multiplicity m then λ has m linearly independent eigenvectors.

Proof – See <http://www.oup.co.uk/companion/singh>.

■

Theorem (7.27). If $\mathbf{A}$ is an n by n symmetric matrix then $\mathbf{A}$ is orthogonally diagonalizable.

How do we prove this?

We consider two cases:

(i) $\mathbf{A}$ has distinct eigenvalues (ii) $\mathbf{A}$ does *not* have distinct eigenvalues

Proof.

Case (i): Let the symmetric matrix $\mathbf{A}$ have distinct eigenvalues $\lambda_1, \lambda_2, \ldots$ and λ_n. Then by Proposition (7.23): Let $\mathbf{A}$ be a symmetric matrix with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots$ and $\mathbf{v}_n$. Then these eigenvectors are orthogonal.

So the eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots$ and $\mathbf{v}_n$ belonging to distinct eigenvalues $\lambda_1, \lambda_2, \ldots$ and λ_n are orthogonal. Because they are orthogonal, they are linearly independent and so we have n linearly independent eigenvectors. By Theorem (7.16):

An n by n matrix $\mathbf{A}$ is diagonalizable $\Leftrightarrow$ it has n independent eigenvectors.

We have that the matrix $\mathbf{A}$ is diagonalizable. Let $\mathbf{Q} = (\widehat{\mathbf{v}_1} \quad \widehat{\mathbf{v}_2} \quad \cdots \quad \widehat{\mathbf{v}_n})$, then the columns of the matrix $\mathbf{Q}$ are orthonormal, which means it is an orthogonal matrix. Thus we have $\mathbf{Q}^{-1}\mathbf{AQ} = \mathbf{D}$, so matrix $\mathbf{A}$ is orthogonally diagonalizable.

Case (ii): Let the symmetric matrix $\mathbf{A}$ have an eigenvalue λ with multiplicity $m > 1$. By the above Lemma (7.26) λ has m linearly independent eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots$ and $\mathbf{u}_m$. These are a basis for the eigenspace E_λ. By the Gram–Schmidt process we can convert these m vectors into orthonormal basis vectors for E_λ.

We can repeat this process for any other eigenvectors belonging to eigenvalues of $\mathbf{A}$ which have a multiplicity of more than 1.

All the remaining eigenvalues are *distinct*, so the eigenvectors are orthogonal.

Thus *all* the vectors are orthogonal. By repeating the procedure outlined in case (i), we conclude that matrix $\mathbf{A}$ is orthogonally diagonalizable.

■

By combining Theorems (7.25) and (7.27) what can we conclude?

If $\mathbf{A}$ is orthogonally diagonalizable then $\mathbf{A}$ is a symmetric matrix and the other way round; that is, if $\mathbf{A}$ is symmetric matrix then $\mathbf{A}$ is orthogonally diagonalizable. We have the main result of this section which has a special name – **spectral theorem**:

Spectral theorem (7.28). Matrix $\mathbf{A}$ is orthogonally diagonalizable $\Leftrightarrow$ $\mathbf{A}$ is a symmetric matrix.

Proof – By the above Theorems (7.25) and (7.27).

■

This means that if we have a symmetric matrix then we can orthogonally diagonalize it. This is the spectral theorem for real matrices.

Example 7.28

Determine an orthogonal matrix $\mathbf{Q}$ which orthogonally diagonalizes $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 2 & 0 \end{pmatrix}$.

Solution

Since matrix $\mathbf{A}$ is symmetric, we can find an orthogonal matrix $\mathbf{Q}$ which diagonalizes $\mathbf{A}$. Verify that the characteristic polynomial is given by:

$$\lambda^2 - 3\lambda - 4 = 0 \Rightarrow (\lambda + 1)(\lambda - 4) = 0$$
$$\lambda_1 = -1 \text{ and } \lambda_2 = 4$$

Let $\mathbf{u}$ and $\mathbf{v}$ be eigenvectors (verify) belonging to $\lambda_1 = -1$ and $\lambda_2 = 4$ respectively:

$$\mathbf{u} = \begin{pmatrix} 1 \\ -2 \end{pmatrix} \text{ and } \mathbf{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

The given matrix $\mathbf{A}$ is symmetric and we have distinct eigenvalues, therefore $\mathbf{u}$ and $\mathbf{v}$ are orthogonal (that is $\mathbf{u} \cdot \mathbf{v} = 0$). Remember, the question says that we have to find an orthogonal matrix $\mathbf{Q}$.

What is an orthogonal matrix?

A square matrix the columns of which form an orthonormal set.

What is an orthonormal set?

A set that is *orthogonal* and *normalized*. The eigenvectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal but we need to normalize them, which means make their norm (length) to be 1.

How?

Divide by the norm (length) of the eigenvector, which in both cases is $\sqrt{5}$ because

$$\|\mathbf{u}\| = \sqrt{1^2 + (-2)^2} = \sqrt{5} \text{ and } \|\mathbf{v}\| = \sqrt{1^2 + 2^2} = \sqrt{5}$$

Normalizing gives

$$\widehat{\mathbf{u}} = \frac{1}{\|\mathbf{u}\|}\mathbf{u} = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ -2 \end{pmatrix} \text{ and } \widehat{\mathbf{v}} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

(continued...)

Thus our orthogonal matrix $\mathbf{Q}$ is given by

$$\mathbf{Q} = (\widehat{\mathbf{u}} \quad \widehat{\mathbf{v}}) = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

$\mathbf{Q}$ is an orthogonal matrix, therefore $\mathbf{Q}^{-1} = \mathbf{Q}^T$. Verify that $\mathbf{Q}^T\mathbf{AQ} = \mathbf{D}$ or $\mathbf{QD} = \mathbf{AQ}$ where $\mathbf{D}$ is a diagonal matrix with eigenvalues along the leading diagonal:

$$\mathbf{Q}^T\mathbf{AQ} = \mathbf{D} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix} \quad \begin{bmatrix} -1,\ 4 \text{ are the eigenvalues} \\ \text{of the given matrix } \mathbf{A} \end{bmatrix}$$

7.4.5 Procedure for orthogonal diagonalization

We can work through a procedure to find an orthogonal matrix which diagonalizes a given matrix.

The procedure for orthogonal diagonalization of a symmetric matrix $\mathbf{A}$ is as follows:

1. Determine the eigenvalues of $\mathbf{A}$.
2. Find the corresponding eigenvectors.
3. If any of the eigenvalues are repeated then check that the associated eigenvectors are orthogonal. If they are *not* orthogonal then place them into an orthogonal set by using the Gram–Schmidt process described in chapter 4.
4. Normalize all the eigenvectors.
5. Form the orthogonal matrix $\mathbf{Q}$ whose columns are the orthonormal eigenvectors.
6. Check that $\mathbf{QD} = \mathbf{AQ}$, where $\mathbf{D}$ is the diagonal matrix whose entries along the leading diagonal are the eigenvalues of matrix $\mathbf{A}$.

Example 7.29

Determine an orthogonal matrix $\mathbf{Q}$ which diagonalizes $\mathbf{A} = \begin{pmatrix} 2 & 2 & -2 \\ 2 & -1 & 4 \\ -2 & 4 & -1 \end{pmatrix}$.

Solution

Steps 1 and 2:

This is the same matrix as Example 7.27. Thus we have

$$\left\{ \mathbf{u} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right\} \text{ belong to } \lambda_{1,\,2} = 3 \text{ and } \mathbf{w} = \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix} \text{ to } \lambda_3 = -6$$

Step 3:

We need to check that eigenvectors $\mathbf{u}$ and $\mathbf{v}$ belonging to repeated eigenvalue $\lambda_{1,\,2} = 3$ are orthogonal.

Remember that in Example 7.27 we showed that $\mathbf{u} \cdot \mathbf{v} = -4$, therefore eigenvectors $\mathbf{u}$ and $\mathbf{v}$ are *not* orthogonal. We need to convert these $\mathbf{u}$ and $\mathbf{v}$ into an orthogonal set, say $\mathbf{q}_1$ and $\mathbf{q}_2$ respectively.

How do we convert $\mathbf{u}$ and $\mathbf{v}$ into the orthogonal set $\mathbf{q}_1$ and $\mathbf{q}_2$?

By using the Gram—Schmidt process (4.16):

$$\mathbf{q}_1 = \mathbf{u} \quad \text{and} \quad \mathbf{q}_2 = \mathbf{v} - \frac{\mathbf{v} \cdot \mathbf{q}_1}{\|\mathbf{q}_1\|^2}\mathbf{q}_1 \qquad (*)$$

How do we evaluate $\mathbf{q}_1$ and $\mathbf{q}_2$?

By substituting $\mathbf{u}$ and $\mathbf{v}$ and evaluating $\mathbf{v} \cdot \mathbf{q}_1$ and $\|\mathbf{q}_1\|^2$. We have

$$\mathbf{q}_1 = \mathbf{u} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \text{ and } \mathbf{v} \cdot \mathbf{q}_1 = \mathbf{v} \cdot \mathbf{u} = \mathbf{u} \cdot \mathbf{v} = -4 \;\; \text{[already evaluated]}$$

$$\|\mathbf{q}_1\|^2 = \left\| \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \right\|^2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = 2^2 + 1^2 + 0^2 = 5 \quad (**)$$

Substituting the above $\mathbf{v} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{q}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{v} \cdot \mathbf{q}_1 = -4$ and $\|\mathbf{q}_1\|^2 = 5$ into $\mathbf{q}_2 = \mathbf{v} - \frac{\mathbf{v} \cdot \mathbf{q}_1}{\|\mathbf{q}_1\|^2}\mathbf{q}_1$:

$$\begin{aligned}\mathbf{q}_2 = \mathbf{v} - \frac{\mathbf{v} \cdot \mathbf{q}_1}{\|\mathbf{q}_1\|^2}\mathbf{q}_1 &= \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} - \frac{(-4)}{5}\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} -2 + 8/5 \\ 0 + 4/5 \\ 1 - 0 \end{pmatrix} = \begin{pmatrix} -2/5 \\ 4/5 \\ 1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -2 \\ 4 \\ 5 \end{pmatrix} = \mathbf{q}_2\end{aligned}$$

Thus we have $\mathbf{q}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{q}_2 = \frac{1}{5}\begin{pmatrix} -2 \\ 4 \\ 5 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix}$, which are orthogonal to each other.

What else do we need to do?

Step 4:

We need to normalize these eigenvectors by dividing by the norm (length) of each. We have already established above in (**) that $\|\mathbf{q}_1\|^2 = 5$. Taking the square root gives $\|\mathbf{q}_1\| = \sqrt{5}$.

Remember, we can ignore any fractions (scalars) because vectors are orthogonal independent of scalars. Thus for the eigenvector $\mathbf{q}_2$, we can ignore the fraction 1/5 and call this $\mathbf{q}_2^*$:

$$\|\mathbf{q}_2^*\|^2 = \begin{pmatrix} -2 \\ 4 \\ 5 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ 4 \\ 5 \end{pmatrix} = (-2)^2 + 4^2 + 5^2 = 45$$

Taking the square root gives $\|\mathbf{q}_2^*\| = \sqrt{45} = 3\sqrt{5}$. Similarly we have

$$\|\mathbf{w}\|^2 = \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix} = (-1)^2 + 2^2 + (-2)^2 = 9$$

(continued...)

Taking the square root gives the norm $\|\mathbf{w}\| = 3$. Normalizing means we divide each vector by its norm (length):

$$\widehat{\mathbf{q}}_1 = \frac{1}{\sqrt{5}}\begin{pmatrix}2\\1\\0\end{pmatrix}, \quad \widehat{\mathbf{q}}_2^* = \frac{1}{3\sqrt{5}}\begin{pmatrix}-2\\4\\5\end{pmatrix} \quad \text{and} \quad \widehat{\mathbf{w}} = \frac{1}{3}\begin{pmatrix}-1\\2\\-2\end{pmatrix}$$

Step 5:
What is the orthogonal matrix $\mathbf{Q}$ equal to?

$$\mathbf{Q} = \begin{pmatrix}\widehat{\mathbf{q}}_1 & \widehat{\mathbf{q}}_2^* & \widehat{\mathbf{w}}\end{pmatrix} = \begin{pmatrix}2/\sqrt{5} & -2/3\sqrt{5} & -1/3\\ 1/\sqrt{5} & 4/3\sqrt{5} & 2/3\\ 0 & 5/3\sqrt{5} & -2/3\end{pmatrix}$$

Step 6:
Check $\mathbf{QD} = \mathbf{AQ}$, where $\mathbf{D}$ is a diagonal matrix with entries on the leading diagonal given by the eigenvalues of the matrix $\mathbf{A}$.
What is $\mathbf{D}$ equal to?

$$\mathbf{D} = \begin{pmatrix}3 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & -6\end{pmatrix} \quad \left[\begin{array}{l}\text{eigenvalues of the given matrix}\\ \mathbf{A} \text{ are } \lambda_1 = \lambda_2 = 3 \text{ and } \lambda_3 = -6\end{array}\right]$$

Summary

Definition (7.23). In general, a matrix $\mathbf{A}$ is *orthogonally diagonalizable* if there is an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{Q}^{-1}\mathbf{AQ} = \mathbf{Q}^T\mathbf{AQ} = \mathbf{D}$, where $\mathbf{D}$ is a diagonal matrix.

Spectral Theorem (7.28).

Matrix $\mathbf{A}$ is orthogonally diagonalizable $\Leftrightarrow$ $\mathbf{A}$ is a symmetric matrix.

EXERCISES 7.4

(Brief solutions at end of book. Full solutions available at <http://www.oup.co.uk/companion/singh>.)

In this exercise check your numerical answers using MATLAB.

1. For the following matrices find an orthogonal matrix $\mathbf{Q}$ which diagonalizes the given matrix. Also check that $\mathbf{Q}^T\mathbf{AQ} = \mathbf{D}$ where $\mathbf{D}$ is a diagonal matrix.

   (a) $\mathbf{A} = \begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$ (c) $\mathbf{A} = \begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}$ (d) $\mathbf{A} = \begin{pmatrix}5 & 12\\ 12 & -5\end{pmatrix}$

2. For the following matrices find an orthogonal matrix $\mathbf{Q}$ which diagonalizes the given matrix. Also check that $\mathbf{Q}^T\mathbf{AQ} = \mathbf{D}$ where $\mathbf{D}$ is a diagonal matrix.

   (a) $\mathbf{A} = \begin{pmatrix}9 & 3\\ 3 & 1\end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix}3 & \sqrt{2}\\ \sqrt{2} & 2\end{pmatrix}$

   (c) $\mathbf{A} = \begin{pmatrix}-5 & \sqrt{3}\\ \sqrt{3} & -3\end{pmatrix}$ (d) $\mathbf{A} = \begin{pmatrix}5 & \sqrt{12}\\ \sqrt{12} & 1\end{pmatrix}$

3. For the following matrices find an orthogonal matrix $\mathbf{Q}$ which diagonalizes the given matrix. By using MATLAB or otherwise check that $\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \mathbf{D}$ where $\mathbf{D}$ is a diagonal matrix.

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

4. For the following matrices find an orthogonal matrix $\mathbf{Q}$ which diagonalizes the given matrix. By using MATLAB or otherwise check that $\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \mathbf{D}$.

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

5. Let $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Show that $\mathbf{A}^{10} = 2^9\mathbf{A}$. Also prove that $\mathbf{A}^m = 2^{m-1}\mathbf{A}$ where m is a positive integer.

6. Show that if $\mathbf{A}$ is a diagonal matrix then orthogonal diagonalising matrix $\mathbf{Q} = \mathbf{I}$.

7. Prove that (a) the zero matrix $\mathbf{O}$ and (b) the identity matrix $\mathbf{I}$ are orthogonally diagonalisable.

8. Prove that $\mathbf{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \neq \mathbf{O}$ is orthogonally diagonalisable and find the orthogonal matrix $\mathbf{Q}$ which diagonalizes the matrix $\mathbf{A}$.

   [Hint: If the quadratic $x^2 + px + q = 0$ has roots a and b then $a + b = -p$.]

9. Let $\mathbf{A}$ be a symmetric invertible matrix. If $\mathbf{Q}$ orthogonally diagonalizes the matrix $\mathbf{A}$ show that $\mathbf{Q}$ also diagonalizes the matrix $\mathbf{A}^{-1}$.

10. Prove Proposition (7.23).

SECTION 7.5 Singular Value Decomposition

By the end of this section you will be able to

- understand what is meant by SVD
- find a triple factorization of any matrix

The singular value decomposition (SVD) is one of the most important factorizations of a matrix. SVD factorization breaks the matrix down into useful parts such as orthogonal matrices, and the method can be applied to any matrix; it does *not* need to be a square or symmetric matrix.

The SVD of a matrix gives us an orthogonal basis (axes) for the row and column space of the matrix. If we consider a matrix as a transformation then SVD factorization gives an orthogonal basis (axes) for the start and arrival vector spaces (Fig. 7.12).

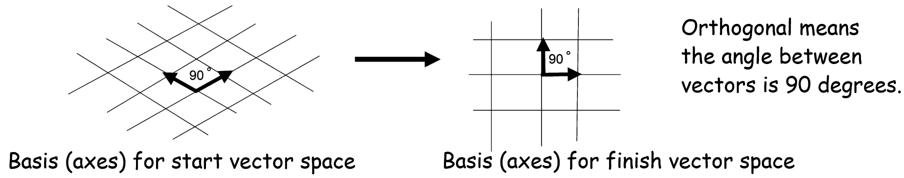


Figure 7.12

A very good application of SVD is given in the article by David Austin in the Monthly Essays on Mathematical Topics of August 2009:

> Netflix, the online movie rental company, is currently offering a $1 million prize for anyone who can improve the accuracy of its movie recommendation system by 10%. Surprisingly, this seemingly modest problem turns out to be quite challenging, and the groups involved are now using rather sophisticated techniques. At the heart of all of them is the singular value decomposition.

First, we look at the geometric significance of SVD factorization.

7.5.1 Geometric interpretation of singular value decomposition (SVD)

To find the SVD of any matrix $\mathbf{A}$ we use the matrix $\mathbf{A}^T\mathbf{A}$ because $\mathbf{A}^T\mathbf{A}$ is a symmetric matrix, as we will show later in this section. Remember, symmetric matrices can be orthogonally diagonalized.

First we define the **singular values** of a matrix $\mathbf{A}$.

Definition (7.29). Let $\mathbf{A}$ be any m by n matrix and $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of $\mathbf{A}^T\mathbf{A}$, then the **singular values** of $\mathbf{A}$ denoted by $\sigma_1, \sigma_2, \ldots, \sigma_n$ are the numbers:

$$\sigma_1 = \sqrt{\lambda_1},\ \sigma_2 = \sqrt{\lambda_2}, \ldots, \sigma_n = \sqrt{\lambda_n} \quad \text{[positive root only]}$$

Example 7.30

Find the eigenvalues and eigenvectors of $\mathbf{A}^T\mathbf{A}$ where $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

Solution

Since the given matrix $\mathbf{A}$ is a symmetric matrix so $\mathbf{A}^T = \mathbf{A}$. We have $\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A} = \mathbf{A}^2$.

We found the eigenvalues t_1 and t_2 with the normalized eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ of this matrix $\mathbf{A}$ in Exercises 7.4 question 1(c):

$$t_1 = 3,\ \mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad t_2 = 1, \mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

What are the eigenvalues and eigenvectors of matrix $\mathbf{A}^2$?

By Proposition (7.8)(a): If m is a natural number then λ^m is an eigenvalue of the matrix $\mathbf{A}^m$ with the *same* eigenvector $\mathbf{u}$.

Let λ_1 and λ_2 be the eigenvalues of matrix $\mathbf{A}^2$, then by using this proposition we have

$$\lambda_1 = 3^2 = 9, \mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \quad \text{and} \quad \lambda_2 = 1^2 = 1, \mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$$

Example 7.31

(i) Find the singular values $\sigma_1 = \sqrt{\lambda_1}$ and $\sigma_2 = \sqrt{\lambda_2}$ of the matrix $\mathbf{A}$ given in the above Example 7.30.

(ii) Determine $\sigma_1\mathbf{u}_1 = \mathbf{A}\mathbf{v}_1$ and $\sigma_2\mathbf{u}_2 = \mathbf{A}\mathbf{v}_2$, where $\mathbf{v}_1$ and $\mathbf{v}_2$ are the normalized eigenvectors belonging to the eigenvalues of matrix $\mathbf{A}^T\mathbf{A}$.

Solution

(i) From Example 7.30 we have the eigenvalues $\lambda_1 = 9$ and $\lambda_2 = 1$. Taking the square root:

$$\sigma_1 = \sqrt{9} = 3 \quad \text{and} \quad \sigma_2 = \sqrt{1} = 1 \qquad \text{[positive root only]}$$

The singular values of matrix $\mathbf{A}$ are 3 and 1. Since $\mathbf{A}$ is a symmetric matrix, the singular values of $\mathbf{A}$ are the eigenvalues of $\mathbf{A}$. This would *not* be the case if $\mathbf{A}$ was a non-symmetric matrix.

(ii) Substituting $\mathbf{A} = \begin{pmatrix}2 & 1\\1 & 2\end{pmatrix}$, $\mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ and $\sigma_1 = 3$ into $\sigma_1\mathbf{u}_1 = \mathbf{A}\mathbf{v}_1$ gives

$$3\mathbf{u}_1 = \mathbf{A}\mathbf{v}_1 = \begin{pmatrix}2 & 1\\1 & 2\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}3\\3\end{pmatrix}$$

Similarly $\sigma_2\mathbf{u}_2 = \mathbf{A}\mathbf{v}_2$ is

$$(1)\,\mathbf{u}_2 = \mathbf{A}\mathbf{v}_2 = \begin{pmatrix}2 & 1\\1 & 2\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$$

We have

$$\mathbf{A}\mathbf{v}_1 = 3\mathbf{u}_1 \quad \text{and} \quad \mathbf{A}\mathbf{v}_2 = \mathbf{u}_2$$

What does this mean?

It means that the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(\mathbf{v}) = \mathbf{A}\mathbf{v}$ transforms the vector $\mathbf{v}_1$ to 3 times the vector $\mathbf{u}_1$ in the same direction. The vector $\mathbf{v}_2$ is transformed under the matrix $\mathbf{A}$ to the same vector, $\mathbf{v}_2 = \mathbf{u}_2$ (Fig. 7.13).

The unit circle in Fig. 7.13(a) is transformed to an ellipse in Fig. 7.13(b) under the matrix $\mathbf{A}$. Note that the unit circle in Fig. 7.13(a) is stretched by the factors $\sigma_1 = 3$ and $\sigma_2 = 1$ in the direction of $\mathbf{u}_1$ and $\mathbf{u}_2$ respectively. Remember, these factors $\sigma_1 = 3$ and $\sigma_2 = 1$ are the singular values of matrix $\mathbf{A}$. Also observe that vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal (perpendicular) and normalized (length equals 1). In SVD factorization, orthogonal vectors get transformed to orthogonal vectors – this is why this factorization is the most useful.

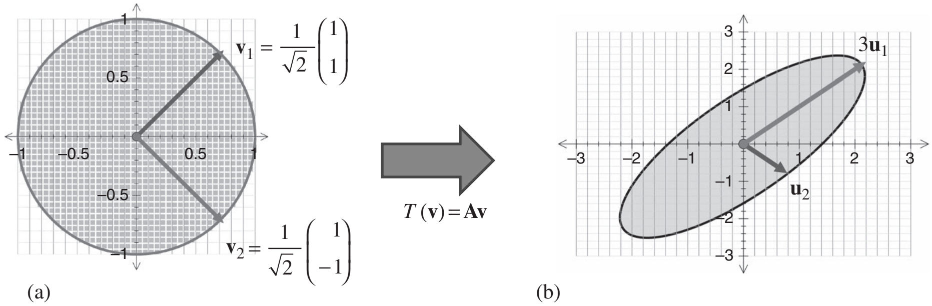


Figure 7.13

Actually the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ form an orthonormal (perpendicular unit) basis (axes) for the arrival vector space $\mathbb{R}^2$, similarly $\mathbf{v}_1$ and $\mathbf{v}_2$ form an orthonormal basis (axes) for the start vector space $\mathbb{R}^2$.

The eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ give us the direction of the semi-axes and the singular values σ_1 and σ_2 give us the length of the semi-axes.

What do we mean by semi-axes?

We illustrate the semi-axes for an ellipse in Fig. 7.14.

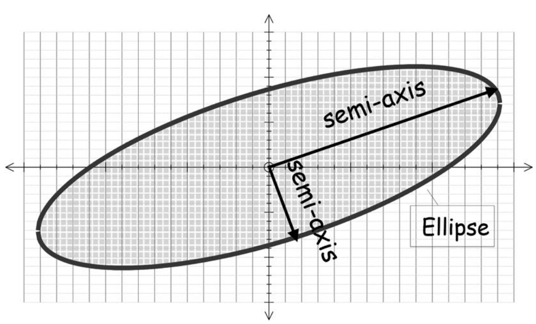


Figure 7.14

We can write $\mathbf{A}\mathbf{v}_1 = 3\mathbf{u}_1$ and $\mathbf{A}\mathbf{v}_2 = \mathbf{u}_2$ in matrix form as

$$\mathbf{A}(\mathbf{v}_1 \;\; \mathbf{v}_2) = (3\mathbf{u}_1 \;\; \mathbf{u}_2) = (\mathbf{u}_1 \;\; \mathbf{u}_2)\begin{pmatrix}3 & 0\\0 & 1\end{pmatrix} \qquad (*)$$

If we let $\mathbf{V} = (\mathbf{v}_1 \;\; \mathbf{v}_2)$, $\mathbf{U} = (\mathbf{u}_1 \;\; \mathbf{u}_2)$ and $\mathbf{D} = \begin{pmatrix}3 & 0\\0 & 1\end{pmatrix}$ then $(*)$ can be written as

$$\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{D} \qquad (\dagger)$$

Matrices $\mathbf{U}$ and $\mathbf{V}$ are orthogonal because $\mathbf{V} = (\mathbf{v}_1 \;\; \mathbf{v}_2)$ and $\mathbf{U} = (\mathbf{u}_1 \;\; \mathbf{u}_2)$ contain the orthonormal vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{u}_1$ and $\mathbf{u}_2$ which are illustrated in the above Fig. 7.13.

Right multiplying (†) by the inverse of $\mathbf{V}$ (remember for an orthogonal matrix $\mathbf{V}^{-1} = \mathbf{V}^T$) which is $\mathbf{V}^T$:

$$(\mathbf{AV})\,\mathbf{V}^T = \mathbf{A}\underbrace{\left(\mathbf{VV}^T\right)}_{=\mathbf{I}} = \mathbf{A} = \mathbf{UDV}^T$$

Hence we have factorized matrix $\mathbf{A} = \mathbf{UDV}^T$. Matrix $\mathbf{A}$ is broken into a *triple* matrix with the diagonal matrix $\mathbf{D}$ sandwiched between the two orthogonal matrices $\mathbf{U}$ and $\mathbf{V}^T$.

This is a singular value decomposition (SVD) of matrix $\mathbf{A}$.

Example 7.32

Check that $\mathbf{A} = \mathbf{UDV}^T$ for $\mathbf{A}$ of Example 7.30, that is the matrix $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

Solution

What are the orthogonal matrices $\mathbf{U}$, $\mathbf{V}$ and the diagonal matrix $\mathbf{D}$ equal to?

By the results of Example 7.31 we have $\mathbf{u}_1 = \frac{1}{3\sqrt{2}}\begin{pmatrix} 3 \\ 3 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\sigma_1 = 3$ and $\sigma_2 = 1$:

$$\mathbf{U} = (\,\mathbf{u}_1\ \ \mathbf{u}_2\,) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{V} = (\,\mathbf{v}_1\ \ \mathbf{v}_2\,) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Carrying out the matrix multiplication $\mathbf{U} \times \mathbf{D} \times \mathbf{V}^T$:

$$\begin{aligned} \mathbf{UDV}^T &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^T \\ &= \frac{1}{2}\begin{pmatrix} 3 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \mathbf{A} \end{aligned}$$

Note, the similarity to orthogonal diagonalization from the last section, $\mathbf{A} = \mathbf{QDQ}^T$. As the given matrix $\mathbf{A}$ is a symmetric matrix, then $\mathbf{U} = \mathbf{Q}$, $\mathbf{D} = \mathbf{D}$ and $\mathbf{Q}^T = \mathbf{V}^T$.

What use is this section on SVD if we can simply apply the orthogonal diagonalization technique of the last section?

Well, for SVD you do *not* need square or symmetric matrix. (Recall that symmetric matrices must be *square* matrices because if the number of rows does *not* equal the number of columns then transposing changes the shape of the matrix.) SVD can be applied any matrix.

7.5.2 Introduction to singular value decomposition (SVD)

In the previous section, and again above, we factorized only symmetric matrices. In this section we extend the factorization or decomposition to any matrix. We factorize the matrix $\mathbf{A}$ where $\mathbf{A}$ is m by n and $m \geq n$. The results in this section are also true if $m < n$ but we have chosen $m \geq n$ for convenience. Note that the results in this section are valid for ANY matrix $\mathbf{A}$.

Singular value decomposition theorem (7.30).

We can decompose any given matrix $\mathbf{A}$ of size m by n with positive singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > 0$ where $k \leq n$, into $\mathbf{UDV}^T$, that is

$$\mathbf{A} = \mathbf{UDV}^T$$

where $\mathbf{U}$ is an m by m *orthogonal* matrix, $\mathbf{D}$ is an m by n matrix and $\mathbf{V}$ is an n by n *orthogonal* matrix. The values of m (rows) and n (columns) is the size of the given matrix $\mathbf{A}$.

We have the situation shown in Fig. 7.15.

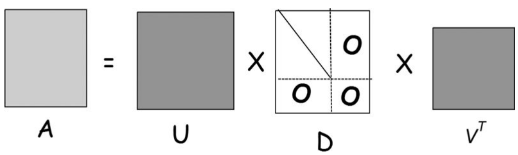


Figure 7.15

The matrix $\mathbf{D}$ looks a bit odd but it is a diagonal-like matrix. The matrix $\mathbf{D}$ has the positive singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > 0$ of matrix $\mathbf{A}$ starting from the top left hand corner of the matrix and working diagonally towards the bottom right. The symbol $\mathbf{O}$ represents the zero matrix of an appropriate size. Matrix $\mathbf{D}$ is of shape (Fig. 7.16).

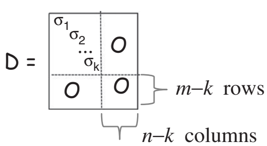


Figure 7.16

We need to be careful with the size of the matrices in Theorem (7.30).

If the given matrix $\mathbf{A}$ is a 3 by 2 then what size are the matrices $\mathbf{U}$, $\mathbf{D}$ and $\mathbf{V}$ in the above formula (7.30)?

$\mathbf{U}$ is a 3 by 3 matrix, $\mathbf{D}$ is a 3 by 2 matrix and $\mathbf{V}$ is a 2 by 2 matrix.

Example 7.33

Find the eigenvalues and eigenvectors of $\mathbf{A}^T\mathbf{A}$ where $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -2 & 1 \end{pmatrix}$.

Solution

First, we carry out the matrix multiplication:

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 1 & 1 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 3 \end{pmatrix}$$

Since $\mathbf{A}^T\mathbf{A}$ is a diagonal matrix, the entries on the leading diagonal are the eigenvalues

$$\lambda_1 = 6 \text{ and } \lambda_2 = 3.$$

Verify that the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ belonging to these eigenvalues $\lambda_1 = 6$ and $\lambda_2 = 3$ are

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Note that we *cannot* find eigenvalues of matrix $\mathbf{A}$ because $\mathbf{A}$ is a non-square matrix. This is why the question says find the eigenvalues of $\mathbf{A}^T\mathbf{A}$.

Note that in the above example $\mathbf{A}^T\mathbf{A}$ is a symmetric matrix.

Is this always the case?

Yes.

Proposition (7.30). Let $\mathbf{A}$ be any matrix. Then $\mathbf{A}^T\mathbf{A}$ is a symmetric matrix.

Proof.

Remember, a matrix $\mathbf{X}$ is symmetric if $\mathbf{X}^T = \mathbf{X}$.

To show that $\mathbf{A}^T\mathbf{A}$ is a symmetric matrix, we need to prove $\left(\mathbf{A}^T\mathbf{A}\right)^T = \mathbf{A}^T\mathbf{A}$:

$$\begin{aligned} \left(\mathbf{A}^T\mathbf{A}\right)^T &= \mathbf{A}^T\left(\mathbf{A}^T\right)^T && \left[\text{by (1.19) (d) } (\mathbf{XY})^T = \mathbf{Y}^T\mathbf{X}^T\right] \\ &= \mathbf{A}^T\mathbf{A} && \left[\text{by (1.19) (a) } \left(\mathbf{X}^T\right)^T = \mathbf{X}\right] \end{aligned}$$

Hence $\mathbf{A}^T\mathbf{A}$ is a symmetric matrix.

■

What do we know about diagonalizing a symmetric matrix?

From the previous section, we have the spectral theorem (7.28):

Matrix $\mathbf{A}$ is orthogonally diagonalizable $\Leftrightarrow$ $\mathbf{A}$ is a symmetric matrix.

This means that we can orthogonally diagonalize the matrix $\mathbf{A}^T\mathbf{A}$ because it is a symmetric matrix. This is why we examine the matrix $\mathbf{A}^T\mathbf{A}$ for the purposes of finding the SVD of $\mathbf{A}$.

We show that the eigenvalues of this matrix $\mathbf{A}^T\mathbf{A}$ are positive or zero:

Proposition (7.31). Let $\mathbf{A}$ be any matrix. Then the eigenvalues of $\mathbf{A}^T\mathbf{A}$ are positive or zero.

Proof – Exercises 7.5.

■

Example 7.34

(i) Find the singular values $\sigma_1 = \sqrt{\lambda_1}$ and $\sigma_2 = \sqrt{\lambda_2}$ of the matrix $\mathbf{A}$ given in the above Example 7.33.

(ii) Determine $\sigma_1 \mathbf{u}_1 = \mathbf{A}\mathbf{v}_1$ and $\sigma_2 \mathbf{u}_2 = \mathbf{A}\mathbf{v}_2$, where $\mathbf{v}_1$ and $\mathbf{v}_2$ are the normalized eigenvectors belonging to the eigenvalues of matrix $\mathbf{A}^T\mathbf{A}$.

Solution

(i) From Example 7.33 we have the eigenvalues $\lambda_1 = 6$ and $\lambda_2 = 3$. Taking the square root:

$$\sigma_1 = \sqrt{6} \quad \text{and} \quad \sigma_2 = \sqrt{3}$$

The singular values of matrix $\mathbf{A}$ are $\sqrt{6}$ and $\sqrt{3}$. Note that matrix $\mathbf{A}$ does *not* have eigenvalues because it is a non-square matrix but has singular values which are the square roots of the eigenvalues of $\mathbf{A}^T\mathbf{A}$.

(ii) Substituting $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -2 & 1 \end{pmatrix}$, eigenvector $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\sigma_1 = \sqrt{6}$ into $\sigma_1 \mathbf{u}_1 = \mathbf{A}\mathbf{v}_1$:

$$\sqrt{6}\mathbf{u}_1 = \mathbf{A}\mathbf{v}_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

Similarly for $\sigma_2 \mathbf{u}_2 = \mathbf{A}\mathbf{v}_2$ we have

$$\sqrt{3}\mathbf{u}_2 = \mathbf{A}\mathbf{v}_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

The matrix $\mathbf{A}$ transforms the two-dimensional eigenvectors $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to three-dimensional vectors $(1\ 1\ -2)^T$ and $(1\ \ 1\ \ 1)^T$ respectively. This transformation is $T : \mathbb{R}^2 \to \mathbb{R}^3$, such that $T(\mathbf{v}) = \mathbf{A}\mathbf{v}$ where $\mathbf{A}$ is the given 3 by 2 matrix.

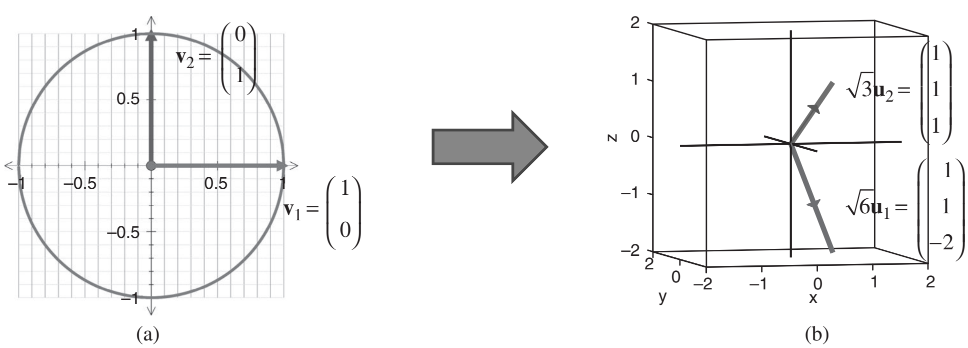


Figure 7.17

The transformation of the unit circle in Fig. 7.17(a) under the given matrix $\mathbf{A}$ is a two-dimensional ellipse in 3d space $\mathbb{R}^3$, but not illustrated above because of the limitation of software available.

Since the given $\mathbf{A}$ is a 3 by 2 matrix, by using the above formula (7.30) we find that $\mathbf{U}$ is a 3 by 3 matrix. This means that $\mathbf{U} = (\mathbf{u}_1 \;\; \mathbf{u}_2 \;\; \mathbf{u}_3)$.

In the above Example 7.34 we have found $\mathbf{u}_1$ and $\mathbf{u}_2$ but what is $\mathbf{u}_3$ equal to?

The vector $\mathbf{u}_3$ needs to be orthogonal (perpendicular) to both vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ because $\mathbf{U}$ is an orthogonal matrix. This means vector $\mathbf{u}_3$ must satisfy both $\mathbf{u}_1 \cdot \mathbf{u}_3 = 0$ and $\mathbf{u}_2 \cdot \mathbf{u}_3 = 0$. Let $\mathbf{u}_3 = (x \;\; y \;\; z)^T$. We can ignore the scalars because the vectors will be orthogonal (perpendicular) independently of their scalars. We need to solve

$$\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \quad \text{and} \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

[Remember, for orthogonal (perpendicular) vectors, the dot product is zero]

In matrix form, and solving by inspection, yields

$$\begin{pmatrix} 1 & 1 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ gives } x = 1, \; y = -1 \text{ and } z = 0$$

Normalizing the vector gives $\mathbf{u}_3 = \dfrac{1}{\sqrt{2}}(1 \; -1 \; 0)^T$. Hence

$$\mathbf{U} = \left(\mathbf{u}_1 \;\; \mathbf{u}_2 \;\; \mathbf{u}_3 \right) = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{6} & 1/\sqrt{3} & -1/\sqrt{2} \\ -2/\sqrt{6} & 1/\sqrt{3} & 0 \end{pmatrix}$$

Note that the column vectors $\left\{ \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \right\}$ of matrix $\mathbf{U}$ is a basis (axes) for $\mathbb{R}^3$ (Fig. 7.18).

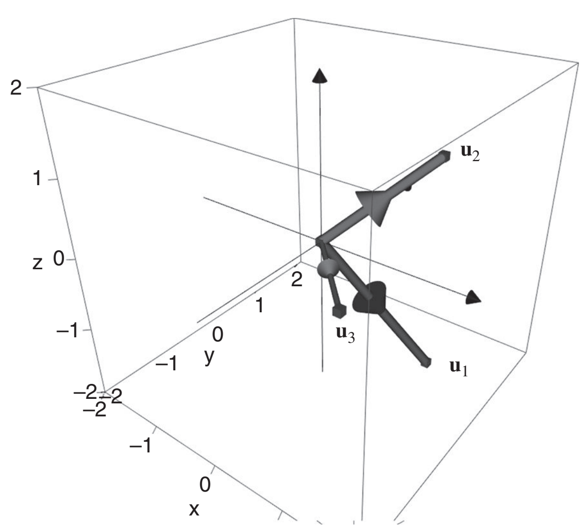


Figure 7.18

Example 7.35

Check that $\mathbf{A} = \mathbf{UDV}^T$ for the matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -2 & 1 \end{pmatrix}$ given in the above Example 7.33.

Solution

We need to break the given matrix $\mathbf{A}$ into $\mathbf{A} = \mathbf{UDV}^T$ where **U** is 3 by 3, $\mathbf{D}$ is 3 by 2 and **V** is 2 by 2.

What are the matrices $\mathbf{U}$, $\mathbf{D}$ *and* $\mathbf{V}$ *equal to?*

By the results of Example 7.34 and above we have:

$$\mathbf{U} = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3) = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{6} & 1/\sqrt{3} & -1/\sqrt{2} \\ -2/\sqrt{6} & 1/\sqrt{3} & 0 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{3} \\ 0 & 0 \end{pmatrix}$$

$$\text{and } \mathbf{V} = (\mathbf{v}_1 \quad \mathbf{v}_2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I} \quad [\text{identity matrix}]$$

By transposing the last matrix $\mathbf{V}^T = \mathbf{I}^T = \mathbf{I}$, let us check that this triple factorization actually works; that is $\mathbf{UDV}^T = \mathbf{UDI} = \mathbf{UD} = \mathbf{A}$:

$$\mathbf{UDV}^T = \mathbf{UD} = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{6} & 1/\sqrt{3} & -1/\sqrt{2} \\ -2/\sqrt{6} & 1/\sqrt{3} & 0 \end{pmatrix} \times \begin{pmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{3} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -2 & 1 \end{pmatrix} = \mathbf{A}$$

Hence a non-square matrix such as $\mathbf{A}$ can be broken into $\mathbf{UDV}^T$.

7.5.3 Proof of the singular value decomposition (SVD)

Singular value decomposition can be applied to any matrix $\mathbf{A}$.

Proposition (7.32). Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ be the eigenvectors of $\mathbf{A}^T\mathbf{A}$ such that they belong to the positive eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_k > 0$. Then

(i) for $j = 1, 2, 3, \ldots, k$ we have $\|\mathbf{Av}_j\| = \sigma_j$ where $\sigma_j = \sqrt{\lambda_j}$ is the singular value of $\mathbf{A}$.

(ii) $\{\mathbf{Av}_1, \mathbf{Av}_2, \ldots, \mathbf{Av}_k\}$ is an *orthogonal* set of vectors.

Note, the following:

(i) σ_j gives the size of the vector $\mathbf{Av}_j$, or the length of the semi-axis ($\mathbf{u}_j$) of the ellipse.

(ii) This part means that orthogonal (perpendicular) vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ are transformed to orthogonal (perpendicular) vectors $\{\mathbf{Av}_1, \mathbf{Av}_2, \ldots, \mathbf{Av}_k\}$ under the matrix $\mathbf{A}$.

Proof of (i).

By the above proposition (7.30): Let $\mathbf{A}$ be any matrix. Then $\mathbf{A}^T\mathbf{A}$ is a symmetric matrix.

$\mathbf{A}^T\mathbf{A}$ is a symmetric matrix, so we can orthogonally diagonalize this because of the spectral theorem (7.28): $\mathbf{A}$ is orthogonally diagonalizable $\Leftrightarrow$ $\mathbf{A}$ is a symmetric matrix.

This means that the eigenvectors belonging to the positive eigenvalues of $\mathbf{A}^T\mathbf{A}$, given by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are orthonormal (perpendicular unit) vectors. Consider $\|\mathbf{Av}_1\|^2$. Then

$$
\begin{aligned}
\|\mathbf{Av}_1\|^2 = \mathbf{Av}_1 \cdot \mathbf{Av}_1 = (\mathbf{Av}_1)^T\,\mathbf{Av}_1 \qquad & \left[\text{by (2.4) } \mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v}\right] \\
= \mathbf{v}_1^T\mathbf{A}^T\mathbf{Av}_1 \qquad & \left[\text{because } (\mathbf{XY})^T = \mathbf{Y}^T\mathbf{X}^T\right] \\
= \mathbf{v}_1^T\,(\lambda_1\mathbf{v}_1) \qquad & \left[\lambda_1 \text{ and } \mathbf{v}_1 \text{ are e.values and e.vectors of } \mathbf{A}^T\mathbf{A}\right. \\
& \left.\text{so } \mathbf{A}^T\mathbf{Av}_1 = \lambda_1\mathbf{v}_1\right] \\
= \lambda_1\mathbf{v}_1^T\mathbf{v}_1 \qquad & \\
= \lambda_1\,(\mathbf{v}_1\cdot\mathbf{v}_1) = \lambda_1 \qquad & \left[\text{because } \mathbf{v}_1 \text{ is normalized so } (\mathbf{v}_1\cdot\mathbf{v}_1) = \|\mathbf{v}_1\|^2 = 1\right]
\end{aligned}
$$

Taking the square root of this result $\|\mathbf{Av}_1\|^2 = \lambda_1$ gives

$$\|\mathbf{Av}_1\| = \sqrt{\lambda_1} = \sigma_1$$

Similarly for $j = 2, 3, \ldots, n$ we have $\left\|\mathbf{Av}_j\right\| = \sigma_j$. This completes our proof for part (i). ■

Proof of (ii).

Required to prove that the vectors in the set $\{\mathbf{Av}_1, \mathbf{Av}_2, \ldots, \mathbf{Av}_k\}$ are orthogonal (perpendicular). We prove that any two arbitrary *different* vectors $\mathbf{Av}_i$ and $\mathbf{Av}_j$ where $i \neq j$ in the set are orthogonal, which means that we need to show that $\mathbf{Av}_i \cdot \mathbf{Av}_j = 0$:

$$
\begin{aligned}
\mathbf{Av}_i \cdot \mathbf{Av}_j = (\mathbf{Av}_i)^T\,\mathbf{Av}_j = \mathbf{v}_i^T\mathbf{A}^T\mathbf{Av}_j & \\
= \mathbf{v}_i^T\lambda_j\mathbf{v}_j \qquad & \left[\lambda_j \text{ and } \mathbf{v}_j \text{ are e.values and e.vectors of } \mathbf{A}^T\mathbf{A}\right. \\
& \left.\text{so } \mathbf{A}^T\mathbf{Av}_j = \lambda_j\mathbf{v}_j\right] \\
= \lambda_j\left(\mathbf{v}_i^T\mathbf{v}_j\right) = \lambda_j\left(\mathbf{v}_i\cdot\mathbf{v}_j\right) = 0 \qquad & \left[\mathbf{v}_i \text{ and } \mathbf{v}_j \text{ are orthogonal so}\right. \\
& \left.\mathbf{v}_i\cdot\mathbf{v}_j = 0\right]
\end{aligned}
$$

We have $\mathbf{Av}_i \cdot \mathbf{Av}_j = 0$, therefore $\mathbf{Av}_i$ and $\mathbf{Av}_j$ are orthogonal (perpendicular). Hence the set of vectors $\{\mathbf{Av}_1, \mathbf{Av}_2, \ldots, \mathbf{Av}_k\}$ are orthogonal to each other. ■

Next we prove the main result in this section, which was stated above and also repeated here:

Singular Value Decomposition Theorem (7.30).

We can decompose any given matrix $\mathbf{A}$ of size m by n with positive singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > 0$ where $k \leq n$, into $\mathbf{UDV}^T$, that is

$$\mathbf{A} = \mathbf{UDV}^T$$

where $\mathbf{U}$ is an m by m *orthogonal* matrix, $\mathbf{D}$ is an m by n matrix and $\mathbf{V}$ is an n by n *orthogonal* matrix.

Proof.
The size of matrix $\mathbf{A}$ is m by n. This means that the transformation T given by matrix $\mathbf{A}$ is $T : \mathbb{R}^n \to \mathbb{R}^m$ such that $T(\mathbf{v}) = \mathbf{Av}$ (Fig. 7.19).

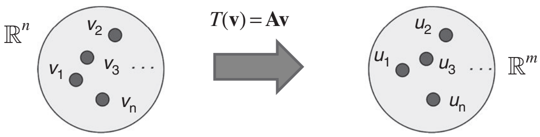


Figure 7.19

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ be the eigenvectors of $\mathbf{A}^T\mathbf{A}$ belonging to the positive eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$ respectively. Then by the above Proposition (7.32) part (ii)

$$\mathbf{Av}_1, \mathbf{Av}_2, \ldots, \mathbf{Av}_k \quad (^*)$$

is an orthogonal set of vectors. We want to convert this into an orthonormal set which means we need to normalize each of the vectors in the list (*). Convert the first vector $\mathbf{Av}_1$ into a unit vector $\mathbf{u}_1$ say, (length of 1):

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{Av}_1\|}\mathbf{Av}_1 = \frac{1}{\sigma_1}\mathbf{Av}_1 \qquad \left[\text{by (7.32) part(i) } \|\mathbf{Av}_j\| = \sigma_j\right]$$

Similarly converting the remaining vectors $\mathbf{Av}_2, \ldots, \mathbf{Av}_k$ into unit vectors we have

$$\mathbf{u}_j = \frac{1}{\sigma_j}\mathbf{Av}_j \text{ for } j = 2, 3, 4, \ldots, k$$

The vectors in $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ constitute an orthonormal set of vectors. We need to produce a matrix $\mathbf{U}$ which is of size m by m. If $k < m$ then the vectors in this set $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ are the first k vectors of the matrix $\mathbf{U}$. However, we need m vectors because

$$\mathbf{U} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k & \mathbf{u}_{k+1} & \cdots & \mathbf{u}_m \end{pmatrix}$$

We extend the above set S to an orthonormal set $S' = \left\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_m\right\}$. This S' is an orthonormal (perpendicular unit) basis for R^m. Let

$$\mathbf{U} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_m \end{pmatrix}$$

Multiplying the above result $\mathbf{u}_j = \frac{1}{\sigma_j}\mathbf{Av}_j$ by σ_j for $j = 1, 2, 3, \ldots, k$ gives

$$\sigma_j\mathbf{u}_j = \mathbf{Av}_j \quad (\dagger)$$

The remaining singular values are zero, that is $\sigma_{k+1} = \sigma_{k+2} = \cdots = \sigma_n = 0$. We have

$$\mathbf{A}\mathbf{v}_j = \sigma_j \mathbf{u}_j = 0\mathbf{u}_j \text{ for } j = k+1, \ldots, n \qquad (\dagger\dagger)$$

Collecting all these together we have

$$\mathbf{A}\mathbf{v}_1 = \sigma_1\mathbf{u}_1,\ \mathbf{A}\mathbf{v}_2 = \sigma_2\mathbf{u}_2, \ldots,\ \mathbf{A}\mathbf{v}_k = \sigma_k\mathbf{u}_k,\ \mathbf{A}\mathbf{v}_{k+1} = 0\mathbf{u}_{k+1}, \ldots, \mathbf{A}\mathbf{v}_n = 0\mathbf{u}_n$$

In matrix form we have

$$\begin{aligned}\mathbf{AV} &= \mathbf{A}\left(\mathbf{v}_1 \cdots \mathbf{v}_k \quad \mathbf{v}_{k+1} \cdots \mathbf{v}_n\right) && \left[\text{V is an n by n orthogonal matrix}\right]\\ &= \left(\mathbf{A}\mathbf{v}_1 \cdots \mathbf{A}\mathbf{v}_k \quad \mathbf{A}\mathbf{v}_{k+1} \cdots \mathbf{A}\mathbf{v}_n\right)\\ &= \left(\sigma_1\mathbf{u}_1 \cdots \sigma_k\mathbf{u}_k \quad 0\mathbf{u}_{k+1} \cdots 0\mathbf{u}_n\right) && \left[\text{by } (\dagger) \text{ and } (\dagger\dagger)\right]\end{aligned}$$

The scalars $\sigma_1, \sigma_2, \ldots$ in front of the $\mathbf{u}$'s can be placed in the diagonal-like matrix $\mathbf{D}$ and the vectors $\mathbf{u}$'s into the above matrix $\mathbf{U} = \left(\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_m\right)$.

Hence the above derivation

$$\mathbf{AV} = \left(\ \sigma_1\mathbf{u}_1 \quad \cdots \quad \sigma_k\mathbf{u}_k \quad 0\mathbf{u}_{k+1} \quad \cdots \quad 0\mathbf{u}_n\ \right)$$

can be written as:

$$\begin{aligned}\left(\sigma_1\mathbf{u}_1 \cdots \sigma_k\mathbf{u}_k \quad 0\mathbf{u}_{k+1} \cdots 0\mathbf{u}_n\right) &= \underbrace{\left(\mathbf{u}_1 \quad \mathbf{u}_2 \cdots \mathbf{u}_m\right)}_{=\mathbf{U}} \underbrace{\left(\begin{array}{ccc|c} \sigma_1 & & O & \\ & \ddots & & O \\ O & & \sigma_k & \\ \hline & O & & O \end{array}\right)}_{=\mathbf{D}}\\ &= \mathbf{UD} = \mathbf{AV}\end{aligned}$$

Since $\mathbf{V}$ is an orthogonal matrix so $\mathbf{V}^{-1} = \mathbf{V}^T$. Right multiplying both sides of $\mathbf{UD} = \mathbf{AV}$ by $\mathbf{V}^T$ gives us our required result $\mathbf{UDV}^T = \mathbf{A}$.

■

Proposition (7.33). Let $\mathbf{A}$ be an m by n matrix with factorization given by $\mathbf{A} = \mathbf{UDV}^T$. Let matrix $\mathbf{A}$ have $k \le n$ positive singular values. Then we have the following:

(a) The set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ form an orthonormal basis for the column space of matrix $\mathbf{A}$.

(b) The set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ form an orthonormal basis for the row space of matrix $\mathbf{A}$.

(c) The set of vectors $\left\{\mathbf{v}_{k+1}, \mathbf{v}_{k+2}, \ldots, \mathbf{v}_n\right\}$ form an orthonormal basis for the null space of matrix $\mathbf{A}$.

Proof – Exercises 7.5.

■

Summary

We can break any matrix $\mathbf{A}$ into a triple factorization $\mathbf{A} = \mathbf{U} \times \mathbf{D} \times \mathbf{V}^T$.

EXERCISES 7.5

(Brief solutions at end of book. Full solutions available at <http://www.oup.co.uk/companion/singh>.)

1. Determine the matrices $\mathbf{U}$, $\mathbf{D}$ and $\mathbf{V}$ such that $\mathbf{A} = \mathbf{UDV}^T$ for the following:

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

(d) $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ (e) $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix}$ (f) $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 3 & 3 \end{pmatrix}$

2. Prove Proposition (7.31).
3. Let matrix $\mathbf{A}$ have k positive singular values. Show that the rank of matrix $\mathbf{A}$ is k.
   Hint: You may use the following result:
   If matrix $\mathbf{X}$ is invertible then $rank(\mathbf{XA}) = rank(\mathbf{A})$ and $rank(\mathbf{AX}) = rank(\mathbf{A})$.
4. Prove Proposition (7.33).
5. Prove that the singular values $\sigma_1, \sigma_2, \ldots, \sigma_n$ of an m by n matrix $\mathbf{A}$ are unique.
6. Prove that the column vectors of the orthogonal matrix $\mathbf{U}$ in $\mathbf{A} = \mathbf{UDV}^T$ are the eigenvectors of $\mathbf{AA}^T$.
7. Prove that the singular values of $\mathbf{A}$ and $\mathbf{A}^T$ are identical.
8. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation given by $T(\mathbf{x}) = \mathbf{Ax}$, where $\mathbf{A}$ is an m by n matrix. Let the singular value decomposition $\mathbf{A} = \mathbf{UDV}^T$ with $k \leq n$ positive singular values of matrix $\mathbf{A}$. Prove the following results:
   (a) The set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ form an orthonormal basis for the range of T.
   (b) The set of vectors $\{\mathbf{v}_{k+1}, \mathbf{v}_{k+2}, \ldots, \mathbf{v}_n\}$ form an orthonormal basis for the kernel of T.

MISCELLANEOUS EXERCISES 7

(Brief solutions at end of book. Full solutions available at <http://www.oup.co.uk/companion/singh>.)

In this exercise you may check your numerical answers using MATLAB.

7.1. If $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}$.

(a) Find all the eigenvalues of $\mathbf{A}$.

(b) Find a non-singular matrix $\mathbf{Q}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{Q}^{-1}\mathbf{AQ} = \mathbf{D}$ (that is $\mathbf{A} = \mathbf{QDQ}^{-1}$).

(c) For the matrix $\mathbf{A}$ find $\mathbf{A}^5$.

Purdue University, USA

7.2. Let $\mathbf{A} = \begin{pmatrix} 4 & 5 \\ -3 & -4 \end{pmatrix}$. Compute $\mathbf{A}^{1\,000\,001}$.

Harvey Mudd College, California, USA

7.3. Find the eigenvalues and bases for the corresponding eigenspaces for the matrix

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

Harvey Mudd College, California, USA

7.4. Answer each of the following by filling in the blank. No explanation is necessary.

(a) Let $\mathbf{A}$ be an invertible $n \times n$ matrix with eigenvalue λ. Then ______ is an eigenvalue of $\mathbf{A}^{-1}$.

(b) An $n \times n$ matrix $\mathbf{A}$ is diagonalizable if and only if $\mathbf{A}$ has n ___.

Illinois State University, USA

7.5. Let $\mathbf{A}$ be the matrix

$$\begin{pmatrix} 7 & 5 \\ 3 & -7 \end{pmatrix}$$

(a) Find matrices S and Λ such that $\mathbf{A}$ has factorization of the form

$$\mathbf{A} = \mathbf{S}\Lambda\mathbf{S}^{-1}$$

where $\mathbf{S}$ is invertible and Λ is diagonal: $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2)$.

(b) Find a matrix $\mathbf{B}$ such that $\mathbf{B}^3 = \mathbf{A}$: (Hint: first find such a matrix for Λ. Then use the formula above.)

Massachusetts Institute of Technology, USA

7.6. (a) Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix}$$

(i) Find the eigenvalues and eigenvectors of $\mathbf{A}$.

(ii) Find a matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP}$ is diagonal.

(iii) Find the eigenvalues and the determinant of $\mathbf{A}^{2008}$.

(b) Consider the matrix

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Is $\mathbf{B}$ diagonalisable? Justify your answer.

Loughborough University, UK

7.7. Do **one** of the following.

(a) Is the matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 10 & 2 \end{bmatrix}$ diagonalizable? If not, explain why not. If so, find an invertible matrix $\mathbf{S}$ for which $\mathbf{S}^{-1}\mathbf{AS}$ is diagonal.

(b) The matrices $\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$ are similar. Exhibit a matrix $\mathbf{S}$ for which $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$.

University of Puget Sound, USA

7.8. Consider the matrix

$$\mathbf{B} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Do there exist matrices $\mathbf{P}$ and $\mathbf{D}$ such that $\mathbf{B} = \mathbf{PDP}^T$ where $\mathbf{P}^{-1} = \mathbf{P}^T$ and $\mathbf{D}$ is a diagonal matrix? Why? If these matrices exist then write down a possible $\mathbf{P}$ and the corresponding $\mathbf{D}$.

New York University, USA

7.9. Consider the 3×3 matrix $\mathbf{A} = \begin{pmatrix} 4 & 2 & -2 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix}$.

Already computed are the eigenpairs $\left(2, \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}\right), \quad \left(4, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right)$.

(a) Find the remaining eigenpairs of $\mathbf{A}$.

(b) Display an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{AP} = \mathbf{PD}$.

University of Utah, USA (part question)

7.10. Prove that, if λ_1 and λ_2 are *distinct* eigenvalues of a symmetric matrix $\mathbf{A}$, then the corresponding eigenspaces are orthogonal.

Harvey Mudd College, California, USA

7.11. (a) Find the eigenvalues and corresponding eigenvectors of $\mathbf{A} = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}$.

(b) Determine which of the following vectors

$$\vec{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \vec{w} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 2 \end{bmatrix}$$

is an eigenvector of the matrix $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$ and find the corresponding eigenvalue.

University of New Brunswick, Canada

7.12. (a) Find the eigenvalues and corresponding eigenvectors for

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

(b) Is $\mathbf{A}$ diagonalizable? Give reasons for your answer.

University of New Brunswick, Canada

7.13. (a) (i) Show that the eigenvalues and corresponding eigenvectors of

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

are given by

$$\lambda_1 = 0, \mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \lambda_2 = 0, \mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \text{ and } \lambda_3 = 3, \mathbf{w} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

(ii) Find an orthogonal matrix $\mathbf{Q}$ which diagonalises the matrix $\mathbf{A}$.

(iii) Use the Cayley–Hamilton theorem or otherwise to find $\mathbf{A}^3$.

(b) Let $\mathbf{B}$ be an n by n real matrix. Prove that $\mathbf{B}$ and $\mathbf{B}^T$ have the same eigenvalues.

University of Hertfordshire, UK

7.14. (a) Define what is meant by an eigenvector and an eigenvalue for a real $n \times n$ matrix.

(b) Let $\mathbf{A} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 0 \end{pmatrix}$. Show that the vector $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is an eigenvector for $\mathbf{A}$. What is the corresponding eigenvalue?

(c) Show that the matrix $\mathbf{A}$ given in (b) is diagonalizable and hence find an invertible 3×3 matrix $\mathbf{P}$ (and $\mathbf{P}^{-1}$) such that $\mathbf{P}^{-1}\mathbf{AP}$ is diagonal.

City University, London, UK

7.15. (a) Let $\mathbf{A}$ be the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 2 \\ 8 & 11 & -8 \\ 4 & 4 & -1 \end{pmatrix}$$

By finding a basis of eigenvectors, determine an invertible matrix $\mathbf{P}$ and its inverse $\mathbf{P}^{-1}$ such that $\mathbf{P}^{-1}\mathbf{AP}$ is diagonal.

(b) State the Cayley–Hamilton theorem and verify it for the above matrix.

City University, London, UK

7.16. Let $\mathbf{A} \in M_4(\mathbb{R})$ be the following matrix

$$\mathbf{A} = \begin{pmatrix} 0 & -3 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & -9 \\ 1 & -1 & 1 & 10 \end{pmatrix}$$

(a) Show that the characteristic polynomial of $\mathbf{A}$ is given by

$$P_A(t) = \left(t^2 + t + 3\right)\left(t^2 - 10t + 9\right)$$

(b) Compute all real eigenvalues of $\mathbf{A}$. Choose one real eigenvalue λ of $\mathbf{A}$ and find a basis of its eigenspace E_λ.

Jacobs University, Germany

7.17. Find the eigenvalues of the matrix $\mathbf{A} = \begin{pmatrix} 0 & 1 & 2 & 0 & 0 \\ -1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & -3 & 0 \\ 1 & 2 & 3 & 4 & 0 \end{pmatrix}$. To save time, **do not** find the eigenvectors.

University of Utah, USA

7.18. (Calculators are **not allowed** for this question). Let $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

(a) Find the eigenvalues for $\mathbf{A}$.

(b) For $\mathbf{A}$, find an eigenvector for each of the eigenvalues. To make it easier to grade, choose eigenvectors with integer coordinates where the integers are as small as possible.

(c) Use your eigenvectors to make a basis of $\mathbb{R}^3$. Choose the first basis vector to be the eigenvector associated with the largest eigenvalue and the third basis vector to be the eigenvector associated with the smallest eigenvalue. Call this basis β. We have a linear transformation given to us by $\mathbf{A}$. What is the matrix $\mathbf{D}$ when we use coordinates from this new basis, $\mathbf{D} : [x]_\beta \to [\mathbf{A}x]_\beta$?

(d) We know that there is a matrix $\mathbf{S}$ such that $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{D}$. Find $\mathbf{S}$.

(e) Find $\mathbf{S}^{-1}$.

(f) Modify the basis β so that you can orthogonally diagonalize $\mathbf{A}$.

(g) Find the new $\mathbf{S}$ needed to orthogonally diagonalize $\mathbf{A}$.

(h) Find the inverse of this last $\mathbf{S}$.

Johns Hopkins University, USA

7.19. (a) If $a \neq c$, find the eigenvalue matrix Λ and eigenvector matrix $\mathbf{S}$ in

$$\mathbf{A} = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} = \mathbf{S}\Lambda\mathbf{S}^{-1}$$

(b) Find the *four entries* in the matrix $\mathbf{A}^{1000}$.

Massachusetts Institute of Technology, USA

7.20. (a) Define what is meant by saying that an $n \times n$ matrix $\mathbf{A}$ is *diagonalizable*.

(b) Define what is meant by an eigenvector and an eigenvalue of an $n \times n$ matrix $\mathbf{A}$.

(c) Define the algebraic multiplicity and the geometric multiplicity of an eigenvalue λ_0 of a matrix $\mathbf{A}$.

[Note that geometric multiplicity of an eigenvalue is **not** discussed in this book but have a look at other sources to find the definition of this.]

(d) Suppose that a 3×3 matrix $\mathbf{A}$ has characteristic polynomial $(\lambda - 1)(\lambda - 2)^2$. What is the algebraic multiplicity of each eigenvalue of $\mathbf{A}$? State a necessary and sufficient condition for the diagonalizability of $\mathbf{A}$ using geometric multiplicity.

(e) Each of the following matrices has characteristic polynomial $(\lambda - 1)(\lambda - 2)^2$.

(i) $\begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & 2 \\ 1 & 0 & 3 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ -1 & -1 & 3 \end{bmatrix}$

Determine whether each of the matrices is diagonalizable. In each case, if the matrix is diagonalizable, find a diagonalizing matrix.

University of Manchester, UK

7.21. (a) Define what it means for vectors $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ in $\mathbb{R}^3$ to be orthonormal.

(b) Apply the Gram–Schmidt process to the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 5 \\ 4 \\ 6 \end{bmatrix}$$

in $\mathbb{R}^3$ to find an orthonormal set in $\mathbb{R}^3$.

(c) Consider the matrix

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

(i) Obtain the characteristic polynomial and find the eigenvalues of $\mathbf{A}$.

(ii) Find a complete set of linearly independent eigenvectors corresponding to each eigenvalue.

(d) Let the non-zero vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$ be an orthogonal set. Prove that they are linearly independent.

University of Southampton, UK

7.22. (a) Define the terms orthogonal and orthonormal applied to a set of vectors in a vector space on which an inner product is defined.

(b) State the relationship between an *orthogonal matrix* and its transpose. Prove that the set of columns of an orthogonal matrix forms an orthonormal set of vectors.

(c) (i) Show that $(1, 0, 1, 0)$ and $(1, 0, -1, 0)$ are eigenvectors of the matrix

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

and find the corresponding eigenvalues.

(ii) Find the two other eigenvalues and corresponding eigenvectors of $\mathbf{A}$.

(iii) Find matrices $\mathbf{P}, \mathbf{Q}, \Lambda$ such that $\mathbf{PQ} = \mathbf{I}$ and $\mathbf{PAQ} = \Lambda$ is diagonal.

Queen Mary, University of London, UK

[The wording has been modified for the next two questions so that it is compatible with the main text. You will need to look at the website material for chapter 7 to attempt these questions.]

7.23. Express the following quadratic in its diagonal form:

$$3x^2 + 4xy + 6y^2 = aX^2 + bY^2$$

Write X and Y in terms of x and y.

Columbia University, USA

7.24. Express the following quadratic in its diagonal form:

$$2xy + 4xz + 4yz + 3z^2 = aX^2 + bY^2 + cZ^2$$

Write X, Y and Z in terms of x, y and z.

Columbia University, USA

Brief Solutions

SOLUTIONS TO EXERCISES 1.1

1. (a) Linear (b) Not linear (c) Not linear (d) Not linear
(e) Linear (f) Linear (g) Linear (h) Linear
(i) Linear (j) Not linear (k) Linear (l) Linear
(m) Linear

2. (a) $x = 1,\ y = 1$ (b) $x = 1,\ y = -1$ (c) $x = -29,\ y = -31$
(d) $x = \frac{3}{4},\ y = \frac{1}{4}$ (e) $x = \frac{3}{4\pi},\ y = -\frac{1}{4}$ (f) $x = \frac{1}{e},\ y = -\frac{1}{e}$

3. (a) $x = y = z = 1$ (b) $x = -2,\ y = 1$ and $z = -3$
(c) $x = 3,\ y = -4$ and $z = \frac{1}{2}$ (d) $x = 3,\ y = -3$ and $z = 2$

4. (a) Unique (b) Infinite (c) No solution (d) No solution
(e) No solution (f) Unique

5. (a) Unique (b) Unique (c) No solution

SOLUTIONS TO EXERCISES 1.2

1. (a) $x = 6$ and $y = 1$ (b) $x = 2,\quad y = -4$ and $z = -3$
(c) $x = 1,\ y = 3$ and $z = 2$ (d) $x = 1/2,\quad y = 1/8$ and $z = 1/4$
(e) $x = -6.2,\quad y = 30$ and $z = -10$

2. (a) $x = 1,\quad y = 4, z = 1$ (b) $x = -1, y = 2, z = -1$ (c) $x = y = z = 1$

3. (a) $x = 1,\quad y = 2$ and $z = 3$ (b) $x = -2,\quad y = 1$ and $z = -1$
(c) $x = -1/2,\quad y = 1$ and $z = -2$ (d) $x = -3,\quad y = 1$ and $z = 1/2$

SOLUTIONS TO EXERCISES 1.3

2. (e) 1 (f) 1 (g) 2 (h) 5

3. (a) 15 (b) 15 (c) 14 (d) 30

5. (a) $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ (b) $\begin{pmatrix} -4 \\ 2 \end{pmatrix}$ (c) $\begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$ (d) $\begin{pmatrix} -5/2 \\ 3/2 \end{pmatrix}$ (e) $\begin{pmatrix} 0 \\ 2/3 \end{pmatrix}$

6. (a) $\begin{pmatrix} 0 \\ 3 \end{pmatrix}$ (b) $\begin{pmatrix} 0 \\ 3/2 \end{pmatrix}$ (c) $\begin{pmatrix} 2 \\ -1/2 \end{pmatrix}$ (d) $\begin{pmatrix} -2 \\ 1/2 \end{pmatrix}$

7. $x=-\frac{1}{2},\ y=-\frac{9}{2}$

9. (a) $\begin{pmatrix}4\\4\\10\end{pmatrix}$ (b) $\begin{pmatrix}-15\\25\\40\end{pmatrix}$ (c) $\begin{pmatrix}36\\4\\28\end{pmatrix}$ (d) $\begin{pmatrix}-24\\8\\2\end{pmatrix}$ (e) $\begin{pmatrix}-13\\-21\\-48\end{pmatrix}$

10. (a) $\begin{pmatrix}-10\\1\\12\end{pmatrix}$ (b) $\begin{pmatrix}-8\\3\\-4\end{pmatrix}$ (c) $\begin{pmatrix}-21\\7\\6\end{pmatrix}$ (d) $\begin{pmatrix}17\\-11\\26\end{pmatrix}$

11. $x=7,\ y=-11$ and $z=1$

13. (a) $\begin{pmatrix}2\\0\\8\\3\end{pmatrix}$ (b) $\begin{pmatrix}-4\\6\\-4\\-3\end{pmatrix}$ (c) $\begin{pmatrix}-7\\4\\-5\\4\end{pmatrix}$ (d) $\begin{pmatrix}x-1\\y+6\\z-1\\a-6\end{pmatrix}$

$x=-2, y=0, z=-8$ and $a=-3$

SOLUTIONS TO EXERCISES 1.4

1. (a) $\begin{pmatrix}7&1\\8&2\end{pmatrix}$ (b) $\begin{pmatrix}7&1\\8&2\end{pmatrix}$ (c) $\begin{pmatrix}18&-3\\15&9\end{pmatrix}$

(d) $\begin{pmatrix}18&-3\\15&9\end{pmatrix}$ (e) $\begin{pmatrix}15&4\\19&3\end{pmatrix}$ (h) $\begin{pmatrix}1\\-4\end{pmatrix}$

(i) $\begin{pmatrix}-7\\-2\end{pmatrix}$ (k) $\begin{pmatrix}17\\-8\end{pmatrix}$. Parts (f), (g) and (j) **cannot** be evaluated.

2. (a) $\begin{pmatrix}0&0&0\\0&0&0\end{pmatrix}$ (b) **A** (c) $\begin{pmatrix}42\\73\\81\end{pmatrix}$

(e) $\begin{pmatrix}12&10&8\\9&5&2\\3&3&6\end{pmatrix}$ (f) $\begin{pmatrix}12&10&8\\9&5&2\\3&3&6\end{pmatrix}$ (h) $\begin{pmatrix}1\\2\\3.5\end{pmatrix}$

(i) $\begin{pmatrix}44&44&3\\72&71&1\\29&12&-43\end{pmatrix}$ (j) $\begin{pmatrix}111&41&127\\24&-13&-20\\-1&-12&-26\end{pmatrix}$ (k) $\begin{pmatrix}-67&3&-124\\48&84&21\\30&24&-17\end{pmatrix}$

(d) (g) and (l) Impossible

3. (a) $\begin{pmatrix}2&4\\3&9\end{pmatrix}$ (b) $\begin{pmatrix}6&7\\2&3\end{pmatrix}$ (c) $\begin{pmatrix}a&b\\c&d\end{pmatrix}$ (d) $\begin{pmatrix}2&3&6\\1&4&5\\0&9&7\end{pmatrix}$

The result is always the first (or left hand) matrix.

4. (a) $\begin{pmatrix}1&0\\0&1\end{pmatrix}$ (b) $\begin{pmatrix}1&0\\0&1\end{pmatrix}$ (c) $\begin{pmatrix}1&0\\0&1\end{pmatrix}$

5. and **6.** $\begin{pmatrix}0&0\\0&0\end{pmatrix}$ and $\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}$. Multiplying two non-zero matrices gives a zero matrix.

7. $\begin{pmatrix} a_{\boxed{11}} & a_{\boxed{12}} & a_{\boxed{13}} & a_{\boxed{14}} \\ a_{\boxed{21}} & a_{\boxed{22}} & a_{\boxed{23}} & a_{\boxed{24}} \end{pmatrix}$

8. (a) $\mathbf{A}^2 = \mathbf{A}^3 = \mathbf{A}^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{x}_n = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\mathbf{x} = \mathbf{x}$

(b) $\mathbf{A}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, $\mathbf{A}^3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $\mathbf{A}^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{x}_n = \mathbf{A}^r\mathbf{x}$ where r is the reminder after dividing n by 4. If the reminder $r = 0$ then $\mathbf{A}^r\mathbf{x} = \mathbf{x}$.

(c) $\mathbf{A}^2 = \mathbf{A}^3 = \mathbf{A}^4 = \mathbf{A}$ and $\mathbf{x}_n = \mathbf{A}\mathbf{x}$.

9. $\mathbf{A} = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 2 & 4 & 4 \end{pmatrix}$ (a) $\begin{pmatrix} -3 & -2 & -2 & -3 \\ 2 & 2 & 4 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & 6 & 6 & 3 \\ 6 & 6 & 12 & 12 \end{pmatrix}$

(c) $\begin{pmatrix} 2 & 2 & 4 & 4 \\ -1 & -2 & -2 & -1 \end{pmatrix}$ (d) $\begin{pmatrix} -1 & -2 & -2 & -1 \\ 2 & 2 & 4 & 4 \end{pmatrix}$

10. $\mathbf{AF} = \begin{pmatrix} 1.2 & 1.6 & 2.6 & 2.52 & 1.92 & 1.8 & 2.4 & 2.32 & 1.72 & 1.6 \\ 1 & 3 & 3 & 2.6 & 2.6 & 2 & 2 & 1.6 & 1.6 & 1 \end{pmatrix}$

11. (a) $x = 0, \quad y = 0$ (b) $x = 0, \quad y = 0$

(c) $x = -4r$, $y = r$ where r is any real number.

12. (i) $\mathbf{u} - 2\mathbf{v}$ (ii) $\frac{1}{4}(5\mathbf{v} - 2\mathbf{u})$ (iii) No (iv) $2\mathbf{u} - 7\mathbf{v}$

SOLUTIONS TO EXERCISES 1.5

1. (a) $\begin{pmatrix} 1 & 1 & 12 \\ 2 & -10 & 13 \end{pmatrix}$ (b) $\begin{pmatrix} 10 & 6 & 20 \\ -4 & -11 & 19 \end{pmatrix}$ (c) $\begin{pmatrix} 10 & 6 & 20 \\ -4 & -11 & 19 \end{pmatrix}$

(d) $\begin{pmatrix} 10 & 6 & 20 \\ -4 & -11 & 19 \end{pmatrix}$ (e) $\begin{pmatrix} 9 & 5 & 8 \\ -6 & -1 & 6 \end{pmatrix}$ (f) $\begin{pmatrix} -5 & 15 & 35 \\ 10 & -45 & 30 \end{pmatrix}$

(g) Same as part (f). (h) $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

2. (a) $\begin{pmatrix} -5 & 1 & 2 \\ -1 & 3 & -2 \end{pmatrix}$ (b) $\begin{pmatrix} 4 & -3 & -5 \\ -3 & -8 & -4 \end{pmatrix}$

(c) $\begin{pmatrix} 9 & -4 & -7 \\ -2 & -11 & -2 \end{pmatrix}$ and $\begin{pmatrix} 11 & 0 & -1 \\ 6 & -1 & 10 \end{pmatrix}$

3. (a) $\begin{pmatrix} -8 & 8 & 0 \\ 12 & 5 & -26 \end{pmatrix}$ (b) Impossible (c) $\begin{pmatrix} -53 & 52 & -41 \\ -77 & 98 & -116 \end{pmatrix}$

(d) $\begin{pmatrix} -53 & 52 & -41 \\ -77 & 98 & -116 \end{pmatrix}$ (e) Impossible (f) Impossible (g) $\mathbf{O}_{33}$ (h) $\mathbf{O}_{33}$

(i) $\mathbf{O}_{33}$ (j) Impossible (k) Impossible

4. We have $\mathbf{AI} = \mathbf{IA} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. Note, that $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$.

5. $(\mathbf{AB})_{11} = 27,\ (\mathbf{AB})_{12} = 30,\ (\mathbf{AB})_{21} = 60,\ (\mathbf{AB})_{22} = 66$ and $\mathbf{AB} = \begin{pmatrix} 27 & 30 \\ 60 & 66 \end{pmatrix}$.

6. (a) $\begin{pmatrix} 72 & -216 & -360 \\ -288 & -72 & 504 \end{pmatrix}$ (b) $\begin{pmatrix} 72 & -216 & -360 \\ -288 & -72 & 504 \end{pmatrix}$

(c) $\mathbf{O}_{23}$ (d) $\mathbf{O}_{23}$ (e) $\mathbf{B}$ (f) $\mathbf{B}$ (g) $\mathbf{A}$ (h) $\mathbf{A}$

7. $\lambda = 3$

8. $\mathbf{A} = \mathbf{A}^2 = \mathbf{A}^3 = \mathbf{A}^4 = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, $\mathbf{A}^n = \mathbf{A}$. Also $\mathbf{x}_n = \mathbf{A}\mathbf{x}$.

9. (a) $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}$; (b) $\mathbf{A} = \mathbf{O},\ \mathbf{B} \neq \mathbf{O}$

10. $\begin{pmatrix} 0.65 \\ 0.35 \end{pmatrix}, \begin{pmatrix} 0.635 \\ 0.365 \end{pmatrix}, \begin{pmatrix} 0.636 \\ 0.364 \end{pmatrix}, \begin{pmatrix} 0.636 \\ 0.364 \end{pmatrix}$, and $\begin{pmatrix} 0.636 \\ 0.364 \end{pmatrix}$.

For large k we have $\mathbf{p}_k = \mathbf{T}^k\mathbf{p} = \begin{pmatrix} 0.636 \\ 0.364 \end{pmatrix}$.

SOLUTIONS TO EXERCISES 1.6

1. (a) $\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & -1 \\ 2 & -2 \\ 3 & -3 \end{pmatrix}$ (c) $\begin{pmatrix} -1 \\ 5 \\ 9 \\ 100 \end{pmatrix}$

(d) $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ (e) $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ (f) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

2. (a) $\begin{pmatrix} 35 & -35 \\ 16 & -66 \end{pmatrix}$ (b) $\begin{pmatrix} 28 & -36 & 16 \\ 45 & -30 & 35 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & -5 \\ 5 & 0 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & -5 & -5 \\ 5 & 0 & -1 \\ 5 & 1 & 0 \end{pmatrix}$ (e) $\begin{pmatrix} -6 & 3 & 1 \\ 8 & 9 & -2 \\ 6 & -1 & 5 \end{pmatrix}$ (f) $\begin{pmatrix} -8 & 12 \\ 2 & 10 \end{pmatrix}$

(g) $\begin{pmatrix} 4 & -1 \\ 4 & -5 \\ 13 & 8 \end{pmatrix}$ (h) $\begin{pmatrix} 4 & 4 & 13 \\ -1 & -5 & 8 \end{pmatrix}$ (i) $\begin{pmatrix} 4 & 4 & 13 \\ -1 & -5 & 8 \end{pmatrix}$

(j) $\begin{pmatrix} -27 & 8 & -9 \\ -58 & 32 & -26 \end{pmatrix}$ (k) $\begin{pmatrix} 6 & -8 & -6 \\ -3 & -9 & 1 \\ -1 & 2 & -5 \end{pmatrix}$ (l) $\begin{pmatrix} 6 & -8 & -6 \\ -3 & -9 & 1 \\ -1 & 2 & -5 \end{pmatrix}$

(m) $\begin{pmatrix} 22 & 6 \\ 1 & 31 \end{pmatrix}$ (n) $\begin{pmatrix} 22 & 6 \\ 1 & 31 \end{pmatrix}$

3. (a) 32 (b) 32 (c) 32

15. $x=-1$, $y=-1$ and $z=1$

16. $x=\sin(\theta)+\cos(\theta)$, $y=\cos(\theta)-\sin(\theta)$

SOLUTIONS TO EXERCISES 1.7

1. (a) $x=3/4,\ y=13/4$ and $z=-11/4$
(b) $x=0,\ y=0$ and $z=0$
(c) No Solution (d) $x=3t,\ y=2-2t$ and $z=t$
(e) $x=-7t,\ y=-6t,\ z=-t,\ w=t$
(f) $x=11t-20/3$, $y=7-8t$, $z=-1/3$, $w=t$
(g) $x=\dfrac{1}{20}(33-144t-20s)$, $y=\dfrac{1}{5}(19t-3)$, $z=t$ and $w=s$.
(h) No Solution
(i) $x=\dfrac{1}{9}(71t-18s-47)$, $y=5t-3$, $z=\dfrac{1}{3}(26-32t)$, $w=s$ and $u=t$
(j) $x=-(1+2t+3s)$, $y=\dfrac{1}{9}(15t-6s-13)$, $z=\dfrac{1}{27}(14+6t+3s)$,
$w=t$ and $u=s$.

2. (a) $x=-9+10t-s$, $y=s$, $z=-7+7t$ and $w=t$
(b) $x_1=-p-q-r,\quad x_2=p,\ x_3=q,\ x_4=r,\ x_5=0$ and $x_6=s$
(c) $x_1=2-3r-6s-2t$, $x_2=r,\quad x_3=1-3s,\ x_4=0,\quad x_5=s,\ x_6=t$

3. $x=4+2s$, $y=6$ and $z=s$

8. $x_1=-10+s+t$, $x_2=7-s$, $x_3=9-t$, $x_4=15-s-t$, $x_5=s$ and $x_6=t$

9. $x_1=-10+s+t$, $x_2=7-s$, $x_3=9-t$, $x_4=15-s-t$, $x_5=s$ and $x_6=t$,
$x_7=8-s-t$, $x_8=s$ and $x_9=t$.

10. $x_1=-2+p+r-t,\ x_2=-7-p+s+t,\ x_3=p,\ x_4=3+q-r+t$,
$x_5=25-q-s-t,\ x_6=q,\ x_7=r,\ x_8=s$ and $x_9=t$

11. $x=5.41-0.095t$, $y=5.41-0.095t$, $z=3.67-0.067t$ and t is free.

SOLUTIONS TO EXERCISES 1.8

1. Matrices **B**, **C**, **D** and **G** are elementary matrices.

2. (a) $\mathbf{E}_1^{-1}=\begin{pmatrix}-1&0\\0&1\end{pmatrix}$ (b) $\mathbf{E}_2^{-1}=\begin{pmatrix}0&1\\1&0\end{pmatrix}$

(c) $\mathbf{E}_3^{-1}=\begin{pmatrix}1&0\\0&-1/2\end{pmatrix}$ (d) $\mathbf{E}_4^{-1}=\begin{pmatrix}-1/5&0&0\\0&1&0\\0&0&1\end{pmatrix}$

(e) $\mathbf{E}_5^{-1}=\begin{pmatrix}1&0&0\\0&-1\big/\sqrt{2}&0\\0&0&1\end{pmatrix}$ (f) $\mathbf{E}_6^{-1}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1/\pi\end{pmatrix}$

3. (a) $\begin{pmatrix} a & b & c \\ -d & -e & -f \\ g & h & i \end{pmatrix}$ (b) $\begin{pmatrix} g & h & i \\ d & e & f \\ a & b & c \end{pmatrix}$

(c) $\begin{pmatrix} ka & kb & kc \\ d & e & f \\ g & h & i \end{pmatrix}$ (d) $\begin{pmatrix} a & b & c \\ d & e & f \\ -\frac{g}{k} & -\frac{h}{k} & -\frac{i}{k} \end{pmatrix}$

4. (a) $\frac{1}{6}\begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}$ (b) $-\frac{1}{28}\begin{pmatrix} 1 & 5 \\ 6 & 2 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & 1 & -1 \\ -1 & 1 & -2 \\ 0 & 0 & -1 \end{pmatrix}$ (e) $\frac{1}{4}\begin{pmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix}$

(f) $\frac{1}{15}\begin{pmatrix} 3 & -10 & 1 \\ 0 & 10 & 5 \\ 3 & -5 & -4 \end{pmatrix}$ (i) $\frac{1}{5}\begin{pmatrix} 131 & -2 & 8 & 15 \\ -43 & 1 & -9 & 0 \\ -25 & 0 & -5 & 0 \\ -30 & 0 & 0 & -5 \end{pmatrix}$

(c) Non-invertible (g) Non-invertible
(h) Non-invertible

5. (a) $x = 1/3,\ y = 4/3$ (b) $x = 1/14$ and $y = -4/7$
(c) No unique solution (d) $x = -2,\ y = -17$ and $z = -5$
(e) $x = 8,\ y = 11$ and $z = 7$ (f) $x = -2,\ y = 4$ and $z = -3$
(g) $x = 222.6,\ y = -66.8,\ z = -39$ and $w = -53$

7. $\mathbf{p} = (220.04 \quad 277.58 \quad 235.40)^T$

SOLUTIONS TO MISCELLANEOUS EXERCISES 1

1.1. (a) $-16\begin{pmatrix} 4 & 5 \\ 4 & 5 \end{pmatrix}$ (b) $-4\begin{pmatrix} 15 & 17 \\ 19 & 21 \end{pmatrix}$ This is because **AB** does not equal **BA**.

1.2. Error in first line because $(\mathbf{AB})^{-1} \neq \mathbf{A}^{-1}\mathbf{B}^{-1}$ and error in line 2 because matrix multiplication is **not commutative**.

1.3. (a) $\mathbf{A}^2 = \mathbf{I}$ and $\mathbf{A}^3 = \mathbf{A}$ (b) $\mathbf{A}^{-1} = \mathbf{A}$ and $\mathbf{A}^{2004} = \mathbf{I}$

1.4. $\mathbf{A}^t = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ -1 & 1 & 3 \end{pmatrix}$. The operations **AB**, **B** + **C**, **A** − **B** and $\mathbf{BC}^t$ are **not** valid.

$\mathbf{CB} = \begin{pmatrix} 0 & -2 \\ -1 & 1 \end{pmatrix}$, $\mathbf{A}^2 = \begin{pmatrix} 0 & -2 & -4 \\ 3 & 3 & 3 \\ 6 & 8 & 10 \end{pmatrix}$

1.5. Prove that $\left(\mathbf{A}^T\mathbf{A}\right)^T = \mathbf{A}^T\mathbf{A}$ using $(\mathbf{XY})^T = \mathbf{Y}^T\mathbf{X}^T$ and $\left(\mathbf{X}^T\right)^T = \mathbf{X}$.

1.6. $a = c$ and $b = d = \frac{2}{3}c$

1.7. An example is $\mathbf{Au} = \mathbf{Av} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{O}$ with $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ and $\mathbf{u} = \begin{pmatrix} -2 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \mathbf{v}$.

1.8. (a) $\mathbf{A}^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$, $\mathbf{B}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, $\mathbf{AB} = \begin{pmatrix} -2 & 1 \\ -4 & 3 \end{pmatrix}$ and $\mathbf{BA} = \begin{pmatrix} 3 & 4 \\ -1 & -2 \end{pmatrix}$

(b) $\mathbf{AB} - \mathbf{BA} = \begin{pmatrix} bg - cf & af + bh - be - df \\ ce + dg - ag - ch & cf - bg \end{pmatrix}$

1.9. $\mathbf{M}^2 = 2\mathbf{M}$, $\mathbf{M}^3 = 4\mathbf{M}$ and $\mathbf{M}^4 = 8\mathbf{M}$. $c(n) = 2^{n-1}$.

1.10. (i) $\mathbf{A}^2 = \frac{1}{2}\left(\frac{2}{3}\right)^2 \mathbf{A}$ (ii) $\mathbf{A}^3 = \frac{1}{2}\left(\frac{2}{3}\right)^3 \mathbf{A}$

To prove the required result use mathematical induction.

1.11. Matrix **B** has four rows.

1.12. (a) The given matrix is **non-invertible (singular)**.
(b) $\mathbf{B}^{-1} = \mathbf{CA}$ so the matrix **B** is invertible.

1.13. (b) $\mathbf{A}^{-1} = \begin{bmatrix} 7/2 & 0 & -3 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$

1.14. $x = 3$, $y = 1$ and $z = 2$.

1.15. (a) Reduced row echelon form is $\left[\begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$.

(b) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 - 3t \\ 1 - 2t \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ -2 \\ 1 \end{bmatrix}$

1.16. $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2 + t \\ 3 - 3t \\ -1 - 2t \\ t \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -3 \\ -2 \\ 1 \end{bmatrix}$

1.17. $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -1 + 11s + 16t \\ 1 - 6s - 6t \\ s \\ t \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 11 \\ -6 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 16 \\ -6 \\ 0 \\ 1 \end{bmatrix}$

1.18. (i) $\begin{bmatrix} 1 & 2 & 5 & 6 \\ 0 & 1 & -12 & -16 \\ 0 & 0 & 13 & 26 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 2 & 5 & 6 \\ 0 & 1 & -12 & -16 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 0 & 0 & -20 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 2 \end{bmatrix}$

1.19. The system is inconsistent because if you carry our row operations you end up with something like $0x_1 + 0x_2 + 0x_3 + 0x_4 = -33$ which means that $0 = -33$

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

1.21. $rref\,(\mathbf{A}) = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

1.22. $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 3+2t-2s \\ 1+s-t \\ s \\ 2t+2 \\ t \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} + s\begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \\ 1 \end{pmatrix}$

1.23. (a) $\mathbf{E}_2\mathbf{E}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 5/2 & 1/2 \end{pmatrix} = \mathbf{A}^{-1}$

(b) $\mathbf{E}_1^{-1}\mathbf{E}_2^{-1} = \begin{pmatrix} 1 & 0 \\ -5 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -5 & 2 \end{pmatrix} = \mathbf{A}$

1.24. $\mathbf{E} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

1.25. Option D which is $\begin{bmatrix} 4 & -4 \\ 5 & 11 \end{bmatrix}$.

1.26. $\mathbf{X} = \begin{bmatrix} -6 & -2 & 2 \\ -5 & 4 & 3 \\ -2 & -3 & 0 \end{bmatrix}$

1.27. (a) $x = -4$, $y = -24$ and $z = 33$ (b) Inverse matrix is $\begin{pmatrix} -1 & 0 & 1 \\ -5 & 1 & 3 \\ 7 & -1 & -4 \end{pmatrix}$.

(c) If our z coefficient is 5/4 then the linear system has no solutions.

1.28. Prove $\left(\mathbf{A}^{-1}\right)^T = \mathbf{A}^{-1}$ using matrix operations.

1.29. Pre- and post-multiply $\mathbf{A}$ by $\mathbf{I} - \mathbf{A}$ and in each case the result should be $\mathbf{I}$.

1.30. (a) To show uniqueness, assume that we have two matrices which are inverses of a matrix $\mathbf{A}$ and then show that they are equal.

(b) Use mathematical induction.

(c) $\mathbf{A}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 2 \\ 1 & 1 & -1 \end{bmatrix}$

1.31. Pre-multiply $\mathbf{AXA}^{-1} = \mathbf{B}$ by $\mathbf{A}^{-1}$ and post-multiply by $\mathbf{A}$. You should get $\mathbf{X} = \mathbf{A}^{-1}\mathbf{BA}$. Take the inverse of both sides to find $\mathbf{X}^{-1} = \mathbf{A}^{-1}\mathbf{B}^{-1}\mathbf{A}$.

1.32. $\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \end{bmatrix} = \begin{bmatrix} p \\ 7-q-2r-4t \\ q \\ 8-3r-5t \\ r \\ 9-6t \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 7 \\ 0 \\ 8 \\ 0 \\ 9 \\ 0 \end{bmatrix} + p\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + q\begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + r\begin{bmatrix} 0 \\ -2 \\ 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ -4 \\ 0 \\ -5 \\ 0 \\ -6 \\ 1 \end{bmatrix}$

1.33. Go both ways ($\Rightarrow$) and ($\Leftarrow$) to show the required result.

1.34. Prove $(\mathbf{AB})\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right)=\mathbf{I}$ by using matrix operations and then apply result of question (33) to show that $\mathbf{B}^{-1}\mathbf{A}^{-1}$ is the inverse of $\mathbf{AB}$.

1.35. (i) Use mathematical induction to show the required result.
(ii) Multiply out the two matrices to show that $\mathbf{DD}^{-1}=\mathbf{I}$.

SOLUTIONS TO EXERCISES 2.1

1. (a) 1 (b) 1 (c) 10 (d) 5 (e) 10 (f) 5
(g) 3.16 (2 dp) (h) 2.24 (2dp) (i) 17 (j) 3.61 (2dp)

2. (a) 15 (b) 15 (c) 14 (d) 30 (e) 14 (f) 30
(g) 3.74 (2dp) (h) 5.48 (2dp) (i) 74 (j) 3.74 (2dp)

3. (a) -28 (b) -28 (c) 39 (d) 39 (e) 39 (f) 39
(g) 6.25 (2dp) (h) 6.25 (2dp) (i) 22 (j) 11.58 (2dp)

5. $x=2s,\ y=3s$ where s is any real number

6. (a) $\sqrt{5}$ (b) 13

9. (a) $\dfrac{1}{\sqrt{53}}(2\ \ -7)^T$ (b) $\dfrac{1}{\sqrt{139}}(-9\ \ 3\ \ 7)^T$ (c) $\dfrac{1}{\sqrt{134}}(-3\ \ 5\ \ 8\ \ 6)^T$
(d) $\dfrac{1}{\sqrt{138}}(-6\ \ 2\ \ 8\ \ 3\ \ 5)^T$

SOLUTIONS TO EXERCISES 2.2

1. (a) $\theta=45^\circ$ (b) $\theta=90^\circ$ (c) $\theta=168.69^\circ$

2. (a) $\theta=55.90^\circ$ (b) $\theta=90^\circ$ (c) $\theta=115.38^\circ$

3. (a) $\theta=41.98^\circ$ (b) $\theta=135^\circ$ (c) $\theta=56.56^\circ$

4. (a) $k=-13/7$ (b) $k=-5/3$ (c) $k=0$

5. (a) $\hat{\mathbf{u}}=\dfrac{1}{\sqrt{13}}\begin{pmatrix}2\\3\end{pmatrix}$ (b) $\hat{\mathbf{u}}=\dfrac{1}{\sqrt{14}}\begin{pmatrix}1\\2\\3\end{pmatrix}$ (c) $\hat{\mathbf{u}}=\dfrac{1}{3}\begin{pmatrix}2\\-2\\1\end{pmatrix}$

(d) $\hat{\mathbf{u}}=\dfrac{1}{\sqrt{5}}(1\ \ \sqrt{2}\ \ -1\ \ 1)^T$ (e) $\hat{\mathbf{u}}=\dfrac{1}{\sqrt{205}}(-2\ \ 10\ \ -10\ \ 1\ \ 0)^T$

6. $k=\dfrac{1}{2}$ or $k=-\dfrac{1}{2}$

7. (d) $\theta=90^\circ$

11. $(1\ \ 1\ \ -2)$ and $(-s-t\ \ s\ \ t)^T$

12. (a) 0.71 (b) 0.89 (c) 6.12 (d) 0.76

SOLUTIONS TO EXERCISES 2.3

1. (a) independent. (b), (c), (d), (e) dependent

2. (a), (b) and (d) independent (c) dependent

3. (a) independent (b) and (c) dependent

SOLUTIONS TO EXERCISES 2.4

1. (a) Span $\mathbb{R}^2$ (b) Span $\mathbb{R}^2$ (c) Does not span $\mathbb{R}^2$ (d) Span $\mathbb{R}^2$

2. (a) Span $\mathbb{R}^3$ (b) Does not span $\mathbb{R}^3$ (c) Span $\mathbb{R}^3$ (d) Span $\mathbb{R}^3$

3. (a) Forms a basis (b) No basis (c) Forms a basis (d) Forms a basis

5. (a) **b** is in the space spanned by columns of **A**.
(b) **b** is not spanned by the columns of **A**.

9. (a) independent. (b) $(x \;\; y \;\; 0)^T$ (c) $\mathbf{w} = (0 \;\; 0 \;\; z)^T$ $(z \neq 0)$

SOLUTIONS TO MISCELLANEOUS EXERCISES 2

2.1. Linearly dependent.

2.2. No such example exists because to span $\mathbb{R}^3$ you need three vectors.

2.3. (a) Linearly independent (b) $2\mathbf{v}_1 + \mathbf{v}_2 = (3, \; 2, \; 1)$ (c) Yes

2.4. $(3, \; 1, \; 1)$ **cannot** be expressed as a linear combination of the vectors $(2, \; 5, \; -1)$, $(1, \; 6, \; 0)$, $(5, \; 2, \; -4)$.

2.5. (b) $-3\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 7 \\ -6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

2.6. (a) $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$ and $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$
(b) (i) No because the set is linearly dependent but it does span $\mathbb{R}^3$.
(ii) Yes the given set of vectors is a basis for $\mathbb{R}^3$.

2.7. A basis is $\left\{\begin{bmatrix} a \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} b \\ b \\ 0 \\ 0 \\ b \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ c \\ 0 \\ c \end{bmatrix}\right\}$.

2.8. The first three vectors are linearly independent.

2.9. (a) Use definition (3.22) and show that **all** scalars are equal to zero.
(b) For all real values of λ providing $\lambda \neq -4$.

2.10. $c = 17$

2.11. By showing $c_1\mathbf{v}_1 + c_2(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{O} \Rightarrow c_1 = c_2 = 0$

2.12. None of the given sets form a basis for $\mathbb{R}^3$.

(a) We only have two vectors. (b) Linear dependence $2\begin{pmatrix}1\\2\\3\end{pmatrix} + \begin{pmatrix}4\\5\\6\end{pmatrix} = \begin{pmatrix}6\\9\\12\end{pmatrix}$

(c) We are given four vectors.

2.13. a. $\mathbf{v} + \mathbf{w} = \mathbf{u}$ b. Yes, they are linearly dependent because $\mathbf{v} + \mathbf{w} = \mathbf{u}$.

2.14. (a) Since the dot product is zero therefore $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is orthogonal.
(b) $\frac{5}{2}\mathbf{u}_1 - \frac{3}{2}\mathbf{u}_2 + 2\mathbf{u}_3 = \mathbf{x}$

2.15. (a) $\mathbf{u} \cdot \mathbf{u} = 13$ (b) $\|\mathbf{u}\| = \sqrt{13}$ (c) $\widehat{\mathbf{u}} = \frac{1}{\sqrt{13}}\begin{bmatrix}2\\0\\3\end{bmatrix}$ (d) $\mathbf{u} \cdot \mathbf{v} = 0$ (e) Yes

(f) $\|\mathbf{v} - \mathbf{x}\| = 6$

2.16. (c) $\theta = \cos^{-1}\left(\frac{1}{2}\right) = \frac{\pi}{3}$

2.17. (c) $\mathbf{v} = \begin{pmatrix}x\\y\\z\end{pmatrix} = s\begin{pmatrix}-1\\0\\1\end{pmatrix} + t\begin{pmatrix}-1\\1\\0\end{pmatrix}$ where $s, t \in \mathbb{R}$. A particular $\mathbf{v} = \begin{pmatrix}-2\\1\\1\end{pmatrix}$.

$\mathbf{w} = \begin{pmatrix}0\\-1\\1\end{pmatrix}$. Also $\widehat{\mathbf{u}} = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}$, $\widehat{\mathbf{v}} = \frac{1}{\sqrt{6}}\begin{pmatrix}-2\\1\\1\end{pmatrix}$ and $\widehat{\mathbf{w}} = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\-1\\1\end{pmatrix}$ where $\{\widehat{\mathbf{u}}, \widehat{\mathbf{v}}, \widehat{\mathbf{w}}\}$ is an orthonormal set of vectors in $\mathbb{R}^3$.

2.18. (a) $\mathbf{u} \cdot \mathbf{v} = 0$, $(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{w}) = 20$ and $(2\mathbf{u} - 3\mathbf{v}) \cdot (\mathbf{u} + 4\mathbf{w}) = -20$.
(b) $|\mathbf{u} \cdot \mathbf{v}| = 2$, $\|\mathbf{u}\|\,\|\mathbf{v}\| = 3\sqrt{2}$ therefore $|\mathbf{u} \cdot \mathbf{v}| = 2 \le 3\sqrt{2} = \|\mathbf{u}\|\,\|\mathbf{v}\|$

2.19. Consider the linear combination $k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 = \mathbf{O}$ then show that the dot product of $\mathbf{O} \cdot \mathbf{v}_1 = \mathbf{O} \cdot \mathbf{v}_2 = \mathbf{O} \cdot \mathbf{v}_3$ gives $k_1 = k_2 = k_3 = 0$.

2.20. Show that $k_1(\mathbf{A}\mathbf{u}_1) + k_2(\mathbf{A}\mathbf{u}_2) + \cdots + k_n(\mathbf{A}\mathbf{u}_n) = \mathbf{O}$ implies $k_1 = k_2 = \cdots k_n = 0$.

2.21. Prove that $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u} - \mathbf{v}\|^2$.

2.22. From $\mathbf{u} \cdot (\mathbf{v} - \mathbf{w}) = 0$ we have

$$u_1(v_1 - w_1) + u_2(v_2 - w_2) + \cdots + u_n(v_n - w_n) = 0$$

Thus $v_1 = w_1, v_2 = w_2, \ldots, v_n = w_n$.

2.23. $(\Leftarrow)$. Show that $k_1\mathbf{u} + k_2\mathbf{v} = \mathbf{O} \quad \Rightarrow k_1 = k_2 = 0$.
$(\Rightarrow)$. By assuming $k_1\mathbf{u} + k_2\mathbf{v} = \mathbf{O}$ and $k_1 = k_2 = 0$ show that $ad - bc \neq 0$.

2.24. Use proof by contradiction. Suppose m linearly independent vectors span $\mathbb{R}^n$.

2.25. Consider $k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 + \cdots + k_n\mathbf{v}_n = \mathbf{O}$ and prove that

$$k_1 = k_2 = k_3 = \cdots = k_n = 0$$

Then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n\}$ are linearly independent which means they form a basis for $\mathbb{R}^n$.

SOLUTIONS TO EXERCISES 3.2

7. S is a subspace of $\mathbb{R}^3$.
18. (a), (b) and (c) are in span$\{\mathbf{f}, \mathbf{g}\}$ but part (d) is not in the span$\{\mathbf{f}, \mathbf{g}\}$.
19. Does not span.

SOLUTIONS TO EXERCISES 3.3

1. (a) and (c) independent (b) and (d) dependent
2. (a) dependent (b), (c), (d) and (e) independent
3. (a), (b), (d) and (e) dependent. Others independent.
6. $2 + 2(t-1) + (t-1)^2$.

SOLUTIONS TO EXERCISES 3.4

1. (a) 5 (b) 7 (c) 11 (d) 13 (e) 9 (f) 16 (g) 6 (h) 4 (i) 6 (j) 0
2. $\dim(S) = 1$
3. $\dim(S) = 2$
4. $\dim(S) = 2$
5. $\dim(S) = 3$
6. (a) mn (b) 3 (c) 4

SOLUTIONS TO EXERCISES 3.5

1. (a) Row vectors – $\begin{pmatrix}1\\2\end{pmatrix}$, $\begin{pmatrix}3\\4\end{pmatrix}$ and column vectors – $\begin{pmatrix}1\\3\end{pmatrix}$ and $\begin{pmatrix}2\\4\end{pmatrix}$.

   (b) Row vectors $(1\ \ 2\ \ 3\ \ 4)^T$, $(5\ \ 6\ \ 7\ \ 8)^T$ and $(9\ \ 10\ \ 11\ \ 12)^T$.
   Column vectors $(1\ \ 5\ \ 9)^T$, $(2\ \ 6\ \ 10)^T$, $(3\ \ 7\ \ 11)^T$ and $(4\ \ 8\ \ 12)^T$.

   (c) Row vectors $(1\ \ 2)^T$, $(3\ \ 4)^T$, $(5\ \ 6)^T$. Column vectors $(1\ 3\ 5)^T$ and $(2\ 4\ 6)^T$.

   (d) Row vectors $(1\ 2\ 3)^T$ and $(4\ 5\ 6)^T$. Column vectors $(1\ 4)^T$, $(2\ 5)^T$ and $(3\ 6)^T$.

   (e) Row vectors $\begin{pmatrix}-1\\2\\5\end{pmatrix}$, $\begin{pmatrix}-3\\7\\0\end{pmatrix}$ and $\begin{pmatrix}-8\\1\\3\end{pmatrix}$.

   Column vectors $\begin{pmatrix}-1\\-3\\-8\end{pmatrix}$, $\begin{pmatrix}2\\7\\1\end{pmatrix}$ and $\begin{pmatrix}5\\0\\3\end{pmatrix}$.

(f) Row vectors $\begin{pmatrix} -5 \\ 2 \\ 3 \end{pmatrix}$, $\begin{pmatrix} 7 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} -7 \\ 6 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -2 \\ 5 \\ 2 \end{pmatrix}$.

Column vectors $\begin{pmatrix} -5 \\ 7 \\ -7 \\ -2 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 1 \\ 6 \\ 5 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix}$.

2. (a) $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$, $rank\,(\mathbf{A}) = 2$ (b) $\left\{\begin{pmatrix} 1 \\ 0 \\ -1 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}\right\}$, $rank\,(\mathbf{B}) = 2$

(c) $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$, $rank\,(\mathbf{C}) = 2$ (d) $\left\{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}\right\}$, $rank\,(\mathbf{D}) = 2$

(e) and (f) $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ with $rank\,(\mathbf{E}) = 3$ and $rank\,(\mathbf{F}) = 3$

3. (a) $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$ (b) $\left\{\begin{pmatrix} 1 \\ 3 \end{pmatrix}\right\}$

(c) and (d) $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ (e) $\left\{\begin{pmatrix} 1 \\ 0 \\ -0.8 \\ -1.6 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1.4 \\ 1.8 \end{pmatrix}\right\}$

4. $\left\{\frac{1}{7}\begin{pmatrix} 7 \\ 0 \\ 8 \\ -4 \end{pmatrix}, \frac{1}{7}\begin{pmatrix} 0 \\ 7 \\ 1 \\ 3 \end{pmatrix}\right\}$

5. For rank, see solution to question 2.

(a) $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$ (b) $\left\{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}\right\}$

(c) $\left\{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}\right\}$ (d) $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$

(e) $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ (f) $\left\{\begin{pmatrix} 128 \\ 0 \\ 0 \\ 61 \end{pmatrix}, \begin{pmatrix} 0 \\ 128 \\ 0 \\ 80 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 128 \\ 73 \end{pmatrix}\right\}$

SOLUTIONS TO EXERCISES 3.6

1. (a) $\{\mathbf{O}\}$ (b) $\left\{ s\begin{pmatrix} -2 \\ 1 \end{pmatrix} \middle| s \in \mathbb{R} \right\}$ (c) $\{\mathbf{O}\}$

(d) $\{\mathbf{O}\}$

2. (a) $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}$ (b) $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

(c) and (d) $\mathbf{x} = \mathbf{O}$

3. (a) $s\begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}$, $rank\,(\mathbf{A}) = 2$, $nullity\,(\mathbf{A}) = 1$

(b) $s\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $rank\,(\mathbf{B}) = 1$, $nullity\,(\mathbf{B}) = 2$

(c) and (d) $\{\mathbf{O}\}$, $rank\,(\mathbf{C}) = rank\,(\mathbf{D}) = 3$, $nullity\,(\mathbf{C}) = nullity\,(\mathbf{D}) = 0$

4. $(s + 2t + 3p + 4q + 5r \quad -2s - 3t - 4p - 5q - 6r \quad s \quad t \quad p \quad q \quad r)^T$

5. (a) $\left\{ \begin{pmatrix} 13 \\ 0 \\ -73 \end{pmatrix}, \begin{pmatrix} 0 \\ 13 \\ 11 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 73 \\ -11 \\ 13 \end{pmatrix} \right\}$, $rank\,(\mathbf{A}) = 2$, $nullity\,(\mathbf{A}) = 1$

(b) $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -5 \end{pmatrix} \right\}$, $\{\mathbf{O}\}$, $rank\,(\mathbf{B}) = 2$, $nullity\,(\mathbf{B}) = 0$

(c) $\left\{ \begin{pmatrix} 1 \\ 3 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$, $rank\,(\mathbf{C}) = 3$,
$nullity\,(\mathbf{C}) = 1$

6. (a) $\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \underbrace{s\begin{pmatrix} -1 \\ -3 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -5 \\ 0 \\ 0 \\ 1 \end{pmatrix}}_{=\mathbf{x}_H} + \underbrace{\begin{pmatrix} 9 \\ -3 \\ 0 \\ 0 \end{pmatrix}}_{=\mathbf{x}_P}$

(b) $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}$

7. (a) infinite (b) infinite (c) no solution (d) unique solution

8. (a) and (b) in the null space. (c) and (d) not in the null space.

SOLUTIONS TO MISCELLANEOUS EXERCISES 3

3.1. $N = \left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} \right\}$

3.2. B

3.3. F

3.4. (a) Use $\mathbf{Av} = \mathbf{b}$ and $\mathbf{Aw} = \mathbf{O}$ (b) $\mathbf{Au} = \mathbf{b}$ and $\mathbf{Ax} = \mathbf{b}$

3.5. (a) Matrix $\mathbf{A}$ is invertible (non-singular) (b) linear independent rows
(c) $rank(\mathbf{A}) = n$

3.6. (a) false (b) true (c) false (d) true

3.7. (a) true (b) true (c) false

3.8. All three statements are true.

3.9. (a) $rank(\mathbf{A}) = 3$ (b) $nullity(\mathbf{A}) = 2$

(c) $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}$ (d) $\left\{ \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

(e) $\left\{ \begin{bmatrix} 2 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$

3.10. (b) $\left\{ s \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \,\middle|\, s \in \mathbb{R} \right\}$ (c) $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent.

3.11. (a) $\{\mathbf{O}\}$ and the space $\mathbb{R}^n$ (c) $-4\mathbf{v}_1 + 3\mathbf{v}_2 + \mathbf{v}_3 = \mathbf{O}$ and dimension is 2.

3.12. (a) $\left[\begin{array}{ccc|c} 1 & 0 & -2 & -3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] = \mathbf{R}$ (b) The system is consistent. (c) $\vec{x} = \underbrace{s \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}}_{=\mathbf{x}_H} + \underbrace{\begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}}_{=\mathbf{x}_P}$

3.13. $\left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$

3.14. (b) $\begin{bmatrix} 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathbf{R}$, $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$, nullity and rank of matrix $\mathbf{A}$ is 2.

Two linearly independent and linear combination of first and last columns of the matrix $\mathbf{A}$.

3.15. (a) (i) Not a subspace (ii) Subspace (iii) Subspace (b) $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ -4 \end{pmatrix}$

(c) (i) Not a basis, linearly dependent, span $\mathbb{R}^3$. (ii) Is a basis

3.16. Apply the definition of subspace.

3.17. (a) $\mathbf{v}$ is in the span of S (b) $\mathbf{v}$ is in the span of S

3.18. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{R}$ (a) $\begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$ (b) $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

(c) *Nullity* $(\mathbf{B}) = 1$ (d) *rank* $(\mathbf{B}) = 3$

3.19. Use the definition of subspace.

3.20. (a) Yes (b) Dim is 3 and a set of basis vectors are $\mathbf{u}_1 = (2\ \ 0\ \ 5\ \ 3)^T$, $\mathbf{u}_2 = (0\ \ 1\ \ 0\ \ 0)^T$ and $\mathbf{u}_3 = (0\ \ 0\ \ 2\ \ 1)^T$

(c) Yes is in V. (d) 1 (e) $\begin{pmatrix} -0.25 \\ 0 \\ -0.5 \\ 1 \end{pmatrix}$

3.21. Use Proposition (4.34).

3.22. Use Proposition (4.34).

3.23. (a) $N = \left\{ t \begin{bmatrix} -3 \\ 2 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 0 \\ 2 \\ 0 \end{bmatrix} \middle|\ s, t \in \mathbb{R} \right\}$

(b) $\mathbf{x} = \mathbf{x}_P + \mathbf{x}_H = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix} + \underbrace{t \begin{bmatrix} -3 \\ 2 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 0 \\ 2 \\ 0 \end{bmatrix}}_{\mathbf{x}_H \text{ solution of part (a)}}$

(c) $n - m$ solutions.

3.24. $c = 17$

3.25. $\mathbf{A} = \begin{pmatrix} 4 & 7 & -6 \\ 5 & 8 & -7 \\ 6 & 9 & -8 \end{pmatrix}$

3.26. (a) Use the trigonometric identities of $\cos(x)$, $\cos(2x)$, $\cos(3x)$.
(b) Expand each of these to show that they are in the *span* $\{1,\ x,\ x^2,\ x^3\}$.

3.27. (a) Write two matrices whose rows are each of the given vectors and then show that they have the same basis which means they are equal.
(b) $k = 0, \quad k = -1$

3.28. All three statements (a), (b) and (c) are false.

3.29. Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly dependent and derive a contradiction. Because these are n linearly independent vectors and dimension of $\mathbb{R}^n$ is n therefore they form a basis.
To show that $\mathbf{A}$ is invertible write an arbitrary vector $\mathbf{u}$ uniquely as

$$c_1\mathbf{A}\mathbf{v}_1 + c_2\mathbf{A}\mathbf{v}_2 + \cdots + c_n\mathbf{A}\mathbf{v}_n = \mathbf{u}$$

Rewrite this as $\mathbf{A}\mathbf{x} = \mathbf{u}$ where $\mathbf{x}$ is unique then $\mathbf{A}$ is invertible.

SOLUTIONS TO EXERCISES 4.1

1. (a) $1/4$ (b) $1/4$ (c) $3/4$ (d) $1/3$ (e) $1/\sqrt{3}$ (f) $1/5$ (g) $1/\sqrt{5}$

2. (a) $2/5$ (b) $2/5$ (c) $6/5$ (d) $2/3$ (e) $\sqrt{2/3}$ (f) $2/7$ (g) $\sqrt{2/7}$

3. (a) -21 (b) -21 (c) 63 (d) 38 (e) $\sqrt{38}$ (f) 90 (g) $\sqrt{90}$

5. (a) 70 (b) 350 (c) 70 (d) $\sqrt{30}$ (e) $\sqrt{174}$ (f) 27 (g) 55 (h) 82 (i) 97

7. (a) $\frac{7}{12}$ (b) $-\frac{1}{12}$ (c) $-\frac{2}{3}$ (d) $-\frac{3}{4}$ (e) $-\frac{1}{12}$ (f) $-\frac{11}{12}$
(g) $-\frac{5}{12}$ (h) $-\frac{7}{12}$ (i) $-10\frac{1}{2}$ (j) $-15\frac{5}{6}$

11. (a) 4 (b) 12 (c) $\sqrt{22}$ (d) $\sqrt{67}$

SOLUTIONS TO EXERCISES 4.2

4. (b) 17.23 (2dp)

5. (a) $k = -8$ (b) $k = -1$

8. (d) $\dfrac{\mathbf{f}}{\|\mathbf{f}\|} = \dfrac{\cos(x)}{\sqrt{\pi}}$ and $\dfrac{\mathbf{g}}{\|\mathbf{g}\|} = \dfrac{\sin(x)}{\sqrt{\pi}}$

9. The orthonormal set is $\left\{\dfrac{1}{\sqrt{\pi}},\ \sqrt{\dfrac{2}{\pi}}\sin(2t),\ \sqrt{\dfrac{2}{\pi}}\sin(4t),\ \sqrt{\dfrac{2}{\pi}}\sin(6t),\cdots\right\}$

11. 1.713

14. (a) $\|\mathbf{u}+\mathbf{v}\| = \sqrt{2}$ (b) $\|\mathbf{u}-\mathbf{v}\| = \sqrt{2}$

15. (ii) $\sqrt{n}$

SOLUTIONS TO EXERCISES 4.3

1. (a) $\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (b) $\mathbf{u}_1 = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\mathbf{u}_2 = \frac{1}{\sqrt{10}}\begin{pmatrix} 3 \\ -1 \end{pmatrix}$

(c) $\mathbf{u}_1 = \frac{1}{\sqrt{13}}\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and $\mathbf{u}_2 = \frac{1}{\sqrt{13}}\begin{pmatrix} 3 \\ -2 \end{pmatrix}$

(d) $\mathbf{u}_1 = \frac{1}{\sqrt{29}}\begin{pmatrix} -2 \\ -5 \end{pmatrix}$ and $\mathbf{u}_2 = \frac{1}{\sqrt{29}}\begin{pmatrix} -5 \\ 2 \end{pmatrix}$

2. $\mathbf{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

3. $\mathbf{u}_1 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{u}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ and $\mathbf{u}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$

4. (a) $\mathbf{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{u}_2 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ and $\mathbf{u}_3 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$

(b) $\mathbf{u}_1 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{u}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ and $\mathbf{u}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$

(c) $\mathbf{u}_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \ \mathbf{u}_2 = \frac{1}{\sqrt{45}}\begin{pmatrix} 4 \\ -2 \\ 5 \end{pmatrix}$, and $\mathbf{u}_3 = \frac{1}{3}\begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}$

5. (a) $\mathbf{u}_1 = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \end{pmatrix}, \ \mathbf{u}_2 = \frac{1}{\sqrt{110}}\begin{pmatrix} 3 \\ 10 \\ -1 \\ 0 \end{pmatrix}$ and $\mathbf{u}_3 = \frac{1}{\sqrt{11}}\begin{pmatrix} 3 \\ -1 \\ -1 \\ 0 \end{pmatrix}$

(b) $\mathbf{w}_1 = \frac{1}{\sqrt{31}}\begin{pmatrix} 1 \\ 1 \\ 5 \\ 2 \end{pmatrix}, \mathbf{w}_2' = \frac{1}{\sqrt{28582}}\begin{pmatrix} -109 \\ 77 \\ 44 \\ -94 \end{pmatrix}$ and $\mathbf{w}_3' = \frac{1}{\sqrt{15555062}}\begin{pmatrix} -2929 \\ -1289 \\ -78 \\ 2304 \end{pmatrix}$

(c) $\mathbf{u}_1 = \frac{1}{\sqrt{30}}\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \mathbf{u}_2 = \frac{1}{\sqrt{3930}}\begin{pmatrix} 53 \\ 16 \\ 9 \\ -28 \end{pmatrix}$ and $\mathbf{u}_3 = \frac{1}{\sqrt{224665}}\begin{pmatrix} 224 \\ -140 \\ -308 \\ 245 \end{pmatrix}$

6. $\widehat{\mathbf{p}_1} = \frac{1}{\sqrt{2}}, \ \widehat{\mathbf{p}_2} = \sqrt{\frac{3}{2}}x$ and $\widehat{\mathbf{p}_3} = \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)$

7. $\mathbf{p}_1 = x^2, \quad \mathbf{p}_2 = x$ and $\mathbf{p}_3 = 1 - \frac{5}{3}x^2$

8. (a) $\frac{4}{3}\mathbf{p}_1 + \mathbf{p}_2 + \frac{2}{3}\mathbf{p}_3$ (b) $-\frac{1}{3}\mathbf{p}_1 + \frac{4}{3}\mathbf{p}_3$ (c) $3\mathbf{p}_1$ (d) $\frac{1}{3}\mathbf{p}_1 + \frac{2}{3}\mathbf{p}_3$

(e) $2\mathbf{p}_1 + 5\mathbf{p}_2$

SOLUTIONS TO EXERCISES 4.4

2. In each case $\mathbf{Q}^{-1} = \mathbf{Q}$.

3. (a) and (b) not orthogonal.

(c) Orthogonal and inverse is $\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{pmatrix}$.

4. Lie on the unit circle and perpendicular.

5. Two matrices $\begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{pmatrix}$ or $\begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \end{pmatrix}$.

6. $\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$. Matrix $\mathbf{Q}$ rotates the vector.

7. (a) $\mathbf{Q} = \frac{1}{\sqrt{6}}\begin{pmatrix} \sqrt{2} & 1 & -\sqrt{3} \\ \sqrt{2} & -2 & 0 \\ \sqrt{2} & 1 & \sqrt{3} \end{pmatrix}$, $\mathbf{R} = \frac{1}{\sqrt{6}}\begin{pmatrix} 6\sqrt{2} & -2\sqrt{2} & 4\sqrt{2} \\ 0 & -2 & -2 \\ 0 & 0 & 4\sqrt{3} \end{pmatrix}$

(b) $x = 1,\ y = 2,\ z = 3$ (c) $x = 2,\ y = 5$ and $z = 4$

SOLUTIONS TO MISCELLANEOUS EXERCISES 4

4.1. False.

4.2. $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{w}_2' = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$ and $\mathbf{w}_3' = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$

4.3. (a) $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \end{bmatrix} \right\}$ (b) $\left\{ \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{10}}\begin{bmatrix} -1 \\ 2 \\ 1 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{15}}\begin{bmatrix} 1 \\ -2 \\ -1 \\ 3 \end{bmatrix} \right\}$

4.4. (b) Show $\mathbf{f}_1$, $\mathbf{f}_2$ and $\mathbf{f}_3$ are linearly independent. (c) $\left\{ \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 6 \\ -10 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix} \right\}$

4.5. (a) Expanding $\langle \mathbf{v}_1, \mathbf{O} \rangle = \langle \mathbf{v}_2, \mathbf{O} \rangle = \langle \mathbf{v}_3, \mathbf{O} \rangle = \cdots = \langle \mathbf{v}_m, \mathbf{O} \rangle = 0$.

(b) Consider $\mathbf{w} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_n\mathbf{v}_n$ and show that $k_1 = (\mathbf{v}_1 \cdot \mathbf{w}), \ldots,$ $k_n = (\mathbf{v}_n \cdot \mathbf{w})$

4.6. (b) $\left\{ \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$

4.7. (a) $\sqrt{39}$ (b) No

(c) Yes (d) $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \right\}$

4.8. (a) $\sqrt{39}$ (b) No (c) Yes (d) $\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$

(e) $\lambda_1 = 1,\ \lambda_2 = \frac{1}{\sqrt{2}},\ \lambda_3 = 0,\ \lambda_4 = -\frac{1}{\sqrt{2}}$

4.9. (a) $\sqrt{17}$

(b) $\widehat{\mathbf{w}} = \frac{\mathbf{w}}{\|\mathbf{w}\|}$ where $\mathbf{w} = \mathbf{u} - \langle \mathbf{u}, \mathbf{q}_1 \rangle \mathbf{q}_1 - \langle \mathbf{u}, \mathbf{q}_2 \rangle \mathbf{q}_2 - \langle \mathbf{u}, \mathbf{q}_3 \rangle \mathbf{q}_3$

(c) Because $\mathbf{u}$ is linearly dependent on $\mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_3$.

4.10. $\mathbf{w}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{w}_2' = \begin{bmatrix} 0 \\ 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}$ and $\mathbf{w}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

4.11. (a) and (b). We have $\mathbf{v} = \begin{bmatrix} 2, & -1 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 1, & 3 \end{bmatrix}$.

4.12. $\left\{ \frac{1}{\sqrt{2}}, \frac{\sqrt{3}\mathbf{x}}{\sqrt{2}} \right\}$

4.13. Check definitions (i) to (iv) of (4.1). (b) $\|\mathbf{A} - \mathbf{B}\| = \sqrt{9} = 3$ (c) 31.65°

4.14. $\{4 - 10t + 4t^2\}$

4.15. $\left\{ \mathbf{1},\ \mathbf{x} - \frac{1}{2},\ \mathbf{x}^2 - \mathbf{x} + \frac{1}{6} \right\}$

4.16. Show that condition (iv) of definition (4.1) fails.

4.17. Let $\mathbf{u} \in V$ where $\mathbf{u} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 + \cdots + k_n\mathbf{v}_n$. Show that $k_j = \langle \mathbf{u}, \mathbf{v}_j \rangle$ for $j = 1, 2, 3, \ldots, n$.

4.18. $\left\{ \frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}\mathbf{x},\ \sqrt{\frac{5}{8}}\left(3\mathbf{x}^2 - 1\right), \sqrt{\frac{7}{8}}\left(5\mathbf{x}^3 - 3\mathbf{x}\right) \right\}$

4.19. $\left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \\ 2 \\ -3 \end{pmatrix} \right\}$

4.20. Let $\mathbf{u} = \mathbf{v}$ then $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle = 0$ gives $\mathbf{u} = \mathbf{O}$.

4.21. $\langle k\mathbf{u}, \mathbf{v} \rangle = k \langle \mathbf{u}, \mathbf{v} \rangle = 0$

4.22. Use the definition of orthonormal vectors.

4.23. (a) Use Proposition (4.5).

(b) Apply Minkowski's inequality (4.7).

SOLUTIONS TO EXERCISES 5.1

1. (a) $(3\ \ 1)^T$ (b) $(5\ \ -1)^T$ (c) $(1\ \ \sqrt{2})^T$ (d) $(-3\ \ -2)^T$

2. (a) $(3\ \ 5)^T$ (b) $(4\ \ -4)^T$ (c) $(4\ \ 6)^T$ (d) $(6\ \ 8)^T$

3. (a) $\begin{pmatrix} 6 \\ 11 \\ 9 \end{pmatrix}$ (b) $\begin{pmatrix} 5 \\ 2 \\ -9 \end{pmatrix}$ (c) $\begin{pmatrix} 3\pi \\ 7\pi \\ 6\pi \end{pmatrix}$ (d) $\begin{pmatrix} 7/6 \\ 17/12 \\ 5/4 \end{pmatrix}$

5. (a) and (d) linear (b), (c) and (e) not linear

6. (a) linear (b) not linear

7. (a) linear (b) linear (c) not linear

SOLUTIONS TO EXERCISES 5.2

1. (a) $\{\mathbf{O}\}$ (b) $\left\{ r\begin{pmatrix} -1 \\ 1 \end{pmatrix} \,\middle|\, r \in \mathbb{R} \right\}$ (c) $\mathbb{R}^2$

(d) $\left\{ \begin{pmatrix} 0 \\ a \\ b \end{pmatrix} \,\middle|\, a \in \mathbb{R},\ b \in \mathbb{R} \right\}$ (e) $\left\{ r\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \,\middle|\, r \in \mathbb{R} \right\}$

2. $\ker(T) = \{\mathbf{O}\}$ and $\text{range}(T) = M_{nn}$

3. $\ker(T) = \{c \mid c \in \mathbb{R}\} = P_0$ and $\text{range}(T) = P_1$

SOLUTIONS TO EXERCISES 5.3

1. (a) (i) $\left\{ r\begin{pmatrix} -2 \\ 1 \end{pmatrix} \,\middle|\, r \in \mathbb{R} \right\}$ (ii) 1 (iii) $\left\{ \begin{pmatrix} a \\ a \end{pmatrix} \,\middle|\, a \in \mathbb{R} \right\}$ (iv) 1

(b) (i) $\{\mathbf{O}\}$ (ii) 0 (iii) $\mathbb{R}^3$ (iv) 3

(c) (i) $\left\{ s\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -5 \\ 0 \\ 1 \end{pmatrix} \,\middle|\, s \in \mathbb{R} \text{ and } t \in \mathbb{R} \right\}$ (ii) 2 (iii) $\mathbb{R}$ (iv) 1

(d) (i) P_0 (ii) 1 (iii) $\left\{ dx^3 + ex^2 + fx \,\middle|\, d,\ e \text{ and } f \in \mathbb{R} \right\}$ (iv) 3

(e) $\{c \mid c \in \mathbb{R}\} = P_0$ (ii) 1 (iii) P_2 (iv) 3

(f) (i) $\left\{ ax^3 + bx^2 + cx + d \,\middle|\, a = -\frac{4}{3}b - 2c - 4d \right\}$ (ii) 3 (iii) $\mathbb{R}$ (iv) 1

(g) (i) $\ker(T) = \left\{ \begin{pmatrix} a & b \\ -a & -b \end{pmatrix} \right\}$ (ii) 2 (iii) P_1 (iv) 2

2. Substitute the values obtained in question 1 into nullity (T) + rank $(T) = n$.
The values of n for each part are as follows:
(a) $n = 2$ (b) $n = 3$ (c) $n = 3$ (d) $n = 4$ (e) $n = 4$ (f) $n = 4$
(g) $n = 4$

3. (i) $B = \{(-3\ \ 1\ \ 0\ \ 0\ \ 0)^T, (3.2\ \ 0\ \ -0.4\ \ 1\ \ 0)^T, (-3.6\ \ 0\ \ 0.2\ \ 0\ \ 1)^T\}$
(ii) $B' = \{(1\ \ 0\ \ 8/3)^T, (0\ \ 1\ \ 2/3)^T\}$

SOLUTIONS TO EXERCISES 5.4

3. (a) and (b) T is not one-to-one nor onto.
(c) T is one-to-one and onto.
(d) T is one-to-one but not onto.

17. $T^{-1}\left[\begin{pmatrix} a \\ b \end{pmatrix}\right] = \frac{1}{3}\begin{pmatrix} a+b \\ a-2b \end{pmatrix}$

SOLUTIONS TO EXERCISES 5.5

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

(c) $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 1 & -5 \end{pmatrix}$ (d) $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 2 & 1 & -1 \end{pmatrix}$

(e) $\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 4 & -1 & 3 \\ 7 & -1 & -1 \end{pmatrix}$ (f) $\mathbf{A} = \begin{pmatrix} 1 & -1 & 1 & -3 \\ -1 & 3 & -7 & -1 \\ 9 & 5 & 6 & 12 \\ 1 & 0 & 0 & 0 \end{pmatrix}$

(g) $\mathbf{A} = \mathbf{O}_4$

2. (a) $\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix} 3 & -2 \\ -1 & 5 \end{pmatrix}$

(c) $\mathbf{A} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$ (d) $\mathbf{A} = \mathbf{O}_3$

(e) $\mathbf{A} = \begin{pmatrix} -3 & -5 & -6 \\ -2 & 7 & 5 \\ 0 & 0 & 0 \end{pmatrix}$

3. $\mathbf{A} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

4. $\mathbf{A} = \begin{pmatrix} 0 & -1 & -2 \\ 0 & 0 & 2 \end{pmatrix}$ and $T\left(2x^2 + 3x + 1\right) = 3 + 4x$

5. (i) $\mathbf{A}=\begin{pmatrix}0&1&0&0\\0&0&2&0\\0&0&0&3\end{pmatrix}$ (ii) $\mathbf{A}=\begin{pmatrix}0&0&0&3\\0&0&2&0\\0&1&0&0\end{pmatrix}$ (iii) $\mathbf{A}=\begin{pmatrix}0&0&1&0\\0&2&0&0\\3&0&0&0\end{pmatrix}$

$T\left(-1+3x-7x^2-2x^3\right)=3-14x-6x^2$.

6. $\mathbf{A}=\begin{pmatrix}3&-1\\7&-3\end{pmatrix}$ and $T\left(\begin{pmatrix}-3\\1\end{pmatrix}\right)=\begin{pmatrix}-2\\-4\end{pmatrix}$

7. $\left\{\mathbf{v}_1=\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix},\ \mathbf{v}_2=\begin{pmatrix}-1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}\right\}$ and $\dfrac{1}{\sqrt{2}}\begin{pmatrix}3\\-1\end{pmatrix}$

8. $\mathbf{A}=\begin{pmatrix}7/3&1\\-10/3&-2\\14/3&3\end{pmatrix}$ and $T\left(\begin{pmatrix}2\\1\end{pmatrix}\right)=\begin{pmatrix}-2\\-1\\5\end{pmatrix}$

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

10. $\mathbf{A}=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$ (i) $-2\sin(x)+5\cos(x)$ (ii) $-n\sin(x)+m\cos(x)$

11. $\mathbf{A}=\begin{pmatrix}1&3&9\\0&1&6\\0&0&1\end{pmatrix}$ and $T\left(q+nx+mx^2\right)=q+3n+9m+(n+6m)x+mx^2$

12. $\mathbf{A}=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&1\end{pmatrix}$ (i) $T\left(-\sin(x)+4\cos(x)-2e^x\right)=-4\sin(x)-\cos(x)-2e^x$

(ii) $T\left(m\sin(x)+n\cos(x)+pe^x\right)=-n\sin(x)+m\cos(x)+pe^x$

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

$T\left(ae^{2x}+bxe^{2x}+cx^2e^{2x}\right)=(2a+b)e^{2x}+(2b+2c)xe^{2x}+2cx^2e^{2x}$

SOLUTIONS TO EXERCISES 5.6

1. (i) $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\2\end{pmatrix}$ (ii) $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$

2. (a) $2ax+b$ (b) $2ax$ (c) $2a$ (d) ax^2+bx

3. $\begin{pmatrix}3&1\\-2&1\end{pmatrix}$ and $\begin{pmatrix}0&-1\\5&4\end{pmatrix}$.

(i) $\begin{pmatrix}8\\3\end{pmatrix}$ (ii) $\begin{pmatrix}-5\\25\end{pmatrix}$ (iii) $\begin{pmatrix}-17\\53\end{pmatrix}$ (iv) $\begin{pmatrix}8\\-1\end{pmatrix}$

4. $\begin{pmatrix} -6 & 3 & -7 \\ 7 & 0 & -2 \\ 11 & -2 & -5 \end{pmatrix}$ and $\begin{pmatrix} 7 & 17 & 26 \\ -2 & 4 & 5 \\ -7 & -16 & -22 \end{pmatrix}$

(i) $\begin{pmatrix} -21 \\ 1 \\ -8 \end{pmatrix}$ (ii) $\begin{pmatrix} 119 \\ 21 \\ -105 \end{pmatrix}$ (iii) $\begin{pmatrix} -151 \\ -85 \\ 243 \end{pmatrix}$ (iv) $\begin{pmatrix} -17 \\ 31 \\ 54 \end{pmatrix}$

5. (a) (i) $\begin{pmatrix} 0 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ (ii) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 0 \end{pmatrix}$ (iii) $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 4 & 1 \end{pmatrix}$ (iv) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$S^{-1}(\mathbf{p}) = S(\mathbf{p}) = ce^x + bxe^x + ax^2e^x$ (b) $2e^x - 2xe^x + x^2e^x + C$

6. (i) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ (ii) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ (iii) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (iv) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$(S \circ T)\left(1 + 2x + 3x^2\right) = 1 + 2x,\quad (T \circ S)\left(1 + 2x + 3x^2\right) = 1 + 2x,$
$(T \circ T)\left(1 + 2x + 3x^2\right) = 1 + 2x + 3x^2$ and $(S \circ S)\left(1 + 2x + 3x^2\right) = 1 + 2x$

7. $\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ and $\mathbf{A}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$.

8. $T^{-1}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \frac{1}{2}\begin{pmatrix} x+z \\ y-z \\ x-y \end{pmatrix}$

9. $T^{-1}(\mathbf{p}) = T^{-1}\left(ax^3 + bx^2 + cx + d\right) = \frac{a}{4}x^3 + \frac{b}{3}x^2 + \frac{c}{2}x + d$

10. $T^{-1}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{pmatrix} x-z \\ 2x+y+z \\ 2x-y \end{pmatrix}$

11. (a) Not Invertible (b) Not Invertible (c) Invertible, $T^{-1}(\mathbf{p}) = \mathbf{p}$

12. (i) No change.

13. (i) $\mathbf{x}_1 = \begin{pmatrix} 0.3333 \\ 0.6667 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 0.5556 \\ 0.4444 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 0.4815 \\ 0.5185 \end{pmatrix}, \mathbf{x}_4 = \begin{pmatrix} 0.5062 \\ 0.4938 \end{pmatrix}$

and $\mathbf{x}_5 = \begin{pmatrix} 0.4979 \\ 0.5021 \end{pmatrix}$.

(ii) $\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$

SOLUTIONS TO MISCELLANEOUS EXERCISES 5

5.2. $\mathbf{S} = \begin{bmatrix} 2 & -3 & 4 \\ -1 & 1 & 0 \end{bmatrix}$

5.3. $f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{pmatrix} x^2 \\ xy \end{pmatrix}$

5.4. (a) $\mathbf{S}=\begin{pmatrix}0 & 3 & 2\\ 3 & -4 & 0\end{pmatrix}$ (b) T is **not** one-to-one. (c) T is **onto**

5.5. (a) S is a subspace of $\mathbb{R}^2$.

(b) (i) $T(\vec{e}_1)=\begin{bmatrix}1\\ -2\end{bmatrix}$, $T(\vec{e}_2)=\begin{bmatrix}-1\\ -3\end{bmatrix}$ (ii) $\mathbf{A}=\begin{bmatrix}1 & -1\\ -2 & -3\end{bmatrix}$ (iii) $\mathbf{w}=\begin{bmatrix}-1\\ -2\end{bmatrix}$

5.6. (a) $\mathbf{A}=\begin{pmatrix}1 & -4 & 2\\ 2 & 7 & -1\\ -1 & -8 & 2\\ 2 & 1 & 1\end{pmatrix}$ (b) $\left\{\begin{pmatrix}1\\ 2\\ -1\\ 2\end{pmatrix}, \begin{pmatrix}0\\ 5\\ -4\\ 3\end{pmatrix}\right\}$

5.7. (d) $\ker\ T=t\begin{pmatrix}-1\\ 1\\ 0\end{pmatrix}$

5.8. (a) $B=\left\{\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}, \begin{pmatrix}0\\ 1\\ 0\end{pmatrix}, \begin{pmatrix}0\\ 0\\ 1\end{pmatrix}\right\}$ (b) T is onto

(c) $\left\{\begin{pmatrix}1\\ -1\\ 0\\ 0\end{pmatrix}\right\}$ (d) T is **not** one-to-one

5.9. $\begin{bmatrix}9\\ -3\\ 2\end{bmatrix}$

5.10. $\begin{bmatrix}1\\ -2\\ -4\end{bmatrix}$

5.11. (a) We have $T\ :\ \mathbb{R}^5\to\mathbb{R}^6$ which gives $m=6$ and $n=5$. (b) 6
(c) T is **not** onto (d) T is one-to-one

5.12. For **all** real h and k provided $h\neq 0$ and $k\neq\frac{3}{2}$.

5.13. (a) zero function (c) $\ker(T)=Ae^{2x}$ and dim is 1.

5.14. (c) $\dim\big(\ker(T)\big)=1$ and $\dim\big(image\,(T)\big)=2$

5.15. Consider $T\left(\begin{bmatrix}x\\ y\end{bmatrix}\right)=\begin{pmatrix}x-y\\ x-y\end{pmatrix}$ with $\mathbf{v}_1=\begin{pmatrix}1\\ 1\end{pmatrix}$ and $\mathbf{v}_2=\begin{pmatrix}2\\ 1\end{pmatrix}$.

5.16. (b) $\mathbf{A}=\begin{pmatrix}0 & 3 & 0\\ 0 & 0 & 2\\ 1 & 0 & 0\\ 0 & 1 & 1\end{pmatrix}$ (c) $15-2t+2t^2+4t^3$ (d) yes

5.17. (b) (ii) $\mathbf{S}=\begin{bmatrix}1 & 2 & 0\\ 0 & 2 & -3\\ 1 & 0 & 1\\ 1 & 0 & 0\end{bmatrix}$

5.18. (b) $\begin{bmatrix} 1/2 \\ 1/2 \\ 1 \\ -1 \end{bmatrix}$ (c) $\begin{bmatrix} 1/2 & -1/2 & 0 & 0 \\ 1/2 & -1/2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

5.19. (a) (i) f is a linear map (ii) f is **not** a linear map

(b) $Rank\,(f) + Nullity\,(f) = n$

(c) f is **not** injective nor surjective

(d) f is **not** injective nor surjective.

A basis for $\ker(f)$ is $\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\}$ and basis for image is $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

5.20. Option C.

5.21. $\mathbf{C} = \mathbf{BA} = \begin{pmatrix} 9 & -6 & 15 \\ 20 & -1 & 0 \\ 15 & 5 & 5 \end{pmatrix}$

5.22. (a) $\begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$ (b) $\det(\mathbf{A}) = 27$

(c) A basis for $\mathrm{Im}(\phi)$ is $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$. No basis for $\ker(\phi)$.

(d) ϕ is invertible since $\det(\mathbf{A}) = 27 \neq 0$

5.23. (b) $[T]_B = \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}$ (c) $[T]_B^{-1} = \frac{1}{2}\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}$ (d) $f(x) = \frac{1}{2}\sin x - \frac{5}{2}\cos x$

5.26. True.

SOLUTIONS TO EXERCISES 6.1

1. (a) 4 (b) 1 (c) 1 (d) 0

2. (a) -8 and -8 (b) -13 and -13 (c) $\det(\mathbf{A}) = 0$ and $\det(\mathbf{B}) = 0$

3. In both cases $\det(\mathbf{A}) = \det(\mathbf{B}) = ad - bc$.

4. 6 and -6.

5. $x = 2,\ y = -2$.

6. $\det\left(\mathbf{A}^2\right) = \det(\mathbf{A}) \times \det(\mathbf{A})$.

8. not linear

SOLUTIONS TO EXERCISES 6.2

1. (a) -117 (b) 1288 (c) -114

3. $x = -13.38, 5.38$

4. $C = \begin{pmatrix} 7 & 42 & -16 \\ -5 & -30 & 1 \\ -15 & -17 & 3 \end{pmatrix}$, $C^T = \begin{pmatrix} 7 & -5 & -15 \\ 42 & -30 & -17 \\ -16 & 1 & 3 \end{pmatrix}$,

$$\mathbf{A}^{-1} = -\frac{1}{73}\begin{pmatrix} 7 & -5 & -15 \\ 42 & -30 & -17 \\ -16 & 1 & 3 \end{pmatrix}$$

5. (a) $\mathbf{A}^{-1} = \begin{pmatrix} 3 & -2 \\ -13 & 9 \end{pmatrix}$ (b) $\mathbf{B}^{-1} = \begin{pmatrix} 5 & -7 \\ -12 & 17 \end{pmatrix}$

(c) $\mathbf{C}^{-1} = -\frac{1}{7}\begin{pmatrix} 1 & -4 \\ -3 & 5 \end{pmatrix}$ (d) $\mathbf{D}^{-1} = -\frac{1}{147}\begin{pmatrix} 33 & 37 & 32 \\ 60 & 45 & 27 \\ 18 & 38 & 13 \end{pmatrix}$

6. (a) -42 (b) 44 (c) 13 (d) -85

7. (a) 13 (b) 34.5 (c) 0

8. The place sign for a_{31}, a_{56}, a_{62} and a_{65} is $1, -1, 1$ and -1 respectively. There is **no** a_{71} entry in a 6 by 6 matrix.

10. $\det(\mathbf{A}) = 0$

11. All real values of k provided $k \neq \sqrt[3]{-10}$.

14. 7

18. 1

SOLUTIONS TO EXERCISES 6.3

1. (a) $\det(\mathbf{A}) = -10$ (b) $\det(\mathbf{B}) = -1$ (c) $\det(\mathbf{C}) = 1$ (d) $\det(\mathbf{D}) = -1$
(e) $\det(\mathbf{E}) = -0.6$ (f) $\det(\mathbf{F}) = 1$

2. (a) 6 (b) 6 (c) 6 (d) -72 (e) 240 000 (f) -945
(g) impossible

3. (a) $\det(\mathbf{A}) = \alpha\beta\gamma$ (b) $\det(\mathbf{B}) = \sin(2\theta)$ (c) $\det(\mathbf{C}) = xyz$

4. -1.38(2 dp) and 3.63 (2 dp).

5. (a) -27 (b) 2 (c) -39

6. (a) 18 (b) 0

8. (a) 7 (b) $-5/378$ (c) 600

9. (a) 96 (b) 995 328 (c) 1

10. invertible

11. (a) and (c) are negative. (b) is zero.

SOLUTIONS TO EXERCISES 6.4

1. (a) $x_1 = 1,\ x_2 = -1$ and $x_3 = 3$ (b) $x_1 = 2,\ x_2 = 1$ and $x_3 = -1$
(c) $x_1 = -2,\ x_2 = 3$ and $x_3 = 2$ (d) $x_1 = -2,\ x_2 = -3,\ x_3 = 1$

2. $x_1 = 1,\ x_2 = -2,\ x_3 = 3,\ x_4 = -4$

3. 15 984

4. $\mathbf{L} = \mathbf{I},\ \mathbf{U} = \mathbf{A}.$

5. (a) $\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 4 & 5 & 1 \end{pmatrix}$, $\mathbf{U} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}$ (b) $\mathbf{A}^{-1} = \begin{pmatrix} 84 & -37 & 7 \\ -58 & 26 & -5 \\ 11 & -5 & 1 \end{pmatrix}$
(c) (i) $x_1 = 31,\ x_2 = -21$ and $x_3 = 4$ (ii) $x_1 = -181,\ x_2 = 131$ and $x_3 = -25$

SOLUTIONS TO MISCELLANEOUS EXERCISES 6

6.1. $\det(\mathbf{A}) = 165$

6.2. $\det(\mathbf{E}) = -20,\ \det(\mathbf{F}) = 30,\ \det(\mathbf{EF}) = -600$ and $\det(\mathbf{E}+\mathbf{F}) = -6$

6.3. 64

6.4. (a) $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ (b) $\det\left(\mathbf{A}^{-1}\right) = \dfrac{1}{\det(\mathbf{A})}$ provided $\det(\mathbf{A}) \neq 0$
(c) $\det(\mathbf{A}+\mathbf{B}) =$ No Formula (d) $\det(3\mathbf{A}) = 3^n \det(\mathbf{A})$
(e) $\det\left(\mathbf{A}^T\right) = \det(\mathbf{A})$

6.5. 56

6.6. $x_2 = \det\begin{pmatrix} 2 & 8 & 1 \\ 3 & 1 & -1 \\ 4 & 10 & 3 \end{pmatrix} \Big/ \det\begin{pmatrix} 2 & 1 & 1 \\ 3 & -2 & -1 \\ 4 & -7 & 3 \end{pmatrix}$

6.7. $x_2 = \dfrac{1}{210}$

6.8. $x_1 = 0, x_2 = -2$ and $x_3 = 2$

6.9. -448

6.10. 0

6.11. (a) $\det(\mathbf{A}) = -10$ (b) $\mathbf{A}^{-1} = \dfrac{1}{10}\begin{pmatrix} -2 & -14 & 6 \\ 1 & 12 & -3 \\ 3 & -4 & 1 \end{pmatrix}$
(c) $\det(\mathbf{A}) = -10$

6.12. (i) $\mathbf{A}^{-1} = \begin{pmatrix} 1 & -1 & 2 \\ 1 & 2 & 0 \\ 0 & -1 & 1 \end{pmatrix}$
(ii) $\left(\mathbf{A}^t\right)^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & -1 \\ 2 & 0 & 1 \end{pmatrix}$, $(3\mathbf{A})^{-1} = \dfrac{1}{3}\begin{pmatrix} 1 & -1 & 2 \\ 1 & 2 & 0 \\ 0 & -1 & 1 \end{pmatrix}$, $\left(\mathbf{A}^2\right)^{-1} = \begin{pmatrix} 0 & -5 & 4 \\ 3 & 3 & 2 \\ -1 & -3 & 1 \end{pmatrix}$

6.13. 3

6.14. For the 2 by 2 matrix: $adj\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $\det(\mathbf{A}) = 0$ which means that $\mathbf{A}^{-1}$ does not exist. 3 by 3 gives $adj\mathbf{A} = \begin{pmatrix} 1 & 18 & -11 \\ 1 & 13 & -8 \\ 0 & -2 & 1 \end{pmatrix}$, $\det \mathbf{A} = -1$ and

$$\mathbf{A}^{-1} = \begin{pmatrix} -1 & -18 & 11 \\ -1 & -13 & 8 \\ 0 & 2 & -1 \end{pmatrix}.$$

6.15. (a) -16 means that $\mathbf{A}$ is invertible. (b) $\det\left(\mathbf{A}^T\right) = -16$ (c) $\det\left(\mathbf{A}^{-1}\right) = -\frac{1}{16}$

6.16. $D_1 = -7, D_2 = 7, D_3 = 0$

6.17. (a) $\det(\mathbf{A}) = -24$ (b) $\det(\mathbf{B}) = 0$ (c) $\det(\mathbf{C}) = -2$

6.18. (a) $\det(\mathbf{A}) = -17$, $\det(\mathbf{B}) = -12$, $\det(\mathbf{AB}) = 204$, $\det\left(\mathbf{A}^3\right) = -4913$
(b) Take the determinants of $\mathbf{AA}^{-1} = \mathbf{I}$ to show the required result.

6.19. (a) $-(u_1v_1 + u_2v_2 + u_3v_3)$ (b) $100 - (u_1v_1 + u_2v_2 + u_3v_3)$

6.20. $(a-b)(a-c)(c-b)$

6.21. 16

6.22. $\frac{9}{4}$

6.23. (a) (i) 8 (ii) 18 (b) (i) 162 (ii) $-\frac{3}{5}$ (iii) $-\frac{12}{5}$

6.24. (a) (i) 24 (ii) $2abc$ (iii) 432 (b) (i) 4 (ii) -2 (iii) $-\frac{3}{2}$

6.25. (a) (i) 8/9 (ii) 12 (iii) 1 (b) 2

6.26. (a) 3 (b) 21 (c) 8 (d) 7

6.27. *adfpru*

6.28. $x = \frac{1}{3a-2}$, $y = \frac{3a-3}{3a-2}$, $z = \frac{1-a}{3a-2}$

6.29. Apply this $\det(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3\cdots\mathbf{A}_n) = \det(\mathbf{A}_1)\det(\mathbf{A}_2)\det(\mathbf{A}_3)\cdots\det(\mathbf{A}_n)$ to get $\det\left(\mathbf{A}^5\right) = \det(\mathbf{A})\det(\mathbf{A})\cdots\det(\mathbf{A}) = 0$

6.30. (a) $\det(\mathbf{A}_2) = 2!$, $\det(\mathbf{A}_3) = 3! = 6$ and $\det(\mathbf{A}_4) = 4! = 24$ (b) $\det(\mathbf{A}_n) = n!$

6.31. $\mathbf{X}^{-1} = \frac{1}{ACF}\begin{bmatrix} CF & -BF & BE-CD \\ 0 & AF & -AE \\ 0 & 0 & AC \end{bmatrix}$

6.32. $y = \frac{1}{C}$

6.33. -125

6.34. $k \neq 3,\ k \neq 4$ or $k \neq -1$

6.35. $2^4/15^7$

6.36. Not possible

6.37. Use properties of determinants.

6.38. Convert the given matrix into an upper triangular matrix by swapping rows. The number of swaps is $\lfloor n/2 \rfloor$, which means we multiply the resulting determinant which is the product of the leading diagonal by $(-1)^{\lfloor n/2 \rfloor}$.

SOLUTIONS TO EXERCISES 7.1

1. (a) $\lambda_1 = -4,\ \mathbf{u} = \begin{pmatrix} -3/11 \\ 1 \end{pmatrix}$ and $\lambda_2 = 7,\ \mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

(b) $\lambda_1 = 1,\ \mathbf{u} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\lambda_2 = 3,\ \mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

(c) $\lambda_1 = -3,\ \mathbf{u} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ and $\lambda_2 = 3,\ \mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

2. (a) $\lambda_1 = 4, \mathbf{u} = s\begin{pmatrix} 20 \\ 9 \\ 15 \end{pmatrix}$ and $\lambda_2 = -5, \mathbf{v} = s\begin{pmatrix} -2/3 \\ 0 \\ 1 \end{pmatrix}$

3. $\lambda_1 = 2, \mathbf{u} = s\begin{pmatrix} -1 \\ 1 \end{pmatrix}, E_2 = \left\{ s\begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$ and $\lambda_2 = 4 \quad \mathbf{v} = s\begin{pmatrix} 1 \\ 1 \end{pmatrix}, E_4 = \left\{ s\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$

4. $\lambda_1 = -2, \mathbf{u} = s\begin{pmatrix} 2 \\ 7 \end{pmatrix}, E_{-2} = \left\{ s\begin{pmatrix} 2 \\ 7 \end{pmatrix} \right\}$ and $\lambda_2 = 3, \mathbf{v} = s\begin{pmatrix} 1 \\ 1 \end{pmatrix}, E_3 = \left\{ s\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$

A basis vector for E_{-2} is $\begin{pmatrix} 2 \\ 7 \end{pmatrix}$ and a basis vector for E_3 is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

5. (a) The eigenvalues of $\mathbf{A}$ are $\lambda_1 = -7, \quad \lambda_2 = 6$.
(b) The eigenvalues of $\mathbf{B}$ are $\lambda_3 = -14, \quad \lambda_4 = 12$.
(c) If t is the eigenvalue of matrix $\mathbf{B}$ and λ is the eigenvalue of matrix $\mathbf{A}$ then $t = 2\lambda$.

8. $\lambda_1 = 1, \mathbf{u} = s\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$, basis vector is $\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$. $\quad \lambda_2 = 4, \mathbf{v} = s\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$,

basis vector is $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. $\lambda_3 = -1, \quad \mathbf{w} = s\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, basis vector is $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

SOLUTIONS TO EXERCISES 7.2

The variables are non-zero in the following answers.

1. A basis vector for E_3 is $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ which corresponds to $\lambda_{1,\,2} = 3$.

2. A basis vector for E_{-5} is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ which corresponds to $\lambda_{1,\,2} = -5$.

5. A basis vector for E_1 is $\begin{pmatrix}0\\0\\1\end{pmatrix}$ which correspond to $\lambda_{1,\,2,\,3} = 1$.

6. (a) Basis vectors for E_5 is $B = \left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right\}$ which corresponds to $\lambda_{1,\,2} = 5$ and a basis vector for E_2 is $B = \left\{\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ which corresponds to $\lambda_3 = 2$.

(b) For $\lambda_1 = 1$, $B = \left\{\begin{pmatrix}1\\0\\0\end{pmatrix}\right\}$. For $\lambda_2 = 5$, $B = \left\{\begin{pmatrix}1\\2\\0\end{pmatrix}\right\}$ and for $\lambda_3 = 9$,

$B = \left\{\begin{pmatrix}7\\4\\16\end{pmatrix}\right\}$

(c) For $\lambda_{1,\,2,\,3} = -2$ and a basis is $B = \left\{\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$

7. (a) $\lambda_{1,\,2} = 7$, $\mathbf{u} = s\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + t\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$ and $\lambda_{3,\,4} = 5$, $\mathbf{v} = s\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix}$

(b) $\lambda_{1,\,2,\,3} = 1$, $\mathbf{u} = \begin{pmatrix}s\\t\\r\\0\end{pmatrix} = s\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + t\begin{pmatrix}0\\1\\0\\0\end{pmatrix} + r\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$ and $\lambda_4 = 3$, $\mathbf{v} = s\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$

(c) $\lambda_{1,\,2,\,3,\,4} = 3$, $\mathbf{u} = s\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + t\begin{pmatrix}0\\1\\0\\0\end{pmatrix} + r\begin{pmatrix}0\\0\\1\\0\end{pmatrix} + q\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$

11. (a) (i) Eigenvalues $\lambda_1 = 1,\ \lambda_2 = 2,\ \lambda_3 = 3$ and $\lambda_4 = 4$.
(ii) Eigenvalues of $\mathbf{A}^5$ are $(\lambda_1)^5 = 1,\ (\lambda_2)^5 = 32,\ (\lambda_3)^5 = 243$ and $(\lambda_4)^5 = 1024$
(iii) Eigenvalues of $\mathbf{A}^{-1}$ are $(\lambda_1)^{-1} = 1,\ (\lambda_2)^{-1} = \frac{1}{2},\ (\lambda_3)^{-1} = \frac{1}{3}$ and $(\lambda_4)^{-1} = \frac{1}{4}$
(iv) $\det(\mathbf{A}) = 24$ (v) $tr(\mathbf{A}) = 10$
(b) (i) $\lambda_1 = -1,\ \lambda_2 = 6,\ \lambda_3 = -8$ and $\lambda_4 = 3$.
(ii) $(\lambda_1)^5 = -1,\ (\lambda_2)^5 = 7776,\ (\lambda_3)^5 = -32768$ and $(\lambda_4)^5 = 243$
(iii) Eigenvalues of $\mathbf{A}^{-1}$ are $(\lambda_1)^{-1} = -1,\ (\lambda_2)^{-1} = \frac{1}{6},\ (\lambda_3)^{-1} = -\frac{1}{8}$ and $(\lambda_4)^{-1} = \frac{1}{3}$
(iv) $\det(\mathbf{A}) = 144$ (v) $tr(\mathbf{A}) = 0$
(c) (i) $\lambda_1 = 2,\ \lambda_2 = -4,\ \lambda_3 = -7$ and $\lambda_4 = 0$.

(ii) Eigenvalues of $\mathbf{A}^5$ $(\lambda_1)^5 = 32,\ (\lambda_2)^5 = -1024,\ (\lambda_3)^5 = -16807$ and $(\lambda_4)^5 = 0$

(iii) $\mathbf{A}^{-1}$ does *not* exist.

(iv) $\det(\mathbf{A}) = 0$ (v) $tr(\mathbf{A}) = -9$

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

13. $\mathbf{A}^2 = \begin{pmatrix} 51 & 50 \\ 30 & 31 \end{pmatrix}$ and $\mathbf{A}^3 = \begin{pmatrix} 456 & 455 \\ 273 & 274 \end{pmatrix}$

14. $\mathbf{A}^{-1} = -\frac{1}{4}\left(\mathbf{A}^2 - 4\mathbf{A} - \mathbf{I}\right)$ and $\mathbf{A}^4 = 17\mathbf{A}^2 - 16\mathbf{I}$

SOLUTIONS TO EXERCISES 7.3

1. (i) $\lambda_1 = 1, \mathbf{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\lambda_2 = 2, \mathbf{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (ii) $\mathbf{P} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I}, \mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \mathbf{A}$

(b) (i) $\lambda_1 = 0, \mathbf{u} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $\lambda_2 = 2, \mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (ii) $\mathbf{P} = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$

(c) (i) $\lambda_1 = 3, \mathbf{u} = \begin{pmatrix} 1 \\ -4 \end{pmatrix}$ and $\lambda_2 = 4, \mathbf{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (ii) $\mathbf{P} = \begin{pmatrix} 1 & 0 \\ -4 & 1 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$

(d) (i) $\lambda_1 = 1, \mathbf{u} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ and $\lambda_2 = 4, \mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (ii) $\mathbf{P} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$

2. (i) (a) $\begin{pmatrix} 1 & 0 \\ 0 & 32 \end{pmatrix}$ (b) $\begin{pmatrix} 16 & 16 \\ 16 & 16 \end{pmatrix}$ (c) $\begin{pmatrix} 243 & 0 \\ 3124 & 1024 \end{pmatrix}$ (d) $\begin{pmatrix} 342 & 682 \\ 341 & 683 \end{pmatrix}$

(ii) $\begin{pmatrix} 1/\sqrt{3} & 0 \\ 2 - 4/\sqrt{3} & 1/2 \end{pmatrix}$

3. (a) (i) $\lambda_1 = 1, \mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \lambda_2 = 2, \mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\lambda_3 = 3, \mathbf{w} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

(ii) $\mathbf{P} = \mathbf{I}, \mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ (iii) $\mathbf{A}^4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 81 \end{pmatrix}$

(b) (i) $\lambda_1 = -1, \mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \lambda_2 = 4, \mathbf{v} = \begin{pmatrix} 4 \\ 5 \\ 0 \end{pmatrix}$ and $\lambda_3 = 5, \mathbf{w} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$

(ii) $\mathbf{P} = \begin{pmatrix} 1 & 4 & 2 \\ 0 & 5 & 3 \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ (iii) $\mathbf{A}^4 = \begin{pmatrix} 1 & 204 & 636 \\ 0 & 256 & 1107 \\ 0 & 0 & 625 \end{pmatrix}$

(c) (i) $\lambda_1 = 2, \mathbf{u} = \begin{pmatrix} -12 \\ 4 \\ 1 \end{pmatrix}, \lambda_2 = 5, \mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$ and $\lambda_3 = 6, \mathbf{w} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

(ii) $\mathbf{P} = \begin{pmatrix} -12 & 0 & 0 \\ 4 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix}$ (iii) $\mathbf{A}^4 = \begin{pmatrix} 16 & 0 & 0 \\ 203 & 625 & 0 \\ 554 & 1342 & 1296 \end{pmatrix}$

4. All (a) (b) and (c) are diagonalizable.

5. $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$

6. $\mathbf{A}$ is diagonalizable, $\mathbf{D} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, $\mathbf{P} = \begin{pmatrix} 1 & 5 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

$\mathbf{A}^3 = \begin{pmatrix} -203 & 97.5 & 0 \\ -156 & 70 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

7. (a) Only one eigenvalue $\lambda = 3$ and one independent eigenvector $\mathbf{u} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

(b) Only one independent e.vector $\mathbf{u} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ to the only eigenvalue $\lambda = -4$.

(c) Only 1 linearly independent eigenvector $\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^T$

8. (i) $\mathbf{A}^{11} = \begin{pmatrix} -2050 & 1366 \\ -6147 & 4097 \end{pmatrix}$ (ii) $\frac{1}{2}\begin{pmatrix} -5 & 2 \\ -9 & 4 \end{pmatrix}$

12. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 3 & 5 \\ 0 & 2 \end{pmatrix} t + \begin{pmatrix} 9 & 25 \\ 0 & 4 \end{pmatrix} \frac{t^2}{2!} + \begin{pmatrix} 27 & 95 \\ 0 & 8 \end{pmatrix} \frac{t^3}{3!} + \begin{pmatrix} 81 & 325 \\ 0 & 16 \end{pmatrix} \frac{t^4}{4!} + \cdots$

13. $\mathbf{P} = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}$, $\mathbf{P}^{-1} = \frac{1}{\lambda_1 - \lambda_2}\begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$

where $\lambda_1 = \frac{1+\sqrt{5}}{2}$ and $\lambda_2 = \frac{1-\sqrt{5}}{2}$.

19. $\mathbf{O}$

SOLUTIONS TO EXERCISES 7.4

1. (a) $\mathbf{Q} = \mathbf{I}$ and $\mathbf{D} = \mathbf{A}$ (b) $\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$

(c) $\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ (d) $\mathbf{Q} = \frac{1}{\sqrt{13}}\begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 13 & 0 \\ 0 & -13 \end{pmatrix}$

2. (a) $\mathbf{Q} = \frac{1}{\sqrt{10}}\begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 0 & 0 \\ 0 & 10 \end{pmatrix}$ (b) $\mathbf{Q} = \frac{1}{\sqrt{6}}\begin{pmatrix} 2 & \sqrt{2} \\ \sqrt{2} & -2 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$

(c) $\mathbf{Q} = \frac{1}{\sqrt{12}}\begin{pmatrix} -3 & \sqrt{3} \\ \sqrt{3} & 3 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} -6 & 0 \\ 0 & -2 \end{pmatrix}$ (d) $\mathbf{Q} = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}$,

$\mathbf{D} = \begin{pmatrix} -1 & 0 \\ 0 & 7 \end{pmatrix}$

3. (a) $\mathbf{Q}=\mathbf{I}$ and $\mathbf{D}=\mathbf{A}$ (b) $\mathbf{Q}=\begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \end{pmatrix}, \mathbf{D}=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 6 \end{pmatrix}$

(c) $\mathbf{Q}=\dfrac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}, \mathbf{D}=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

4. (a) $\mathbf{Q}=\begin{pmatrix} -2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}, \mathbf{D}=\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$

(b) $\mathbf{Q}=\begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}, \mathbf{D}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$

(c) $\mathbf{Q}=\begin{pmatrix} -1/\sqrt{2} & 1/3\sqrt{2} & 2/3 \\ 1/\sqrt{2} & 1/3\sqrt{2} & 2/3 \\ 0 & -4/3\sqrt{2} & 1/3 \end{pmatrix}, \mathbf{D}=\begin{pmatrix} -9 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

8. $\mathbf{Q}=\begin{pmatrix} \dfrac{b}{\sqrt{b^2+(\lambda_1-a)^2}} & \dfrac{\lambda_2-c}{\sqrt{(\lambda_2-c)^2+b^2}} \\ \dfrac{\lambda_1-a}{\sqrt{b^2+(\lambda_1-a)^2}} & \dfrac{b}{\sqrt{(\lambda_2-c)^2+b^2}} \end{pmatrix}$ where λ_1 and λ_2 are eigenvalues of $\mathbf{A}$.

9. By taking the inverse of $\mathbf{Q}^T\mathbf{AQ}=\mathbf{D}$ show that $\mathbf{D}^{-1}=\mathbf{Q}^{-1}\mathbf{A}^{-1}\mathbf{Q}$.

SOLUTIONS TO EXERCISES 7.5

1. (a) $\mathbf{U}=\mathbf{I}, \mathbf{D}=\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \mathbf{V}=\mathbf{I}$ (b) $\mathbf{U}=\dfrac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix}$,

$\mathbf{V}^T=\dfrac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$

(c) $\mathbf{U}=\begin{pmatrix} 1/\sqrt{30} & -2/\sqrt{5} & 1/\sqrt{6} \\ 2/\sqrt{30} & 1/\sqrt{5} & 2/\sqrt{6} \\ 5/\sqrt{30} & 0 & -1/\sqrt{6} \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} \sqrt{6} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \mathbf{V}^T=\dfrac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$

(d) $\mathbf{U}=\begin{pmatrix} 6/\sqrt{180} & -2/\sqrt{5} \\ 12/\sqrt{180} & 1/\sqrt{5} \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$,

$\mathbf{V}^T=\begin{pmatrix} 1/\sqrt{30} & 2/\sqrt{30} & 5/\sqrt{30} \\ -2/\sqrt{5} & 1/\sqrt{5} & 0 \\ 1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \end{pmatrix}$

(e) $\mathbf{U}=\dfrac{1}{\sqrt{10}}\begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} \sqrt{20} & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{V}^T=\dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

(f) $\mathbf{U}=\dfrac{1}{\sqrt{10}}\begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} \sqrt{30} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{V}^T=\begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{pmatrix}$

SOLUTIONS TO MISCELLANEOUS EXERCISES 7

7.1. (a) $\lambda_1 = 1$ and $\lambda_2 = 3$ (b) $\mathbf{Q} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$

(c) $\mathbf{A}^5 = \begin{pmatrix} 1 & 121 \\ 0 & 243 \end{pmatrix}$

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

7.3. $\lambda_{1,2} = -2$ and $\lambda_3 = 4$, E_{-2} is $\left\{\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\right\}$, E_4 is $\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$

7.4. λ^{-1} and linearly independent eigenvectors.

7.5. (a) $\mathbf{S} = \begin{pmatrix} 5 & 1 \\ 1 & -3 \end{pmatrix}$, $\Lambda = \begin{pmatrix} 8 & 0 \\ 0 & -8 \end{pmatrix}$ (b) $\mathbf{B} = \frac{1}{4}\begin{pmatrix} 7 & 5 \\ 3 & -7 \end{pmatrix} = \frac{1}{4}\mathbf{A}$

7.6. (a) (i) $\lambda_1 = 3$, $\mathbf{u} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ and $\lambda_2 = 1$, $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (ii) $\mathbf{P} = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$

(iii) The eigenvalues of $\mathbf{A}^{2008}$ are $\lambda_1 = 1$, $\lambda_2 = 3^{2008}$ and $\det\left(\mathbf{A}^{2008}\right) = 3^{2008}$.

(b) We have three linearly independent eigenvectors for a 3 by 3 matrix so the matrix is diagonalisable.

7.7. (a) $\mathbf{S} = \begin{pmatrix} -1 & 0 \\ 10 & 1 \end{pmatrix}$ (b) $\mathbf{S} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

7.8. Eigenvalues are distinct $\lambda_1 = \sqrt{2}$ and $\lambda_2 = -\sqrt{2}$ therefore the matrix $\mathbf{B}$ is diagonalisable. The matrix $\mathbf{P} = \begin{pmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{pmatrix}$ and $\mathbf{D} = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{pmatrix}$.

7.9. (a) $\left(4, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right)$ (b) $\mathbf{P} = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ and $\mathbf{D} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$

7.10. Show that $\mathbf{u}^T\mathbf{v} = 0$ by taking the transpose of $\mathbf{Au} = \lambda_1\mathbf{u}$.

7.11. (a) $\lambda_1 = 2$, $\mathbf{u} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ and $\lambda_2 = 7$, $\mathbf{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ (b) $\begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$ with eigenvalue 2

7.12. (a) $\lambda_1 = 1$, $\mathbf{u} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$, $\lambda_2 = 2$, $\mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ and $\lambda_3 = 0$, $\mathbf{w} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$

(b) We have distinct eigenvalues therefore $\mathbf{A}$ is diagonalizable.

7.13. (a) (i) Substitute these into $\mathbf{Ax} = \lambda\mathbf{x}$. (ii) $\mathbf{Q} = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \end{pmatrix}$

(iii) $\mathbf{A}^3 = \begin{pmatrix} 9 & 9 & 9 \\ 9 & 9 & 9 \\ 9 & 9 & 9 \end{pmatrix}$

(b) Show that they have the same $p_{B^T}(\lambda) = p_B(\lambda)$.

7.14. (b) $\lambda_1 = 3$ (c) $\mathbf{P} = \begin{pmatrix} 1 & -1 & -2 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$, $\mathbf{P}^{-1} = \frac{1}{4}\begin{pmatrix} 1 & 2 & 1 \\ -1 & -2 & 3 \\ -1 & 2 & -1 \end{pmatrix}$

7.15. (a) $\mathbf{P} = \begin{pmatrix} -1 & 1 & -1 \\ 4 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$, $\mathbf{P}^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ -2 & -2 & 3 \\ -4 & -3 & 4 \end{pmatrix}$, $\mathbf{P}^{-1}\mathbf{AP} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

(b) Check $p(\mathbf{A}) = (\mathbf{A} - 5\mathbf{I})(\mathbf{A} - 3\mathbf{I})^2 = \mathbf{O}$

7.16. (b) Real eigenvalues of $\mathbf{A}$ are $t_1 = 1$ and $t_2 = 9$. A basis for E_1 is $\begin{pmatrix} 0 \\ 0 \\ 9 \\ -1 \end{pmatrix}$.

7.17. $\lambda_{1,2} = 1$, $\lambda_{3,4} = -1$ and $\lambda_5 = -3$

7.18. (a) $\lambda_1 = 0$, $\lambda_2 = 1$ and $\lambda_3 = 3$

(b) $\mathbf{u} = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$

(c) $\beta = \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \right\}$, $\mathbf{D} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

(d) $\mathbf{S} = (\mathbf{w} \;\; \mathbf{v} \;\; \mathbf{u}) = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}$

(e) $\mathbf{S}^{-1} = \frac{1}{6}\begin{pmatrix} 1 & 2 & 1 \\ 3 & 0 & -3 \\ -2 & 2 & -2 \end{pmatrix}$

(f) $\beta' = \left\{ \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \right\}$

(g) $\mathbf{S}' = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \end{pmatrix}$

(h) $(\mathbf{S}')^{-1} = \begin{pmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ -1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \end{pmatrix}$

7.19. (a) $\mathbf{S} = \begin{pmatrix} 1 & b \\ 0 & c-a \end{pmatrix}$, $\Lambda = \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix}$

(b) $\mathbf{A}^{1000} = \begin{pmatrix} a^{1000} & b\left(c^{999} + c^{998}a + c^{997}a^2 + \cdots + ca^{998} + a^{999}\right) \\ 0 & c^{1000} \end{pmatrix}$

7.20. (d) The algebraic multiplicity of $\lambda_1 = 1$ is one and $\lambda_{2,\,3} = 2$ is two.
Matrix **A** is diagonalizable if and only if $\lambda_1 = 1$ has geometric multiplicity of 1 and $\lambda_{2,\,3} = 2$ has geometric multiplicity of 2.

(e) (i) is **not** diagonalizable (ii) $\mathbf{P} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$

7.21. (b) $\widehat{\mathbf{w}_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\widehat{\mathbf{w}'_2} = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ and $\widehat{\mathbf{w}'_3} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

(c) (i) $\lambda^3 - 6\lambda^2 - 15\lambda - 8 = (\lambda + 1)^2(\lambda - 8) = 0$. The eigenvalues are $\lambda_{1,\,2} = -1$ and $\lambda_3 = 8$.

(ii) $\left\{ \begin{pmatrix} 1 \\ -4 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \right\}$

(c) Consider $k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 = \mathbf{O}$. Prove that $k_1 = k_2 = k_3 = 0$.

7.22. (b) See chapter 4.

(c) (i) $\lambda_1 = 4$ and $\lambda_2 = 6$ (ii) $\lambda_3 = 0$, $\mathbf{w} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ and $\lambda_4 = 2$, $\mathbf{x} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$

(iii) $\mathbf{P} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$, $\mathbf{Q} = \mathbf{P}^T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{pmatrix}$,

$$\Lambda = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

7.23. $X = \frac{1}{\sqrt{5}}(2x - y)$ and $Y = \frac{1}{\sqrt{5}}(x + 2y)$. The diagonal form is
$3x^2 + 4xy + 6y^2 = 2X^2 + 7Y^2$

7.24. $X = \frac{1}{\sqrt{3}}(x + y - z)$, $Y = \frac{1}{\sqrt{2}}(y - x)$ and $Z = \frac{1}{\sqrt{6}}(x + y + 2z)$.
The diagonal form is $2xy + 4xz + 4yz + 3z^2 = -X^2 - Y^2 + 5Z^2$.

Index

W

X

Y

Z