Contents 22

Eigenvalues and Eigenvectors

Learning outcomes

In this Workbook you will learn about the matrix eigenvalue problem AX = kX where A is a square matrix and k is a scalar (number). You will learn how to determine the eigenvalues (k) and corresponding eigenvectors (X) for a given matrix A. You will learn of some of the applications of eigenvalues and eigenvectors. Finally you will learn how eigenvalues and eigenvectors may be determined numerically.

Basic Concepts

22.1

Introduction

From an applications viewpoint, eigenvalue problems are probably the most important problems that arise in connection with matrix analysis. In this Section we discuss the basic concepts. We shall see that eigenvalues and eigenvectors are associated with square matrices of order $n \times n$. If n is small (2 or 3), determining eigenvalues is a fairly straightforward process (requiring the solutiuon of a low order polynomial equation). Obtaining eigenvectors is a little strange initially and it will help if you read this preliminary Section first.

Prerequisites

Before starting this Section you should . . .

- have a knowledge of determinants and matrices
- have a knowledge of linear first order differential equations

Learning Outcomes

On completion you should be able to . . .

- obtain eigenvalues and eigenvectors of 2×2 and 3×3 matrices
- state basic properties of eigenvalues and eigenvectors

1. Basic concepts

Determinants

A square matrix possesses an associated determinant. Unlike a matrix, which is an array of numbers, a determinant has a single value.

A two by two matrix $C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$ has an associated determinant

$$\det\,(C) = \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} = c_{11}\,c_{22} - c_{21}\,c_{12}$$

(Note square or round brackets denote a matrix, straight vertical lines denote a determinant.)

A three by three matrix has an associated determinant

$$\det(C) = \begin{vmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{vmatrix}$$

Among other ways this determinant can be evaluated by an "expansion about the top row":

$$\det(C) = c_{11} \begin{vmatrix} c_{22} & c_{23} \\ c_{32} & c_{33} \end{vmatrix} - c_{12} \begin{vmatrix} c_{21} & c_{23} \\ c_{31} & c_{33} \end{vmatrix} + c_{13} \begin{vmatrix} c_{21} & c_{22} \\ c_{31} & c_{32} \end{vmatrix}$$

Note the minus sign in the second term.

Task

Evaluate the determinants

$$\det(A) = \begin{vmatrix} 4 & 6 \\ 3 & 1 \end{vmatrix} \qquad \det(B) = \begin{vmatrix} 4 & 8 \\ 1 & 2 \end{vmatrix} \qquad \det(C) = \begin{vmatrix} 6 & 5 & 4 \\ 2 & -1 & 7 \\ -3 & 2 & 0 \end{vmatrix}$$

Your solution

Answer

$$\det A = 4 \times 1 - 6 \times 3 = -14 \qquad \det B = 4 \times 2 - 8 \times 1 = 0$$

$$\det C = 6 \begin{vmatrix} -1 & 7 \\ 2 & 0 \end{vmatrix} - 5 \begin{vmatrix} 2 & 7 \\ -3 & 0 \end{vmatrix} + 4 \begin{vmatrix} 2 & -1 \\ -3 & 2 \end{vmatrix} = 6 \times (-14) - 5(21) + 4(4 - 3) = -185$$

A matrix such as $B = \begin{bmatrix} 4 & 8 \\ 1 & 2 \end{bmatrix}$ in the previous task which has zero determinant is called a **singular** matrix. The other two matrices A and C are **non-singular**. The key factor to be aware of is as follows:

Key Point 1

Any non-singular $n \times n$ matrix C, for which $\det(C) \neq 0$, possesses an inverse C^{-1} i.e.

$$CC^{-1} = C^{-1}C = I \quad \text{where } I \text{ denotes the } n \times n \text{ identity matrix}$$

A singular matrix does **not** possess an inverse.

Systems of linear equations

We first recall some basic results in linear (matrix) algebra. Consider a system of n equations in n unknowns $x_1, x_2, \ldots, x_n$:

$$\begin{array}{ccccccccc} c_{11}x_1 & + & c_{12}x_2 & + & \ldots & + & c_{1n}x_n & = & k_1 \\ c_{21}x_1 & + & c_{22}x_2 & + & \ldots & + & c_{2n}x_n & = & k_2 \\ \vdots & + & \vdots & + & \ldots & + & \vdots & = & \vdots \\ c_{n1}x_1 & + & c_{n2}x_2 & + & \ldots & + & c_{nn}x_n & = & k_n \end{array}$$

We can write such a system in matrix form:

$$\begin{bmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}, \qquad \text{or equivalently} \qquad CX = K.$$

We see that C is an $n \times n$ matrix (called the coefficient matrix), $X = \{x_1, x_2, \ldots, x_n\}^T$ is the $n \times 1$ column vector of unknowns and $K = \{k_1, k_2, \ldots, k_n\}^T$ is an $n \times 1$ column vector of given constants.

The zero matrix will be denoted by $\underline{O}$.

If $K \neq \underline{O}$ the system is called **inhomogeneous**; if $K = \underline{O}$ the system is called **homogeneous**.

Basic results in linear algebra

Consider the system of equations $CX = K$.

We are concerned with the nature of the solutions (if any) of this system. We shall see that this system only exhibits three solution types:

- The system is consistent and has a unique solution for X
- The system is consistent and has an infinite number of solutions for X
- The system is inconsistent and has no solution for X

There are two basic cases to consider:

det$(C) \neq 0$ or det$(C) = 0$

Case 1: det$(C) \neq 0$

In this case C^{-1} exists and the **unique** solution to $CX = K$ is

$$X = C^{-1}K$$

Case 2: det$(C) = 0$

In this case C^{-1} does not exist.

(a) If $K \neq \underline{O}$ the system $CA = K$ has **no solutions**.

(b) If $K = \underline{O}$ the system $CX = \underline{O}$ has an **infinite number of solutions.**

We note that a homogeneous system

$$CX = \underline{O}$$

has a unique solution $X = \underline{O}$ if det$(C) \neq 0$ (this is called the **trivial solution**) or an infinite number of solutions if det$(C) = 0$.

Example 1

(Case 1) Solve the inhomogeneous system of equations

$$x_1 + x_2 = 1 \qquad 2x_1 + x_2 = 2$$

which can be expressed as $CX = K$ where

$$C = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \qquad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad K = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Solution

Here $\det(C) = -1 \neq 0$.

The system of equations has the **unique solution**: $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Example 2

(Case 2a) Examine the following inhomogeneous system for solutions

$$x_1 + 2x_2 = 1$$
$$3x_1 + 6x_2 = 0$$

Solution

Here det $(C) = \begin{vmatrix} 1 & 2 \\ 3 & 6 \end{vmatrix} = 0$. In this case there are no solutions.

To see this we see the first equation of the system states $x_1 + 2x_2 = 1$ whereas the second equation (after dividing through by 3) states $x_1 + 2x_2 = 0$, a contradiction.

Example 3

(Case 2b) Solve the homogeneous system

$$x_1 + x_2 = 0$$
$$2x_1 + 2x_2 = 0$$

Solution

Here $\det(C) = \begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} = 0$. The solutions are **any** pairs of numbers $\{x_1, x_2\}$ such that $x_1 = -x_2$,

i.e. $X = \begin{bmatrix} \alpha \\ -\alpha \end{bmatrix}$ where α is arbitrary.

There are an **infinite number of solutions**.

A simple eigenvalue problem

We shall be interested in simultaneous equations of the form:

$$AX = \lambda X,$$

where A is an $n \times n$ matrix, X is an $n \times 1$ column vector and λ is a scalar (a constant) and, in the first instance, we examine some simple examples to gain experience of solving problems of this type.

Example 4

Consider the following system with $n = 2$:

$$\begin{aligned} 2x + 3y &= \lambda x \\ 3x + 2y &= \lambda y \end{aligned}$$

so that

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} x \\ y \end{bmatrix}.$$

It appears that there are three unknowns x, y, λ. The obvious questions to ask are: can we find x, y? what is λ?

Solution

To solve this problem we firstly re-arrange the equations (take all unknowns onto one side);

$$(2 - \lambda)x + 3y = 0 \qquad (1)$$

$$3x + (2 - \lambda)y = 0 \qquad (2)$$

Therefore, from equation (2):

$$x = -\frac{(2 - \lambda)}{3}y. \qquad (3)$$

Then when we substitute this into (1)

$$-\frac{(2 - \lambda)^2}{3}y + 3y = 0 \quad \text{which simplifies to} \quad [-(2 - \lambda)^2 + 9]\, y = 0.$$

We conclude that either $y = 0$ or $9 = (2 - \lambda)^2$. There are thus two cases to consider:

Case 1

If $y = 0$ then $x = 0$ (from (3)) and we get the **trivial solution**. (We could have guessed this solution at the outset.)

Case 2

$$9 = (2 - \lambda)^2$$

which gives, on taking square roots:

$$\pm 3 = 2 - \lambda \quad \text{giving} \quad \lambda = 2 \pm 3 \quad \text{so} \quad \lambda = 5 \quad \text{or} \quad \lambda = -1.$$

Now, from equation (3), if $\lambda = 5$ then $x = +y$ and if $\lambda = -1$ then $x = -y$.

We have now completed the analysis. We have found values for λ but we also see that we cannot obtain unique values for x and y: all we can find is the ratio between these quantities. This behaviour is typical, as we shall now see, of an eigenvalue problem.

2. General eigenvalue problems

Consider a given square matrix A. If X is a column vector and λ is a scalar (a number) then the relation.

$$AX = \lambda X \qquad (4)$$

is called an **eigenvalue problem**. Our purpose is to carry out an analysis of this equation in a manner similar to the example above. However, we will attempt a more general approach which will apply to **all** problems of this kind.

Firstly, we can spot an obvious solution (for X) to these equations. The solution $X = 0$ is a possibility (for then both sides are zero). We will not be interested in these **trivial solutions** of the eigenvalue problem. Our main interest will be in the occurrence of **non-trivial solutions** for X. These may exist for special values of λ, called the **eigenvalues** of the matrix A. We proceed as in the previous example:

take all unknowns to one side:

$$(A - \lambda I)X = 0 \qquad (5)$$

where I is a unit matrix with the same dimensions as A. (Note that $AX - \lambda X = 0$ does **not** simplify to $(A - \lambda)X = 0$ as you **cannot** subtract a scalar λ from a matrix A). This equation (5) is a **homogeneous** system of equations. In the notation of the earlier discussion $C \equiv A - \lambda I$ and $K \equiv 0$. For such a system we know that non-trivial solutions will only exist if the determinant of the coefficient matrix is zero:

$$\det(A - \lambda I) = 0 \qquad (6)$$

Equation (6) is called the **characteristic equation** of the eigenvalue problem. We see that the characteristic equation only involves one unknown λ. The characteristic equation is generally a polynomial in λ, with degree being the same as the order of A (so if A is 2×2 the characteristic equation is a quadratic, if A is a 3×3 it is a cubic equation, and so on). For each value of λ that is obtained the corresponding value of X is obtained by solving the original equations (4). These X's are called **eigenvectors**.

N.B. We shall see that eigenvectors are only unique up to a multiplicative factor: i.e. if X satisfies $AX = \lambda X$ then so does kX when k is any constant.

Example 5

Find the eigenvalues and eigenvectors of the matrix $A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$

Solution

The eigenvalues and eigenvectors are found by solving the eigenvalue probelm

$$AX = \lambda X \qquad X = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{i.e.} \quad (A - \lambda I)X = 0.$$

Non-trivial solutions will exist if det $(A - \lambda I) = 0$

that is, det $\left\{ \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} = 0, \qquad \therefore \qquad \begin{vmatrix} 1-\lambda & 0 \\ 1 & 2-\lambda \end{vmatrix} = 0,$

expanding this determinant: $(1 - \lambda)(2 - \lambda) = 0$. Hence the solutions for λ are: $\lambda = 1$ and $\lambda = 2$.

So we have found two values of λ for this 2×2 matrix A. Since these are unequal they are said to be **distinct** eigenvalues.

To each value of λ there corresponds an eigenvector. We now proceed to find the eigenvectors.

Case 1

$\lambda = 1$ (smaller eigenvalue). Then our original eigenvalue problem becomes: $AX = X$. In full this is

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

Simplifying

$$\begin{aligned} x &= x \qquad \text{(a)} \\ x + y &= 0 \qquad \text{(b)} \end{aligned}$$

All we can deduce here is that $x = -y \qquad \therefore \qquad X = \begin{bmatrix} x \\ -x \end{bmatrix}$ for any $x \neq 0$

(We specify $x \neq 0$ as, otherwise, we would have the trivial solution.)

So the eigenvectors corresponding to eigenvalue $\lambda = 1$ are all proportional to $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, e.g. $\begin{bmatrix} 2 \\ -2 \end{bmatrix}$, $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ etc.

Sometimes we write the eigenvector in **normalised** form that is, with modulus or magnitude 1. Here, the normalised form of X is

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \text{which } \textbf{is unique}.$$

Solution (contd.)

Case 2 Now we consider the larger eigenvalue $\lambda = 2$. Our original eigenvalue problem $AX = \lambda X$ becomes $AX = 2X$ which gives the following equations:

$$\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2 \begin{bmatrix} x \\ y \end{bmatrix}$$

i.e.

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

These equations imply that $x = 0$ whilst the variable y may take any value whatsoever (except zero as this gives the trivial solution).

Thus the eigenvector corresponding to eigenvalue $\lambda = 2$ has the form $\begin{bmatrix} 0 \\ y \end{bmatrix}$, e.g. $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$ etc.

The normalised eigenvector here is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

In conclusion: the matrix $A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$ has two eigenvalues and two associated normalised eigenvectors:

$$\lambda_1 = 1, \qquad \lambda_2 = 2$$

$$X_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad X_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Example 6

Find the eigenvalues and eigenvectors of the 3×3 matrix

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

Solution

The eigenvalues and eigenvectors are found by solving the eigenvalue problem

$$AX = \lambda X \qquad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Proceeding as in Example 5:

$(A - \lambda I)X = 0$ and non-trivial solutions for X will exist if $\quad \det\,(A - \lambda I) = 0$

Solution (contd.)

that is,

$$\det \left\{ \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} = 0$$

$$\text{i.e.} \quad \begin{vmatrix} 2-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 2-\lambda \end{vmatrix} = 0.$$

Expanding this determinant we find:

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

that is,

$$(2-\lambda)\left\{(2-\lambda)^2 - 1\right\} - (2-\lambda) = 0$$

Taking out the common factor $(2-\lambda)$:

$$(2-\lambda)\left\{4 - 4\lambda + \lambda^2 - 1 - 1\right\}$$

which gives: $\quad (2-\lambda)\,[\lambda^2 - 4\lambda + 2] = 0.$

This is easily solved to give: $\lambda = 2$ or $\lambda = \dfrac{4 \pm \sqrt{16-8}}{2} = 2 \pm \sqrt{2}.$

So (typically) we have found three possible values of λ for this 3×3 matrix A.

To each value of λ there corresponds an eigenvector.

Case 1: $\lambda = 2 - \sqrt{2}$ (lowest eigenvalue)

Then $AX = (2-\sqrt{2})X$ implies

$$\begin{aligned} 2x - y &= (2-\sqrt{2})x \\ -x + 2y - z &= (2-\sqrt{2})y \\ -y + 2z &= (2-\sqrt{2})z \end{aligned}$$

Simplifying

$$\begin{aligned} \sqrt{2}x - y &= 0 && \text{(a)} \\ -x + \sqrt{2}y - z &= 0 && \text{(b)} \\ -y + \sqrt{2}z &= 0 && \text{(c)} \end{aligned}$$

We conclude the following:

(c) $\Rightarrow y = \sqrt{2}z \qquad$ (a) $\Rightarrow y = \sqrt{2}x$

$\therefore \quad$ these two relations give $\quad x = z \quad$ then $\quad$ (b) $\Rightarrow -x + 2x - x = 0$

The last equation gives us no information; it simply states that $0 = 0$.

Solution (contd.)

$\therefore \qquad X = \begin{bmatrix} x \\ \sqrt{2}x \\ x \end{bmatrix}$ for any $x \neq 0$ (otherwise we would have the trivial solution). So the eigenvectors corresponding to eigenvalue $\lambda = 2 - \sqrt{2}$ are all proportional to $\begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}$.

In normalised form we have an eigenvector $\dfrac{1}{2}\begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}$.

Case 2: $\lambda = 2$

Here $AX = 2X$ implies $\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

i.e.

$$\begin{aligned} 2x - y &= 2x \\ -x + 2y - z &= 2y \\ -y + 2z &= 2z \end{aligned}$$

After simplifying the equations become:

$$\begin{aligned} -y &= 0 &\text{(a)} \\ -x - z &= 0 &\text{(b)} \\ -y &= 0 &\text{(c)} \end{aligned}$$

(a), (c) imply $y = 0$: (b) implies $x = -z$

$\therefore \qquad$ eigenvector has the form $\begin{bmatrix} x \\ 0 \\ -x \end{bmatrix}$ for any $x \neq 0$.

That is, eigenvectors corresponding to $\lambda = 2$ are all proportional to $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$.

In normalised form we have an eigenvector $\dfrac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$.

Solution (contd.)

Case 3: $\lambda = 2 + \sqrt{2}$ (largest eigenvalue)

Proceeding along similar lines to cases 1,2 above we find that the eigenvectors corresponding to $\lambda = 2 + \sqrt{2}$ are each proportional to $\begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix}$ with normalised eigenvector $\dfrac{1}{2}\begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix}$.

In conclusion the matrix A has three distinct eigenvalues:

$$\lambda_1 = 2 - \sqrt{2}, \qquad \lambda_2 = 2 \qquad \lambda_3 = 2 + \sqrt{2}$$

and three corresponding normalised eigenvectors:

$$X_1 = \frac{1}{2}\begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}, \qquad X_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad X_3 = \frac{1}{2}\begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix}$$

Exercise

Find the eigenvalues and eigenvectors of each of the following matrices A:

(a) $\begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 \\ -8 & 11 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 0 & -2 \\ 0 & 4 & 0 \\ -2 & 0 & 5 \end{bmatrix}$ (d) $\begin{bmatrix} 10 & -2 & 4 \\ -20 & 4 & -10 \\ -30 & 6 & -13 \end{bmatrix}$

Answer (eigenvectors are written in normalised form)

(a) 3 and 2; $\begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$ and $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

(b) 3 and 9; $\dfrac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\dfrac{1}{\sqrt{17}}\begin{bmatrix} 1 \\ 4 \end{bmatrix}$

(c) 1, 4 and 6; $\dfrac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$; $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$; $\dfrac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$

(d) 0, -1 and 2; $\dfrac{1}{\sqrt{26}}\begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}$; $\dfrac{1}{\sqrt{5}}\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$; $\dfrac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$

3. Properties of eigenvalues and eigenvectors

There are a number of general properties of eigenvalues and eigenvectors which you should be familiar with. You will be able to use them as a check on some of your calculations.

Property 1: Sum of eigenvalues

For any square matrix A:

sum of eigenvalues = sum of diagonal terms of A (called the **trace** of A)

Formally, for an $n \times n$ matrix A: $\displaystyle\sum_{i=1}^{n} \lambda_i = \text{trace}(A)$

(Repeated eigenvalues must be counted according to their multiplicity.)

Thus if $\lambda_1 = 4, \lambda_2 = 4, \lambda_3 = 1$ then $\displaystyle\sum_{i=1}^{3} \lambda_i = 9$).

Property 2: Product of eigenvalues

For any square matrix A:

product of eigenvalues = determinant of A

Formally: $\displaystyle\lambda_1\lambda_2\lambda_3\cdots\lambda_n = \prod_{i=1}^{n} \lambda_i = \det(A)$

The symbol $\prod$ simply denotes multiplication, as $\sum$ denotes summation.

Example 7

Verify Properties 1 and 2 for the 3×3 matrix:

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

whose eigenvalues were found earlier.

Solution

The three eigenvalues of this matrix are:

$$\lambda_1 = 2 - \sqrt{2}, \qquad \lambda_2 = 2, \qquad \lambda_3 = 2 + \sqrt{2}$$

Therefore

$$\lambda_1 + \lambda_2 + \lambda_3 = (2 - \sqrt{2}) + 2 + (2 + \sqrt{2}) = 6 = \text{trace}(A)$$

whilst $\lambda_1\lambda_2\lambda_3 = (2 - \sqrt{2})(2)(2 + \sqrt{2}) = 4 = \det(A)$

Property 3: Linear independence of eigenvectors

Eigenvectors of a matrix A corresponding to distinct eigenvalues are **linearly independent** i.e. one eigenvector **cannot** be written as a linear sum of the other eigenvectors. The proof of this result is omitted but we illustrate this property with two examples.

We saw earlier that the matrix

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$$

has distinct eigenvalues $\lambda_1 = 1 \qquad \lambda_2 = 2$ with associated eigenvectors

$$X^{(1)} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad X^{(2)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

respectively.

Clearly $X^{(1)}$ is **not** a constant multiple of $X^{(2)}$ and these eigenvectors are **linearly independent**.

We also saw that the 3×3 matrix

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

had the following distinct eigenvalues $\lambda_1 = 2 - \sqrt{2},\ \lambda_2 = 2,\ \lambda_3 = 2 + \sqrt{2}$ with corresponding eigenvectors of the form shown:

$$X^{(1)} = \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}, \qquad X^{(2)} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad X^{(3)} = \begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix}$$

Clearly none of these eigenvectors is a constant multiple of any other. Nor is any one obtainable as a linear combination of the other two. The three eigenvectors are **linearly independent**.

Property 4: Eigenvalues of diagonal matrices

A 2×2 diagonal matrix D has the form

$$D = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

The characteristic equation

$$|D - \lambda I| = 0 \quad \text{is} \quad \begin{vmatrix} a-\lambda & 0 \\ 0 & d-\lambda \end{vmatrix} = 0$$

i.e. $\qquad (a - \lambda)(d - \lambda) = 0$

So the eigenvalues are simply the diagonal elements a and d.

Similarly a 3×3 diagonal matrix has the form

$$D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

having characteristic equation

$$|D - \lambda I| = (a - \lambda)(b - \lambda)(c - \lambda) = 0$$

so again the diagonal elements **are** the eigenvalues.

We can see that a diagonal matrix is a particularly simple matrix to work with. In addition to the eigenvalues being obtainable immediately by inspection it is exceptionally easy to multiply diagonal matrices.

Task

Obtain the products D_1D_2 and D_2D_1 of the diagonal matrices

$$D_1 = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \qquad D_2 = \begin{bmatrix} e & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & g \end{bmatrix}$$

Your solution

Answer

$$D_1D_2 = D_2D_1 = \begin{bmatrix} ae & 0 & 0 \\ 0 & bf & 0 \\ 0 & 0 & cg \end{bmatrix}$$

which of course is also a diagonal matrix.

Exercise

If $\lambda_1, \lambda_2, \ldots \lambda_n$ are the eigenvalues of a matrix A, prove the following:

(a) A^T has eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n$.

(b) If A is upper triangular, then its eigenvalues are exactly the main diagonal entries.

(c) The inverse matrix A^{-1} has eigenvalues $\dfrac{1}{\lambda_1}, \dfrac{1}{\lambda_2}, \ldots \dfrac{1}{\lambda_n}$.

(d) The matrix $A - kI$ has eigenvalues $\lambda_1 - k, \lambda_2 - k, \ldots \lambda_n - k$.

(e) (Harder) The matrix A^2 has eigenvalues $\lambda_1^2, \lambda_2^2, \ldots \lambda_n^2$.

(f) (Harder) The matrix A^k (k a non-negative integer) has eigenvalues $\lambda_1^k, \lambda_2^k, \ldots \lambda_n^k$.

Verify the above results for any 2×2 matrix and any 3×3 matrix found in the previous Exercises on page 13.

N.B. Some of these results are useful in the numerical calculation of eigenvalues which we shall consider later.

Answer

(a) Using the property that for any square matrix A, $\det(A) = \det(A^T)$ we see that if

$$\det(A - \lambda I) = 0 \qquad \text{then} \qquad \det(A - \lambda I)^T = 0$$

This immediately tells us that $\det(A^T - \lambda I) = 0$ which shows that λ is also an eigenvalue of A^T.

(b) Here simply write down a typical upper triangular matrix U which has terms on the leading diagonal $u_{11}, u_{22}, \ldots, u_{nn}$ and above it. Then construct $(U - \lambda I)$. Finally imagine how you would then obtain $\det(U - \lambda I) = 0$. You should see that the determinant is obtained by multiplying together those terms on the leading diagonal. Here the characteristic equation is:

$$(u_{11} - \lambda)(u_{22} - \lambda) \ldots (u_{nn} - \lambda) = 0$$

This polynomial has the obvious roots $\lambda_1 = u_{11},\ \lambda_2 = u_{22},\ \ldots, \lambda_n = u_{nn}$.

(c) Here we begin with the usual eigenvalue problem $AX = \lambda X$. If A has an inverse A^{-1} we can multiply both sides by A^{-1} on the left to give

$$A^{-1}(AX) = A^{-1}\lambda X \qquad \text{which gives} \qquad X = \lambda A^{-1} X$$

or, dividing through by the scalar λ we get

$A^{-1}X = \dfrac{1}{\lambda}X$ which shows that if λ and X are respectively eigenvalue and eigenvector of A then λ^{-1} and X are respectively eigenvalue and eigenvector of A^{-1}.

As an example consider $A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$. This matrix has eigenvalues $\lambda_1 = -1$, $\lambda_2 = 5$ with corresponding eigenvectors $X_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $X_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The reader should verify (by direct multiplication) that $A^{-1} = -\dfrac{1}{5}\begin{bmatrix} 2 & -3 \\ -3 & 2 \end{bmatrix}$ has eigenvalues -1 and $\dfrac{1}{5}$ with respective eigenvectors $X_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $X_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(d) (e) and (f) are proved in similar way to the proof outlined in (c).

Applications of Eigenvalues and Eigenvectors

22.2

Introduction

Many applications of matrices in both engineering and science utilize eigenvalues and, sometimes, eigenvectors. Control theory, vibration analysis, electric circuits, advanced dynamics and quantum mechanics are just a few of the application areas.

Many of the applications involve the use of eigenvalues and eigenvectors in the process of **transforming** a given matrix into a **diagonal** matrix and we discuss this process in this Section. We then go on to show how this process is invaluable in solving coupled differential equations of both first order and second order.

Prerequisites

Before starting this Section you should ...

- have a knowledge of determinants and matrices
- have a knowledge of linear first order differential equations

Learning Outcomes

On completion you should be able to ...

- diagonalize a matrix with distinct eigenvalues using the modal matrix
- solve systems of linear differential equations by the 'decoupling' method

1. Diagonalization of a matrix with distinct eigenvalues

Diagonalization means transforming a non-diagonal matrix into an equivalent matrix which is diagonal and hence is simpler to deal with.

A matrix A with distinct eigenvalues has, as we mentioned in Property 3 in HELM 22.1, eigenvectors which are linearly independent. If we form a matrix P whose columns are these eigenvectors, it can be shown that

$$\det(P) \neq 0$$

so that P^{-1} exists.

The product $P^{-1}AP$ is then a **diagonal** matrix D whose diagonal elements are the eigenvalues of A. Thus if $\lambda_1, \lambda_2, \dots \lambda_n$ are the distinct eigenvalues of A with associated eigenvectors $X^{(1)}, X^{(2)}, \dots, X^{(n)}$ respectively, then

$$P = \left[\begin{array}{ccccccc} X^{(1)} & \vdots & X^{(2)} & \vdots & \cdots & \vdots & X^{(n)} \\ & & & & & & \end{array}\right]$$

will produce a product

$$P^{-1}AP = D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & & \\ 0 & \dots & \dots & \lambda_n \end{bmatrix}$$

We see that the order of the eigenvalues in D matches the order in which P is formed from the eigenvectors.

N.B.

(a) The matrix P is called the **modal matrix** of A

(b) Since D is a diagonal matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ which are the same as those of A, then the matrices D and A are said to be **similar**.

(c) The transformation of A into D using

$$P^{-1}AP = D$$

is said to be a **similarity transformation.**

Example 8

Let $A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$. Obtain the modal matrix P and calculate the product $P^{-1}AP$.

(The eigenvalues and eigenvectors of this particular matrix A were obtained earlier in this Workbook at page 7.)

Solution

The matrix A has two distinct eigenvalues $\lambda_1 = -1$, $\lambda_2 = 5$ with corresponding eigenvectors $X_1 = \begin{bmatrix} x \\ -x \end{bmatrix}$ and $X_2 = \begin{bmatrix} x \\ x \end{bmatrix}$. We can therefore form the modal matrix from the simplest eigenvectors of these forms:

$$P = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

(Other eigenvectors would be acceptable e.g. we could use $P = \begin{bmatrix} 2 & 3 \\ -2 & 3 \end{bmatrix}$ but there is no reason to over complicate the calculation.)

It is easy to obtain the inverse of this 2×2 matrix P and the reader should confirm that:

$$P^{-1} = \frac{1}{\text{det}(P)}\,\text{adj}(P) = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}^T = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

We can now construct the product $P^{-1}AP$:

$$\begin{aligned} P^{-1}AP &= \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 5 \\ 1 & 5 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} -2 & 0 \\ 0 & 10 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix} \end{aligned}$$

which is a diagonal matrix with entries the eigenvalues, as expected. Show (by repeating the method outlined above) that had we defined $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ (i.e. interchanged the order in which the eigenvectors were taken) we would find $P^{-1}AP = \begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix}$ (i.e. the resulting diagonal elements would also be interchanged.)

Task

The matrix $A = \begin{bmatrix} -1 & 4 \\ 0 & 3 \end{bmatrix}$ has eigenvalues -1 and 3 with respective eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

If $P_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $P_2 = \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix}$, $P_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ write down the products $P_1^{-1}AP_1$, $P_2^{-1}AP_2$, $P_3^{-1}AP_3$

(You may not need to do detailed calculations.)

Your solution

Answer

$P_1^{-1}AP_1 = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} = D_1 \qquad P_2^{-1}AP_2 = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} = D_2 \qquad P_3^{-1}AP_3 = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} = D_3$

Note that $D_1 = D_2$, demonstrating that any eigenvectors of A can be used to form P. Note also that since the columns of P_1 have been interchanged in forming P_3 then so have the eigenvalues in D_3 as compared with D_1.

Matrix powers

If $P^{-1}AP = D$ then we can obtain A (i.e. make A the subject of this matrix equation) as follows:

Multiplying on the left by P and on the right by P^{-1} we obtain

$$PP^{-1}APP^{-1} = PDP^{-1}$$

Now using the fact that $PP^{-1} = P^{-1}P = I$ we obtain

$$IAI = PDP^{-1} \quad \text{and so}$$

$$A = PDP^{-1}$$

We can use this result to obtain the **powers** of a square matrix, a process which is sometimes useful in control theory. Note that

$$A^2 = A.A \qquad A^3 = A.A.A. \quad \text{etc.}$$

Clearly, obtaining high powers of A directly would in general involve many multiplications. The process is quite straightforward, however, for a diagonal matrix D, as this next Task shows.

Task

Obtain D^2 and D^3 if $D = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}$. Write down D^{10}.

Your solution

Answer

$$D^2 = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 3^2 & 0 \\ 0 & (-2)^2 \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$$

$$D^3 = \begin{bmatrix} 3^2 & 0 \\ 0 & (-2)^2 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & (-2) \end{bmatrix} = \begin{bmatrix} 3^3 & 0 \\ 0 & (-2)^3 \end{bmatrix} = \begin{bmatrix} 27 & 0 \\ 0 & -8 \end{bmatrix}$$

Continuing in this way: $D^{10} = \begin{bmatrix} 3^{10} & 0 \\ 0 & (-2)^{10} \end{bmatrix} = \begin{bmatrix} 58049 & 0 \\ 0 & 1024 \end{bmatrix}$

We now use the relation $A = PDP^{-1}$ to obtain a formula for powers of A in terms of the easily calculated powers of the diagonal matrix D:

$$A^2 = A.A = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDIDP^{-1} = PD^2P^{-1}$$

Similarly: $A^3 = A^2.A = (PD^2P^{-1})(PDP^{-1}) = PD^2(P^{-1}P)DP^{-1} = PD^3P^{-1}$

The general result is given in the following Key Point:

Key Point 2

For a matrix A with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and associated eigenvectors $X^{(1)}, X^{(2)}, \ldots, X^{(n)}$ then if

$$P = [X^{(1)} : X^{(2)} : \ \ldots \ : X^{(n)}]$$

$D = P^{-1}AP$ is a diagonal matrix such that

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \qquad \text{and} \qquad A^k = PD^kP^{-1}$$

Example 9

If $A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$ find A^{23}. (Use the results of Example 8.)

Solution

We know from Example 8 that if $P = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ then $P^{-1}AP = \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix} = D$

where $P^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

$\therefore \quad A = PDP^{-1}$ and $A^{23} = PD^{23}P^{-1}$ using the general result in Key Point 2

i.e. $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 5^{23} \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

which is easily evaluated.

Exercise

Find a diagonalizing matrix P if

(a) $A = \begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & -2 & 3 \end{bmatrix}$

Verify, in each case, that $P^{-1}AP$ is diagonal, with the eigenvalues of A as its diagonal elements.

Answer

(a) $P = \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}, \quad PAP^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

(b) $P = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix}, \quad PAP^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

2. Systems of first order differential equations

Systems of first order ordinary differential equations arise in many areas of mathematics and engineering, for example in control theory and in the analysis of electrical circuits. In each case the basic unknowns are each a function of the time variable t. A number of techniques have been developed to solve such systems of equations; for example the Laplace transform. Here we shall use eigenvalues and eigenvectors to obtain the solution. Our first step will be to recast the system of ordinary differential equations in the **matrix form** $\dot{X} = AX$ where A is an $n \times n$ coefficient matrix of constants, X is the $n \times 1$ column vector of unknown functions and $\dot{X}$ is the $n \times 1$ column vector containing the **derivatives** of the unknowns.. The main step will be to use the modal matrix of A to diagonalise the system of differential equations. This process will transform $\dot{X} = AX$ into the form $\dot{Y} = DY$ where D is a **diagonal matrix**. We shall find that this new diagonal system of differential equations can be easily solved. This special solution will allow us to obtain the solution of the original system.

Task

Obtain the solutions of the pair of first order differential equations

$$\left.\begin{array}{l} \dot{x} = -2x \\ \dot{y} = -5y \end{array}\right\} \qquad (1)$$

given the **initial conditions**

$$\begin{array}{lll} x(0) = 3 & \text{i.e. } x = 3 & \text{at } t = 0 \\ y(0) = 2 & \text{i.e. } y = 2 & \text{at } t = 0 \end{array}$$

(The notation is that $\dot{x} \equiv \dfrac{dx}{dt}$, $\dot{y} \equiv \dfrac{dy}{dt}$)

[Hint: Recall, from your study of differential equations, that the general solution of the differential equation $\dfrac{dy}{dt} = Ky$ is $y = y_0 e^{Kt}$.]

Your solution

Answer

Using the hint: $x = x_0 e^{-2t}$ $y = y_0 e^{-5t}$ where $x_0 = x(0)$ and $y_0 = y(0)$.

From the given initial condition $x_0 = 3$ $y_0 = 2$ so finally $x = 3e^{-2t}$ $y = 2e^{-5t}$.

In the above Task although we had two differential equations to solve they were really quite separate. We needed no knowledge of matrix theory to solve them. However, we should note that the two differential equations can be written in matrix form.

Thus if $X = \begin{bmatrix} x \\ y \end{bmatrix} \qquad \dot{X} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \quad A = \begin{bmatrix} -2 & 0 \\ 0 & -5 \end{bmatrix}$

the two equations (1) can be written as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

i.e. $\dot{X} = AX$.

Task

Write in matrix form the pair of **coupled** differential equations

$$\left.\begin{array}{l} \dot{x} = 4x + 2y \\ \dot{y} = -x + y \end{array}\right\} \qquad (2)$$

Your solution

Answer

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\dot{X} \quad = \quad A\,X$$

The essential difference between the two pairs of differential equations just considered is that the pair (1) were really separate equations whereas pair (2) were coupled:

- The first equation of (1) involving only the unknown x, the second involving only y. In matrix terms this corresponded to a **diagonal** matrix A in the system $\dot{X} = AX$.
- The pair (2) were coupled in that **both** equations involved **both** x and y. This corresponded to the **non-diagonal** matrix A in the system $\dot{X} = AX$ which you found in the last Task.

Clearly the second system here is more difficult to deal with than the first and this is where we can use our knowledge of diagonalization.

Consider a system of differential equations written in matrix form: $\dot{X} = AX$ where

$$X = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \qquad \text{and} \qquad \dot{X} = \begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \end{bmatrix}$$

We now introduce a new column vector of unknowns $Y = \begin{bmatrix} r(t) \\ s(t) \end{bmatrix}$ through the relation

$$X = PY$$

where P is the modal matrix of A. Then, since P is a matrix of constants:

$$\dot{X} = P\dot{Y} \qquad \text{so} \quad \dot{X} = AX \quad \text{becomes} \quad P\dot{Y} = A(PY)$$

Then, multiplying by P^{-1} on the left, $\dot{Y} = (P^{-1}AP)Y$

But, because of the properties of the modal matrix, we know that $P^{-1}AP$ is a **diagonal matrix**. Thus if λ_1, λ_2 are **distinct** eigenvalues of A then:

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Hence $\dot{Y} = (P^{-1}AP)Y$ becomes

$$\begin{bmatrix} \dot{r} \\ \dot{s} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} r \\ s \end{bmatrix}.$$

That is, when written out we have

$$\begin{aligned} \dot{r} &= \lambda_1 r \\ \dot{s} &= \lambda_2 s. \end{aligned}$$

These equations are **decoupled**. The first equation only involves the unknown function $r(t)$ and has solution $r(t) = Ce^{\lambda_1 t}$. The second equation only involves the unknown function $s(t)$ and has solution $s(t) = Ke^{\lambda_2 t}$. [C, K are arbitrary constants.]

Once r, s are known the original unknowns x, y can be found from the relation $X = PY$.

Note that the theory outlined above is more widely applicable as specified in the next Key Point:

Key Point 3

For any system of differential equations of the form

$$\dot{X} = AX$$

where A is an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, and t is the independent variable the solution is

$$X = PY$$

where P is the modal matrix of A and

$$Y = [C_1 \mathrm{e}^{\lambda_1 t},\ C_2 \mathrm{e}^{\lambda_2 t}, \ldots, C_n \mathrm{e}^{\lambda_n t}]^T$$

Example 10

Find the solution of the coupled differential equations

$$\dot{x} = 4x + 2y$$

$$\dot{y} = -x + y \qquad \text{with initial conditions} \qquad x(0) = 1 \qquad y(0) = 0$$

Here $\dot{x} \equiv \dfrac{dx}{dt}$ and $\dot{y} \equiv \dfrac{dy}{dt}$.

Solution

Here $A = \begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix}$. It is easily checked that A has distinct eigenvalues $\lambda_1 = 3$ $\lambda_2 = 2$ and corresponding eigenvectors $X_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, $X_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Therefore, taking $P = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}$ then $P^{-1}AP = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$

and using Key Point 3, $r(t) = Ce^{3t} \qquad s(t) = Ke^{2t}$.

So
$$\begin{bmatrix} x \\ y \end{bmatrix} \equiv X = PY = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} Ce^{3t} \\ Ke^{2t} \end{bmatrix} = \begin{bmatrix} -2Ce^{3t} + Ke^{2t} \\ Ce^{3t} - Ke^{2t} \end{bmatrix}.$$

Therefore $x = -2Ce^{3t} + Ke^{2t}$ and $y = Ce^{3t} - Ke^{2t}$.

We can now impose the initial conditions $x(0) = 1$ and $y(0) = 0$ to give

$$\begin{aligned} 1 &= -2C + K \\ 0 &= C - K. \end{aligned}$$

Thus $C = K = -1$ and the solution to the original system of differential equations is

$$\begin{aligned} x(t) &= 2e^{3t} - e^{2t} \\ y(t) &= -e^{3t} + e^{2t} \end{aligned}$$

The approach we have demonstrated in Example 10 can be extended to

(a) Systems of first order differential equations with n unknowns (Key Point 3)

(b) Systems of second order differential equations (described in the next subsection).

The only restriction, as we have said, is that the matrix A in the system $\dot{X} = AX$ has distinct eigenvalues.

3. Systems of second order differential equations

The decoupling method discussed above can be readily extended to this situation which could arise, for example, in a mechanical system consisting of coupled springs.
A typical example of such a system with two unknowns has the form

$$\ddot{x} = ax + by \qquad\qquad \ddot{y} = cx + dy$$

or, in matrix form,

$$\ddot{X} = AX \quad \text{where} \quad X = \begin{bmatrix} x \\ y \end{bmatrix} \qquad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad \ddot{x} = \frac{d^2x}{dt^2}, \quad \ddot{y} = \frac{d^2y}{dt^2}$$

Task

Make the substitution $X = PY$ where $Y = \begin{bmatrix} r(t) \\ s(t) \end{bmatrix}$ and P is the modal matrix of A, A being assumed here to have distinct eigenvalues λ_1 and λ_2. Solve the resulting pair of decoupled equations for the case, which arises in practice, where λ_1 and λ_2 are both negative.

Your solution

Answer

Exactly as with a first order system, putting $X = PY$ into the second order system $\ddot{X} = AX$ gives

$$\ddot{Y} = P^{-1}APY \quad \text{that is} \quad \ddot{Y} = DY \quad \text{where} \quad D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \text{ and } \ddot{Y} = \begin{bmatrix} \ddot{r} \\ \ddot{s} \end{bmatrix} \quad \text{so}$$

$$\begin{bmatrix} \ddot{r} \\ \ddot{s} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} r \\ s \end{bmatrix}$$

That is, $\ddot{r} = \lambda_1 r = -\omega_1^2 r$ and $\ddot{s} = \lambda_2 s = -\omega_2^2 s$ (where λ_1 and λ_2 are both negative.)

The two decoupled equations are of the form of the differential equation governing simple harmonic motion. Hence the general solution is

$$r = K\cos\omega_1 t + L\sin\omega_1 t \quad \text{and} \quad s = M\cos\omega_2 t + N\sin\omega_2 t$$

The solutions for x and y are then obtained by use of $X = PY$.

Note that in this second order case four initial conditions, two each for both x and y, are required because four constants K, L, M, N arise in the solution.

Exercises

1. Solve by decoupling each of the following **first order** systems:

(a) $\dfrac{dX}{dt} = AX$ where $A = \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}, \quad X(0) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

(b) $\dot{x}_1 = x_2 \qquad \dot{x}_2 = x_1 + 3x_3 \qquad \dot{x}_3 = x_2$ with $x_1(0) = 2, \quad x_2(0) = 0, \quad x_3(0) = 2$

(c) $\dfrac{dX}{dt} = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix} X$, with $X(0) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

(d) $\dot{x}_1 = x_1 \qquad \dot{x}_2 = -2x_2 + x_3 \qquad \dot{x}_3 = 4x_2 + x_3$ with $x_1(0) = x_2(0) = x_3(0) = 1$

2. Matrix methods can be used to solve systems of **second order** differential equations such as might arise with coupled electrical or mechanical systems. For example the motion of two masses m_1 and m_2 vibrating on coupled springs, neglecting damping and spring masses, is governed by

$$\begin{aligned} m_1\ddot{y}_1 &= -k_1y_1 + k_2(y_2 - y_1) \\ m_2\ddot{y}_2 &= -k_2(y_2 - y_1) \end{aligned}$$

where dots denote derivatives with respect to time.

Write this system as a matrix equation $\ddot{Y} = AY$ and use the decoupling method to find Y if

(a) $m_1 = m_2 = 1, \; k_1 = 3, \; k_2 = 2$

and the initial conditions are $y_1(0) = 1$, $y_2(0) = 2$, $\dot{y}(0) = -2\sqrt{6}$, $\dot{y}_2(0) = \sqrt{6}$

(b) $m_1 = m_2 = 1, \quad k_1 = 6, \; k_2 = 4$

and the initial conditions are $y_1(0) = y_2(0) = 0, \quad \dot{y}_1(0) = \sqrt{2}, \quad \dot{y}_2(0) = 2\sqrt{2}$

Verify your solutions by substitution in each case.

Answers

1. (a) $X = \begin{bmatrix} 2e^{5t} & - & e^{-5t} \\ e^{5t} & + & 2e^{-5t} \end{bmatrix}$ (b) $X = \begin{bmatrix} 2\ \cosh 2t \\ 4\ \sinh 2t \\ 2\ \cosh 2t \end{bmatrix}$

(c) $X = \dfrac{1}{4}\begin{bmatrix} e^{5t} & + & 3e^{t} \\ e^{5t} & - & e^{t} \\ e^{5t} & - & e^{t} \end{bmatrix}$ (d) $X = \dfrac{1}{5}\begin{bmatrix} 5e^{t} & & \\ 2e^{2t} & + & 3e^{-3t} \\ 8e^{2t} & - & 3e^{-3t} \end{bmatrix}$

2. (a) $Y = \begin{bmatrix} \cos t - 2\sin\sqrt{6}t \\ 2\cos t + \sin\sqrt{6}t \end{bmatrix}$ (b) $Y = \begin{bmatrix} \sin\sqrt{2}t \\ 2\sin\sqrt{2}t \end{bmatrix}$

Repeated Eigenvalues and Symmetric Matrices

22.3

Introduction

In this Section we further develop the theory of eigenvalues and eigenvectors in two distinct directions. Firstly we look at matrices where one or more of the eigenvalues is **repeated**. We shall see that this sometimes (but not always) causes problems in the diagonalization process that was discussed in the previous Section. We shall then consider the special properties possessed by **symmetric** matrices which make them particularly easy to work with.

Prerequisites

Before starting this Section you should . . .

- have a knowledge of determinants and matrices
- have a knowledge of linear first order differential equations

Learning Outcomes

On completion you should be able to . . .

- state the conditions under which a matrix with repeated eigenvalues may be diagonalized
- state the main properties of real symmetric matrices

1. Matrices with repeated eigenvalues

So far we have considered the diagonalization of matrices with distinct (i.e. non-repeated) eigenvalues. We have accomplished this by the use of a **non-singular** modal matrix P (i.e. one where $\det P \neq 0$ and hence the inverse P^{-1} exists). We now want to discuss briefly the case of a matrix A with at least one pair of **repeated eigenvalues**. We shall see that for some such matrices diagonalization is possible but for others it is not.

The crucial question is whether we can form a non-singular modal matrix P with the eigenvectors of A as its columns.

Example

Consider the matrix

$$A = \begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix}$$

which has characteristic equation

$$\det(A - \lambda I) = (1-\lambda)(1-\lambda) = 0.$$

So the only eigenvalue is 1 which is repeated or, more formally, has **multiplicity** 2.
To obtain eigenvectors of A corresponding to $\lambda = 1$ we proceed as usual and solve

$$AX = 1X$$

or

$$\begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

implying

$$x = x \qquad \text{and} \qquad -4x + y = y$$

from which $x = 0$ and y is arbitrary.
Thus possible eigenvectors are

$$\begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 3 \end{bmatrix} \quad \ldots$$

However, if we attempt to form a modal matrix P from any two of these eigenvectors,

e.g. $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ then the resulting matrix $P = \begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix}$ has **zero** determinant.

Thus P^{-1} **does not exist** and the similarity transformation $P^{-1}AP$ that we have used previously to diagonalize a matrix is not possible here.

The essential point, at a slightly deeper level, is that the columns of P in this case are **not linearly independent** since

$$\begin{bmatrix} 0 \\ -1 \end{bmatrix} = (-1)\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

i.e. one is a multiple of the other.
This situation is to be contrasted with that of a matrix with non-repeated eigenvalues.

Earlier, for example, we showed that the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$$

has the non-repeated eigenvalues $\lambda_1 = -1, \ \lambda_2 = 5$ with associated eigenvectors

$$X_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad X_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

These two eigenvectors **are linearly independent**.

since $\begin{bmatrix} 1 \\ -1 \end{bmatrix} \neq k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ for any value of $k \neq 0$.

Here the modal matrix

$$P = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

has linearly independent columns: so that $\det P \neq 0$ and P^{-1} exists.

The general result, illustrated by this example, is given in the following Key Point.

Key Point 4

Eigenvectors corresponding to distinct eigenvalues are always linearly independent.

It follows from this that we can **always** diagonalize an $n \times n$ matrix with n **distinct** eigenvalues since it will possess n linearly independent eigenvectors. We can then use these as the columns of P, secure in the knowledge that these columns will be linearly independent and hence P^{-1} will exist. It follows, in considering the case of repeated eigenvalues, that the key problem is whether or not there are still n linearly independent eigenvectors for an $n \times n$ matrix.

We shall now consider two 3×3 cases as illustrations.

Task

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

(a) Obtain the eigenvalues and eigenvectors of A.

(b) Can three linearly independent eigenvectors for A be obtained?

(c) Can A be diagonalized?

Your solution

Answer

(a) The characteristic equation of A is $\det(A-\lambda I) = \begin{vmatrix} -2-\lambda & 0 & 1 \\ 1 & 1-\lambda & 0 \\ 0 & 0 & -2-\lambda \end{vmatrix} = 0$

i.e. $(-2-\lambda)(1-\lambda)(-2-\lambda) = 0$ which gives $\lambda = 1,\ \lambda = -2,\ \lambda = -2.$

For $\lambda = 1$ the associated eigenvectors satisfy $\begin{bmatrix} -2 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ from which $x = 0, z = 0$ and y is arbitrary. Thus an eigenvector is $X = \begin{bmatrix} 0 \\ \alpha \\ 0 \end{bmatrix}$ where α is arbitrary, $\alpha \neq 0.$

For the repeated eigenvalue $\lambda = -2$ we must solve $AY = (-2)Y$ for the eigenvector Y:

$\begin{bmatrix} -2 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2x \\ -2y \\ -2z \end{bmatrix}$ from which $z = 0,\ x + 3y = 0$ so the eigenvectors are of the form $Y = \begin{bmatrix} -3\beta \\ \beta \\ 0 \end{bmatrix} = \beta\begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}$ where $\beta \neq 0$ is arbitrary.

(b) X and Y are certainly linearly independent (as we would expect since they correspond to distinct eigenvalues.) However, there is only one independent eigenvector of the form Y corresponding to the repeated eigenvalue -2.

(c) The conclusion is that since A is 3×3 and we can only obtain **two** linearly independent eigenvectors then A **cannot be diagonalized**.

Task

The matrix $A = \begin{bmatrix} 5 & -4 & 4 \\ 12 & -11 & 12 \\ 4 & -4 & 5 \end{bmatrix}$ has eigenvalues $-3,\ 1,\ 1$. The eigenvector corresponding to the eigenvalue -3 is $X = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ or any multiple.

Investigate carefully the eigenvectors associated with the repeated eigenvalue $\lambda = 1$ and deduce whether A can be diagonalized.

Your solution

Answer

We must solve $AY = (1)Y$ for the required eigenvector

i.e. $$\begin{bmatrix} 5 & -4 & 4 \\ 12 & -11 & 12 \\ 4 & -4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Each equation here gives on simplification $x - y + z = 0$. So we have just one equation in three unknowns so we can choose any **two** values arbitrarily. The choices $x = 1,\ \ y = 0$ (and hence $z = -1$) and $x = 0,\ \ y = 1$ (and hence $z = 1$) for example, give rise to linearly independent eigenvectors $Y_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ $Y_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

We can thus form a non-singular modal matrix P from Y_1 and Y_2 together with X (given)

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 3 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

We can then indeed diagonalize A through the transformation

$$P^{-1}AP = D = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Key Point 5

An $n \times n$ matrix with repeated eigenvalues can be diagonalized provided we can obtain n linearly independent eigenvectors for it. This will be the case if, for each repeated eigenvalue λ_i of multiplicity $m_i > 1$, we can obtain m_i linearly independent eigenvectors.

2. Symmetric matrices

Symmetric matrices have a number of useful properties which we will investigate in this Section.

Task

Consider the following four matrices

$$A_1 = \begin{bmatrix} 3 & 1 \\ 4 & 5 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 3 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 5 & 8 & 7 \\ -1 & 6 & 8 \\ 3 & 4 & 0 \end{bmatrix} \qquad A_4 = \begin{bmatrix} 5 & 8 & 7 \\ 8 & 6 & 4 \\ 7 & 4 & 0 \end{bmatrix}$$

What property do the matrices A_2 and A_4 possess that A_1 and A_3 do not?

Your solution

Answer

Matrices A_2 and A_4 are **symmetric** across the principal diagonal. In other words transposing these matrices, i.e. interchanging their rows and columns, does not change them.

$$A_2^T = A_2 \qquad A_4^T = A_4.$$

This property does not hold for matrices A_1 and A_3 which are **non-symmetric**.

Calculating the eigenvalues of an $n \times n$ matrix with real elements involves, in principle at least, solving an n th order polynomial equation, a quadratic equation if $n = 2$, a cubic equation if $n = 3$, and so on. As is well known, such equations sometimes have only real solutions, but complex solutions (occurring as complex conjugate pairs) can also arise. This situation can therefore arise with the eigenvalues of matrices.

Task

Consider the non-symmetric matrix

$$A = \begin{bmatrix} 2 & -1 \\ 5 & -2 \end{bmatrix}$$

Obtain the eigenvalues of A and show that they form a complex conjugate pair.

Your solution

Answer

The characteristic equation of A is

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

i.e.

$$-(2-\lambda)(2+\lambda) + 5 = 0 \qquad \text{leading to} \qquad \lambda^2 + 1 = 0$$

giving eigenvalues $\pm$ i which are of course complex conjugates.

In particular any 2×2 matrix of the form

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

has complex conjugate eigenvalues $a \pm \mathrm{i}b$.

A 3×3 example of a matrix with some complex eigenvalues is

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

A straightforward calculation shows that the eigenvalues of B are

$\lambda = -1$ (real), $\lambda = \pm\mathrm{i}$ (complex conjugates).

With **symmetric** matrices on the other hand, complex eigenvalues are not possible.

Key Point 6

The eigenvalues of a symmetric matrix with real elements are always real.

The general proof of this result in Key Point 6 is beyond our scope but a simple proof for symmetric 2×2 matrices is straightforward.

Let $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ be any 2×2 symmetric matrix, $a,\ b,\ c$ being real numbers.
The characteristic equation for A is

$$(a - \lambda)(c - \lambda) - b^2 = 0 \qquad \text{or, expanding:} \qquad \lambda^2 - (a + c)\lambda + ac - b^2 = 0$$

from which

$$\lambda = \frac{(a + c) \pm \sqrt{(a + c)^2 - 4ac + 4b^2}}{2}$$

The quantity under the square root sign can be treated as follows:

$$(a + c)^2 - 4ac + 4b^2 = a^2 + c^2 + 2ac - 4ac + b^2 = (a - c)^2 + 4b^2$$

which is always positive and hence λ **cannot be complex**.

Task

Obtain the eigenvalues and the eigenvectors of the symmetric 2×2 matrix

$$A = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}$$

Your solution

Answer

The characteristic equation for A is

$$(4-\lambda)(1-\lambda)+4=0 \qquad \text{or} \qquad \lambda^2-5\lambda=0$$

giving $\lambda=0$ and $\lambda=5$, both of which are of course real and also unequal (i.e. distinct). For the larger eigenvalue $\lambda=5$ the eigenvector $X=\begin{bmatrix} x \\ y \end{bmatrix}$ satisfy

$$\begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} 5x \\ 5y \end{bmatrix} \quad \text{i.e.} \quad -x-2y=0, \qquad -2x-4y=0$$

Both equations tell us that $x=-2y$ so an eigenvector for $\lambda=5$ is $X=\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ or any multiple of this. For $\lambda=0$ the associated eigenvectors satisfy

$$4x-2y=0 \qquad -2x+y=0$$

i.e. $y=2x$ (from both equations) so an eigenvector is $Y=\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ or any multiple.

We now look more closely at the eigenvectors X and Y in the last task. In particular we consider the product X^TY.

Task

Evaluate X^TY from the previous task i.e. where

$$X=\begin{bmatrix} 2 \\ -1 \end{bmatrix} \qquad Y=\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Your solution

Answer

$$X^TY=[2, \quad -1]\begin{bmatrix} 1 \\ 2 \end{bmatrix}=2\times 1-1\times 2=2-2=0$$

$X^TY=0$ means are X and Y are **orthogonal**.

Key Point 7

Two $n\times 1$ column vectors X and Y are **orthogonal** if $X^TY=0$.

Task

We obtained earlier in Section 22.1 Example 6 the eigenvalues of the matrix

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

which, as we now emphasize, is **symmetric**. We found that the eigenvalues were $2, \quad 2+\sqrt{2}, \quad 2-\sqrt{2}$ which are real and distinct. The corresponding eigenvectors were, respectively

$$X = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \qquad Y = \begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix} \qquad Z = \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}$$

(or, as usual, any multiple of these).

Show that these three eigenvectors X, Y, Z are **mutually orthogonal**.

Your solution

Answer

$$X^TY = [1, \quad 0, \quad -1] \begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix} = 1 - 1 = 0$$

$$Y^TZ = [1, \quad -\sqrt{2}, \quad 1] \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix} = 1 - 2 + 1 = 0$$

$$Z^TX = [1, \quad \sqrt{2}, \quad 1] \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = 1 - 1 = 0$$

verifying the mutual orthogonality of these three eigenvectors.

General theory

The following proof that eigenvectors corresponding to distinct eigenvalues of a symmetric matrix are orthogonal is straightforward and you are encouraged to follow it through.

Let A be a symmetric $n \times n$ matrix and let λ_1, λ_2 be two distinct eigenvalues of A i.e. $\lambda_1 \neq \lambda_2$ with associated eigenvectors $X,\ Y$ respectively. We have seen that λ_1 and λ_2 must be real since A is symmetric. Then

$$AX = \lambda_1 X \qquad AY = \lambda_2 Y \tag{1}$$

Transposing the first of there results gives

$$X^T A^T = \lambda_1 X^T \tag{2}$$

(Remember that for any two matrices the transpose of a product is the product of the transposes **in reverse order**.)

We now multiply both sides of (2) on the right by Y (as well as putting $A^T = A$, since A is symmetric) to give:

$$X^T AY = \lambda_1 X^T Y \tag{3}$$

But, using the second eigenvalue equation of (1), equation (3) becomes

$$X^T \lambda_2 Y = \lambda_1 X^T Y$$

or, since λ_2 is just a number,

$$\lambda_2 X^T Y = \lambda_1 X^T Y$$

Taking all terms to the same side and factorising gives

$$(\lambda_2 - \lambda_1) X^T Y = 0$$

from which, since by assumption $\lambda_1 \neq \lambda_2$, we obtain the result

$$X^T Y = 0$$

and the orthogonality has been proved.

Key Point 8

The eigenvectors associated with distinct eigenvalues of a symmetric matrix are **mutually orthogonal**.

The reader familiar with the algebra of vectors will recall that for two vectors whose Cartesian forms are

$$\underline{a} = a_x \underline{i} + a_y \underline{j} + a_z \underline{k} \qquad \underline{b} = b_x \underline{i} + b_y \underline{j} + b_z \underline{k}$$

the scalar (or dot) product is

$$\underline{a} \cdot \underline{b} = a_x b_x + a_y b_y + a_z b_z.$$

Furthermore, if $\underline{a}$ and $\underline{b}$ are mutually perpendicular then $\underline{a} \cdot \underline{b} = 0$. (The word 'orthogonal' is sometimes used instead of perpendicular in the case.) Our result, that two column vectors are orthogonal if $X^T Y = 0$, may thus be considered as a generalisation of the 3-dimensional result $\underline{a} \cdot \underline{b} = 0$.

Diagonalization of symmetric matrices

Recall from our earlier work that

1. We can **always** diagonalize a matrix with distinct eigenvalues (whether these are real or complex).
2. We can **sometimes** diagonalize a matrix with repeated eigenvalues. (The condition for this to be possible is that any eigenvalue of multiplicity m had to have associated with it m linearly independent eigenvectors.)

The situation with symmetric matrices is simpler. Basically we can diagonalize **any** symmetric matrix. To take the discussions further we first need the concept of an **orthogonal** matrix.

A square matrix A is said to be orthogonal if its inverse (if it exists) is equal to its transpose:

$$A^{-1} = A^T \qquad \text{or, equivalently,} \qquad AA^T = A^T A = I.$$

Example

An important example of an orthogonal matrix is

$$A = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}$$

which arises when we use matrices to describe rotations in a plane.

$$\begin{aligned} AA^T &= \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\phi + \sin^2\phi & 0 \\ 0 & \sin^2\phi + \cos^2\phi \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned}$$

It is clear that $A^T A = I$ also, so A is indeed orthogonal.

It can be shown, but we omit the details, that any 2×2 matrix which is orthogonal can be written in one of the two forms.

$$\begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \qquad \text{or} \qquad \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

If we look closely at either of these matrices we can see that

1. The two columns are mutually orthogonal e.g. for the first matrix we have
$$(\cos\phi \quad -\sin\phi) \begin{bmatrix} \sin\phi \\ \cos\phi \end{bmatrix} = \cos\phi\sin\phi - \sin\phi\cos\phi = 0$$
2. Each column has magnitude 1 (because $\sqrt{\cos^2\phi + \sin^2\phi} = 1$)

Although we shall not prove it, these results are necessary and sufficient for any order square matrix to be orthogonal.

Key Point 9

A square matrix A is said to be orthogonal if its inverse (if it exists) is equal to its transpose:

$A^{-1} = A^T$ or, equivalently, $AA^T = A^TA = I$.

A square matrix is orthogonal if and only if its columns are mutually orthogonal and each column has unit magnitude.

Task

For each of the following matrices verify that the two properties above are satisfied. Then check in both cases that $AA^T = A^TA = I$ i.e. that $A^T = A^{-1}$.

(a) $A = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$ (b) $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$

Your solution

Answer

(a) Since $\left(\frac{\sqrt{3}}{2} \quad \frac{1}{2}\right)\begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix} = -\frac{\sqrt{3}}{4} + \frac{\sqrt{3}}{4} = 0$ the columns are orthogonal.

Since $\left|\frac{\sqrt{3}}{2} + \frac{1}{2}\right| = \sqrt{\frac{3}{4} + \frac{1}{4}} = 1$ and $\left|-\frac{1}{2} + \frac{\sqrt{3}}{4}\right| = \sqrt{\frac{1}{4} + \frac{3}{4}} = 1$ each column has unit magnitude.

Straightforward multiplication shows $AA^T = A^TA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$.

(b) Proceed as in (a).

The following is the key result of this Section.

Key Point 10

Any symmetric matrix A can be diagonalized using an orthogonal modal matrix P via the transformation

$$P^TAP = D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

It follows that any $n \times n$ symmetric matrix **must** possess n mutually orthogonal eigenvectors **even if some of the eigenvalues are repeated**.

It should be clear to the reader that Key Point 10 is a very powerful result for any applications that involve diagonalization of a symmetric matrix. Further, if we do need to find the inverse of P, then this is a trivial process since $P^{-1} = P^T$ (Key Point 9).

Task

The symmetric matrix

$$A = \begin{bmatrix} 1 & 0 & \sqrt{2} \\ 0 & 2 & 0 \\ \sqrt{2} & 0 & 0 \end{bmatrix}$$

has eigenvalues $2,\ 2,\ -1$ (i.e. eigenvalue 2 is repeated with multiplicity 2.)

Associated with the non-repeated eigenvalue -1 is an eigenvector:

$$X = \begin{bmatrix} 1 \\ 0 \\ -\sqrt{2} \end{bmatrix} \qquad \text{(or any multiple)}$$

(a) Normalize the eigenvector X:

Your solution

Answer

(a) Normalizing X which has magnitude $\sqrt{1^2 + (-\sqrt{2})^2} = \sqrt{3}$ gives

$$1/\sqrt{3}\begin{bmatrix} 1 \\ 0 \\ -\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ \\ 0 \\ \\ -\sqrt{2/3} \end{bmatrix}$$

(b) Investigate the eigenvectors associated with the repeated eigenvalue 2:

Your solution

Answer

(b) The eigenvectors associated with $\lambda = 2$ satisfy $AY = 2Y$

which gives $\begin{bmatrix} -1 & 0 & \sqrt{2} \\ 0 & 0 & 0 \\ \sqrt{2} & 0 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

The first and third equations give

$$-x + \sqrt{2}z = 0$$

$$\sqrt{2}x - 2z = 0 \qquad \text{i.e. } x = \sqrt{2}z$$

The equations give us no information about y so its value is arbitrary.

Thus Y has the form $Y = \begin{bmatrix} \sqrt{2}\beta \\ \alpha \\ \beta \end{bmatrix}$ where both α and β are arbitrary.

A certain amount of care is now required in the choice of α and β if we are to find an orthogonal modal matrix to diagonalize A.

For any choice

$$X^TY = (1 \quad 0 \quad -\sqrt{2})\begin{bmatrix} \sqrt{2}\beta \\ \alpha \\ \beta \end{bmatrix} = \sqrt{2}\beta - \sqrt{2}\beta = 0.$$

So X and Y are orthogonal. (The normalization of X does not affect this.)

However, we also need two **orthogonal** eigenvectors of the form $\begin{bmatrix} \sqrt{2}\beta \\ \alpha \\ \beta \end{bmatrix}$. Two such are

$$Y^{(1)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad (\text{ choosing } \beta = 0,\ \alpha = 1) \quad \text{and} \quad Y^{(2)} = \begin{bmatrix} \sqrt{2} \\ 0 \\ 1 \end{bmatrix} \quad (\text{ choosing } \alpha = 0,\ \beta = 1)$$

After normalization, these become $\qquad Y^{(1)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad Y^{(2)} = \begin{bmatrix} \sqrt{2/3} \\ 0 \\ 1/\sqrt{3} \end{bmatrix}$

Hence the matrix $\qquad P = \left[X \ \vdots\ Y^{(1)} \ \vdots\ Y^{(2)}\right] = \begin{bmatrix} 1/\sqrt{3} & 0 & \sqrt{2/3} \\ 0 & 1 & 0 \\ -\sqrt{2/3} & 0 & 1/\sqrt{3} \end{bmatrix}$

is orthogonal and diagonalizes A:

$$P^TAP = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Hermitian matrices

In some applications, of which quantum mechanics is one, matrices with **complex** elements arise.

If A is such a matrix then the matrix $\overline{A}^T$ is the **conjugate transpose** of A, i.e. the complex conjugate of each element of A is taken as well as A being transposed. Thus if

$$A = \begin{bmatrix} 2+\text{i} & 2 \\ 3\text{i} & 5-2\text{i} \end{bmatrix} \quad \text{then} \quad \overline{A}^T = \begin{bmatrix} 2-\text{i} & -3\text{i} \\ 2 & 5+2\text{i} \end{bmatrix}$$

An **Hermitian** matrix is one satisfying

$$\overline{A}^T = A$$

The matrix A above is clearly non-Hermitian. Indeed the most obvious features of an Hermitian matrix is that its diagonal elements **must** be real. (Can you see why?) Thus

$$A = \begin{bmatrix} 6 & 4+\text{i} \\ 4-\text{i} & -2 \end{bmatrix}$$

is Hermitian.

A 3×3 example of an Hermitian matrix is

$$A = \begin{bmatrix} 1 & \text{i} & 5-2\text{i} \\ -\text{i} & 3 & 0 \\ 5+2\text{i} & 0 & 2 \end{bmatrix}$$

An Hermitian matrix is in fact a generalization of a symmetric matrix. The key property of an Hermitian matrix is the same as that of a real symmetric matrix – i.e. the eigenvalues are always **real**.

Numerical Determination of Eigenvalues and Eigenvectors

22.4

Introduction

In Section 22.1 it was shown how to obtain eigenvalues and eigenvectors for low order matrices, 2×2 and 3×3. This involved firstly solving the characteristic equation $\det(A - \lambda I) = 0$ for a given $n \times n$ matrix A. This is an n^{th} order polynomial equation and, even for n as low as 3, solving it is not always straightforward. For large n even obtaining the characteristic equation may be difficult, let alone solving it.

Consequently in this Section we give a brief introduction to alternative methods, essentially **numerical** in nature, of obtaining eigenvalues and perhaps eigenvectors.

We would emphasize that in some applications such as Control Theory we might only require one eigenvalue of a matrix A, usually the one largest in magnitude which is called the **dominant** eigenvalue. It is this eigenvalue which sometimes tells us how a control system will behave.

Prerequisites

Before starting this Section you should ...

- have a knowledge of determinants and matrices
- have a knowledge of linear first order differential equations

Learning Outcomes

On completion you should be able to ...

- use the power method to obtain the dominant eigenvalue (and associated eigenvector) of a matrix
- state the main advantages and disadvantages of the power method

1. Numerical determination of eigenvalues and eigenvectors

Preliminaries

Before discussing **numerical** methods of calculating eigenvalues and eigenvectors we remind you of three results for a matrix A with an eigenvalue λ and associated eigenvector X.

- A^{-1} (if it exists) has an eigenvalue $\dfrac{1}{\lambda}$ with associated eigenvector X.
- The matrix $(A - kI)$ has an eigenvalue $(\lambda - k)$ and associated eigenvector X.
- The matrix $(A - kI)^{-1}$, i.e. the inverse (if it exists) of the matrix $(A - kI)$, has eigenvalue $\dfrac{1}{\lambda - k}$ and corresponding eigenvector X.

  Here k is any real number.

Task

The matrix $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 5 \end{bmatrix}$ has eigenvalues $\lambda = 5,\ 3,\ 1$ with associated eigenvectors $\begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ respectively.

The inverse A^{-1} exists and is $\quad A^{-1} = \dfrac{1}{3}\begin{bmatrix} 2 & -1 & -5 \\ -1 & 2 & -5 \\ 0 & 0 & \dfrac{3}{5} \end{bmatrix}$

Without further calculation write down the eigenvalues and eigenvectors of the following matrices:

(a) A^{-1} $\qquad$ (b) $\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 0 & 0 & 6 \end{bmatrix}$ $\qquad$ (c) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 3 \end{bmatrix}^{-1}$

Your solution

Answer

(a) The eigenvalues of A^{-1} are $\frac{1}{5},\ \frac{1}{3},\ 1$. (Notice that the dominant eigenvalue of A yields the smallest magnitude eigenvalue of A^{-1}.)

(b) The matrix here is $A + I$. Thus its eigenvalues are the same as those of A increased by 1 i.e. $6,\ 4,\ 2$.

(c) The matrix here is $(A - 2I)^{-1}$. Thus its eigenvalues are the reciprocals of the eigenvalues of $(A - 2I)$. The latter has eigenvalues $3,\ 1,\ -1$ so $(A - 2I)^{-1}$ has eigenvalues $\frac{1}{3},\ 1,\ -1$.

In each of the above cases the eigenvectors are the same as those of the original matrix A.

The power method

This is a **direct iteration** method for obtaining the **dominant** eigenvalue (i.e. the largest in magnitude), say λ_1, for a given matrix A and also the corresponding eigenvector.

We will not discuss the theory behind the method but will demonstrate it in action and, equally importantly, point out circumstances when it **fails**.

Task

Let $A = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix}$. By solving $\det(A - \lambda I) = 0$ obtain the eigenvalues of A and also obtain the eigenvector associated with the dominant eigenvalue.

Your solution

Answer

$$\det(A-\lambda I) = \begin{vmatrix} 4-\lambda & 2 \\ 5 & 7-\lambda \end{vmatrix} = 0$$

which gives

$$\lambda^2 - 11\lambda + 18 = 0 \qquad \Rightarrow \qquad (\lambda-9)(\lambda-2) = 0$$

so

$\lambda_1 = 9$ (the dominant eigenvalue) and $\lambda_2 = 2.$

The eigenvector $X = \begin{bmatrix} x \\ y \end{bmatrix}$ for $\lambda_1 = 9$ is obtained as usual by solving $AX = 9X$, so

$\begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 9x \\ 9y \end{bmatrix}$ from which $5x = 2y$ so $X = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ or any multiple.

If we normalize here such that the largest component of X is 1

$$X = \begin{bmatrix} 0.4 \\ 1 \end{bmatrix}$$

We shall now demonstrate how the power method can be used to obtain $\lambda_1 = 9$ and $X = \begin{bmatrix} 0.4 \\ 1 \end{bmatrix}$ where $A = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix}$.

- We choose an *arbitrary* 2×1 column vector

$$X^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- We premultiply this by A to give a new column vector $X^{(1)}$:

$$X^{(1)} = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

- We 'normalize' $X^{(1)}$ to obtain a column vector $Y^{(1)}$ with largest component 1: thus

$$Y^{(1)} = \frac{1}{12}\begin{bmatrix} 6 \\ 12 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$$

- We continue the process

$$X^{(2)} = AY^{(1)} = \begin{bmatrix} 4 & 2 \\ 6 & 7 \end{bmatrix}\begin{bmatrix} 1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 9.5 \end{bmatrix}$$

$$Y^{(2)} = \frac{1}{9.5}\begin{bmatrix} 4 \\ 9.5 \end{bmatrix} = \begin{bmatrix} 0.421053 \\ 1 \end{bmatrix}$$

Task

Continue this process for a further step and obtain $X^{(3)}$ and $Y^{(3)}$, quoting values to 6 d.p.

Your solution

Answer

$$X^{(3)} = AY^{(2)} = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix} \begin{bmatrix} 0.421053 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.684210 \\ 9.105265 \end{bmatrix}$$

$$Y^{(3)} = \frac{1}{9.105265} \begin{bmatrix} 0.404624 \\ 1 \end{bmatrix}$$

The first 8 steps of the above **iterative process** are summarized in the following table (the first three rows of which have been obtained above):

Table 1

Step r	$X_1^{(r)}$	$X_2^{(r)}$	α_r	$Y_1^{(r)}$	$Y_2^{(r)}$
1	6	12	12	0.5	1
2	4	9.5	9.5	0.421053	1
3	3.684211	9.105265	9.105265	0.404624	1
4	3.618497	9.023121	9.023121	0.401025	1
5	3.604100	9.005125	9.005125	0.400228	1
6	3.600911	9.001138	9.001138	0.400051	1
7	3.600202	9.000253	9.000253	0.400011	1
8	3.600045	9.000056	9.000056	0.400002	1

In Table 1, α_r refers to the largest magnitude component of $X^{(r)}$ which is used to normalize $X^{(r)}$ to give $Y^{(r)}$. We can see that α_r is converging to 9 which we know is the dominant eigenvalue λ_1 of A. Also $Y^{(r)}$ is converging towards the associated eigenvector $[0.4,\ 1]^T$.

Depending on the accuracy required, we could decide when to stop the iterative process by looking at the difference $|\alpha_r - \alpha_{r-1}|$ at each step.

Task

Using the power method obtain the dominant eigenvalue and associated eigenvector of

$$A = \begin{bmatrix} 3 & -1 & 0 \\ -2 & 4 & -3 \\ 0 & -1 & 1 \end{bmatrix} \quad \text{using a starting column vector} \quad X^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Calculate $X^{(1)}$, $Y^{(1)}$ and α_1:

Your solution

Answer

$$X^{(1)} = AX^{(0)} = \begin{bmatrix} 3 & -1 & 0 \\ -2 & 4 & -3 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$$

so $Y^{(1)} = \frac{1}{2}\begin{bmatrix} 1 \\ -0.5 \\ 0 \end{bmatrix}$ using $\alpha_1 = 2$, the largest magnitude component of $X^{(1)}$.

Carry out the next two steps of this iteration to obtains $X^{(2)}$, $Y^{(2)}$, α_2 and $X^{(3)}$, $Y^{(3)}$, α_3:

Your solution

Answer

$$X^{(2)} = \begin{bmatrix} 3 & -1 & 0 \\ -2 & 4 & -3 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 3.5 \\ -4 \\ 0.5 \end{bmatrix} \qquad Y^{(2)} = -\frac{1}{4}\begin{bmatrix} -0.875 \\ 1 \\ -0.125 \end{bmatrix} \qquad \alpha_2 = -4$$

$$X^{(3)} = \begin{bmatrix} 3 & -1 & 0 \\ -2 & 4 & -3 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} -0.875 \\ 1 \\ -0.125 \end{bmatrix} = \begin{bmatrix} -3.625 \\ 6.125 \\ -1.125 \end{bmatrix} \qquad Y^{(3)} = \frac{1}{6.125}\begin{bmatrix} -0.5918 \\ 1 \\ -0.1837 \end{bmatrix} \qquad \alpha_3 = 6.125$$

After just 3 iterations there is little sign of convergence of the normalizing factor α_r. However the next two values obtained are

$$\alpha_4 = 5.7347 \qquad \alpha_5 = 5.4774$$

and, after 14 iterations, $|\alpha_{14} - \alpha_{13}| < 0.0001$ and the power method converges, albeit slowly, to

$$\alpha_{14} = 5.4773$$

which (correct to 4 d.p.) is the dominant eigenvalue of A. The corresponding eigenvector is

$$\begin{bmatrix} -0.4037 \\ 1 \\ -0.2233 \end{bmatrix}$$

It is clear that the power method requires, for its practical execution, a computer.

Problems with the power method

1. If the initial column vector $X^{(0)}$ is an eigenvector of A other than that corresponding to the dominant eigenvalue, say λ_1, then the method will fail since the iteration will converge to the wrong eigenvalue, say λ_2, after only one iteration (because $AX^{(0)} = \lambda_2 X^{(0)}$ in this case).

2. It is possible to show that the speed of convergence of the power method depends on the ratio

$$\frac{\text{magnitude of dominant eigenvalue } \lambda_1}{\text{magnitude of next largest eigenvalue}}$$

   If this ratio is small the method is slow to converge.

   In particular, if the dominant eigenvalue λ_1 is **complex** the method will fail completely to converge because the complex conjugate $\overline{\lambda}_1$ will also be an eigenvalue and $|\lambda_1| = |\overline{\lambda}_1|$

3. The power method only gives one eigenvalue, the dominant one λ_1 (although this is often the most important in applications).

Advantages of the power method

1. It is simple and easy to implement.

2. It gives the eigenvector corresponding to λ_1 as well as λ_1 itself. (Other numerical methods require separate calculation to obtain the eigenvector.)

Finding eigenvalues other than the dominant

We discuss this topic only briefly.

1. Obtaining the smallest magnitude eigenvalue

If A has dominant eigenvalue λ_1 then its inverse A^{-1} has an eigenvalue $\dfrac{1}{\lambda_1}$ (as we discussed at the beginning of this Section.) Clearly $\dfrac{1}{\lambda_1}$ will be the smallest magnitude eigenvalue of A^{-1}. Conversely if we obtain the **largest** magnitude eigenvalue, say λ_1', of A^{-1} by the power method then the **smallest** eigenvalue of A is the reciprocal, $\dfrac{1}{\lambda_1'}$.

This technique is called the **inverse power method**.

Example

If $A = \begin{bmatrix} 3 & -1 & 0 \\ -2 & 4 & -3 \\ 0 & -1 & 1 \end{bmatrix}$ then the inverse is $A^{-1} = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 3 & 9 \\ 2 & 3 & 10 \end{bmatrix}$.

Using $X^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in the power method applied to A^{-1} gives $\lambda_1' = 13.4090$. Hence the smallest magnitude eigenvalue of A is $\dfrac{1}{13.4090} = 0.0746$. The corresponding eigenvector is $\begin{bmatrix} 0.3163 \\ 0.9254 \\ 1 \end{bmatrix}$.

In practice, finding the inverse of a large order matrix A can be expensive in computational effort. Hence the inverse power method is implemented without actually obtaining A^{-1} as follows.

As we have seen, the power method applied to A utilizes the scheme:

$$X^{(r)} = AY^{(r-1)} \qquad r = 1, 2, \ldots$$

where $Y^{(r-1)} = \dfrac{1}{\alpha_{r-1}}X^{(r-1)}$, α_{r-1} being the largest magnitude component of $X^{(r-1)}$.
For the inverse power method we have

$$X^{(r)} = A^{-1}Y^{(r-1)}$$

which can be re-written as

$$AX^{(r)} = Y^{(r-1)}$$

Thus $X^{(r)}$ can actually be obtained by solving this system of linear equations without needing to obtain A^{-1}. This is usually done by a technique called LU decomposition i.e. writing A (once and for all) in the form

$A = LU \qquad L$ being a lower triangular matrix and U upper triangular.

2. Obtaining the eigenvalue closest to a given number p

Suppose λ_k is the (unknown) eigenvalue of A closest to p . We know that if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A then $\lambda_1 - p,\ \lambda_2 - p, \ldots, \lambda_n - p$ are the eigenvalues of the matrix $A - pI$. Then $\lambda_k - p$ will be the **smallest** magnitude eigenvalue of $A - pI$ but $\dfrac{1}{\lambda_k - p}$ will be the **largest** magnitude eigenvalue of $(A - pI)^{-1}$. Hence if we apply the power method to $(A - pI)^{-1}$ we can obtain λ_k. The method is called the **shifted inverse power** method.

3. Obtaining all the eigenvalues of a large order matrix

In this case neither solving the characteristic equation $\det(A - \lambda I) = 0$ nor the power method (and its variants) is efficient.

The commonest method utilized is called the **QR technique**. This technique is based on **similarity transformations** i.e. transformations of the form

$$B = M^{-1}AM$$

where B has the same eigenvalues as A. (We have seen earlier in this Workbook that one type of similarity transformation is $D = P^{-1}AP$ where P is formed from the eigenvectors of A. However, we are now, of course, dealing with the situation where we are trying to find the eigenvalues and eigenvectors of A.)

In the QR method A is reduced to upper (or lower) triangular form. We have already seen that a triangular matrix has its eigenvalues on the diagonal.

For details of the QR method, or more efficient techniques, one of which is based on what is called a **Householder** transformation, the reader should consult a text on numerical methods.