UNIT 1 THE EIGENVALUE PROBLEM FOR MATRICES

Structure

1.1 INTRODUCTION

You are familiar with matrices from your +2 Mathematics course. You have learnt to define a matrix as a square or rectangular array of numbers that obeys certain laws. You have also performed the operations of addition and multiplication on matrices and determined the transpose and inverse of a matrix. Using these operations, you could obtain solutions of systems of linear equations as well.

We would now like to go deeper into the subject, particularly from the point of view of physics. In physics, matrices occur mainly in two ways: first, in the solution of systems of linear equations, for example, in electrical networks; second, in the solution of eigenvalue problems in classical and quantum mechanics. We shall begin our discussion with one illustration of how matrices occur in physics. Since one of the major applications of matrices is in solving eigenvalue problems in classical and quantum mechanics, we will confine ourselves in this unit mostly to this aspect. For simplicity, we shall consider only square matrices. You will learn to solve the **eigenvalue problem** for simple physical systems. This involves determination of **eigenvalues** and **eigenvectors**. In this connection you will also study the **Cayley-Hamilton Theorem, the minimal equation** and their applications.

In the next unit, we shall extend this discussion to special matrices having certain symmetry properties.

Objectives

After studying this unit, you should be able to:

- determine the eigenvalues and eigenvectors of a given matrix;
- prove some general properties of eigenvalues and eigenvectors;
- state and apply the Cayley-Hamilton Theorem; and
- apply the minimal equation.

Study Guide

To study this unit effectively you should be familiar with standard calculations done with matrices. You should know how to find the determinant of a matrix, rank of a matrix, minor, co-factors, inverse of matrix, and calculate products of matrices etc. You know that matrix equations are nothing but linear algebraic equations written in a compact form. Therefore, if you know how to solve linear algebraic equations (both homogeneous and inhomogeneous), you will find it easy to follow this suit. You may like to brush up your previous knowledge before studying this unit. For this, you may refer to the +2 Mathematics book of NCERT. Its reference is given at the end of the block. While studying this unit, we advise you to go carefully through all steps and work them out yourself. You should also solve the SAQs and terminal questions on your own. Only if you do all calculations yourself will you be able to understand and appreciate the contents of the unit.

1.2 MATRICES IN PHYSICS: AN INTRODUCTION

There are many physical situations in which matrices occur and there can be as many ways of introducing matrices in physics. However, here we have selected the specific example of rotation of coordinate axes because you are familiar with it. In Unit 1 of the course PHE-04, you have studied how vectors transform under the rotation of a two dimensional coordinate system about its origin. So we will begin our discussion by revisiting this problem and see how introducing matrices simplifies the problem.

Refer to Fig. 1.1. It shows the rotation of a two-dimensional coordinate system S about the origin by an angle θ in the counter-clockwise direction. Let S' be the rotated system.

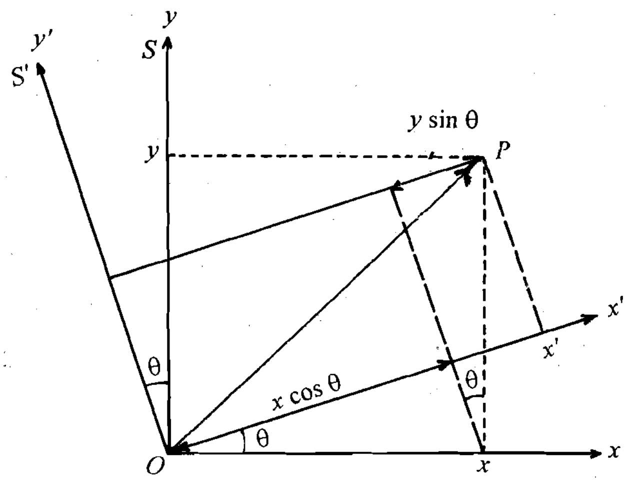

Fig.1.1: Rotation of a two-dimensional coordinate system about its origin

Now consider the vector **OP** in Fig. 1.1. Let the components of **OP** be denoted by (x, y) in S and (x', y') in S'.
Thus

$$\mathbf{OP} = x\,\hat{\mathbf{i}} + y\,\hat{\mathbf{j}}$$

$$= x'\,\hat{\mathbf{i}}' + y'\,\hat{\mathbf{j}}'$$

where $\hat{\mathbf{i}}, \hat{\mathbf{j}}$ and $\hat{\mathbf{i}}', \hat{\mathbf{j}}'$ are the unit vectors along the coordinate axes in S and S', respectively. Now from Fig. 1.1 you can see that the components (x, y) and (x', y') are related through the equations:

$$x' = x\cos\theta + y\sin\theta \qquad (1.1a)$$

$$y' = -x\sin\theta + y\cos\theta \qquad (1.1b)$$

From Sec. 1.3.2 of Unit 1 (PHE-04), you may recall that these equations can be expressed in terms of matrices and products of matrices as follows:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \qquad (1.2)$$

We now use the following notations

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix},$$

$$\mathbf{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

When a matrix has only a single column (row), it is usual to call it a column (row) vector and write it in bold type, e.g., **x**.

and

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

and write Eq. (1.2) in a compact and simple form

$$\mathbf{x}' = A\,\mathbf{x} \qquad (1.3)$$

In Eq. (1.3), $\mathbf{x}'$ is a 2×1 column matrix which represents the vector **OP** in S'. Similarly, $\mathbf{x}$ is a 2×1 column matrix which represents **OP** in S. The matrix A which depends on θ connects the components of **OP** in S and S'. It is called the **rotation matrix**. Eq. (1.3) represents a **single rotation** of the coordinates (x, y) to (x', y') through an angle θ in the counter-clockwise direction.

So you can see that by introducing the concept of matrices we have been able to write the rotation equations (1.1a and b) in a more compact manner. This is a very simple example of how matrices occur in physics. For example, we can use Eq. (1.3) to compute the coordinates of a point P with respect to a rotating merry-go-round when its coordinates in a system fixed to the earth are given to us (see Fig. 1.2).

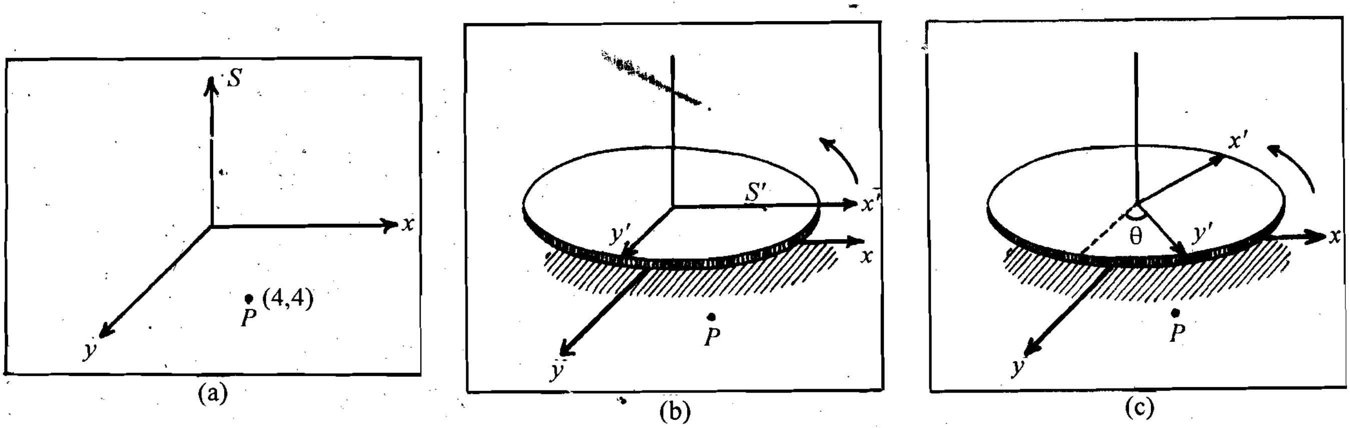

Fig.1.2: (a) Suppose the coordinates of a point P are $x = 4$m, $y = 4$m in the coordinate system S fixed to the ground; (b) Suppose the system S' fixed to the merry-go-round is rotating in the anticlockwise direction; (c) When $\theta = 30°$, say, the coordinates of P in S' are given by

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos 30° & \sin 30° \\ -\sin 30° & \cos 30° \end{pmatrix} \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

or

$$x' = \frac{\sqrt{3}}{2}(4) + \frac{1}{2}(4) = 2\left(\sqrt{3}+1\right)\text{m} \quad \text{and} \quad y' = -\frac{1}{2}(4) + \frac{\sqrt{3}}{2}(4) = 2\left(-1+\sqrt{3}\right)\text{m}$$

Now suppose the coordinate system S' of Fig. 1.1 is further rotated through an angle ϕ, taking the coordinates (x', y') to (x'', y''). Then we can write

$$x'' = x'\cos\phi + y'\sin\phi \qquad (1.4a)$$

$$y'' = -x' \sin\phi + y' \cos\phi \tag{1.4b}$$

In the matrix notation this becomes

$$\mathbf{x}'' = B\,\mathbf{x}' \tag{1.5}$$

where

$$\mathbf{x}'' = \begin{pmatrix} x'' \\ y'' \end{pmatrix}, \quad \mathbf{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{1.6}$$

and

$$B = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \tag{1.7}$$

Now from Eq. (1.3) $\mathbf{x}' = A\,\mathbf{x}$, and we can write

Rule of Matrix Multiplication:

If $C = AB$

$$(C)_{ij} = \sum_{k=1}^{n} (A)_{ik} (B)_{kj}$$

where C, A and B are square matrices of order n. Thus

$$C_{11} = A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} + \ldots + A_{1n}B_{n1}$$

and so on.

$$\mathbf{x}'' = B\,\mathbf{x}' = BA\,\mathbf{x} \equiv C\,\mathbf{x} \tag{1.8}$$

where the product $C \equiv BA$ represents the rotation matrix corresponding to a rotation by an angle $(\theta + \phi)$:

$$C = \begin{pmatrix} \cos(\theta+\phi) & \sin(\theta+\phi) \\ -\sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix} \tag{1.9}$$

You may like to quickly verify this result by substituting x' and y' from Eqs. (1.1a and b) in Eqs. (1.4a and b). However, you should note that this kind of additivity of angles θ and ϕ holds only for rotation in a plane.

SAQ 1

Spend 5 min

Verify that $C \equiv BA$ is the rotation matrix for the rotation by angle $(\theta + \phi)$ in a plane.

From this SAQ you see that the lengthy algebra turns into a very simple problem of (matrix) multiplication if we introduce matrices. For each successive rotation α, β, γ etc. we need to write equations like (1.5) and multiply successive rotation matrices. In this manner, matrices bring in a lot of convenience in representing physical situations and performing calculations on them! This is true particularly for three-dimensional systems.

We can use equations like (1.3) and (1.5) to represent other physical systems. For example, the vectors $\mathbf{x}'$ and $\mathbf{x}$ may describe voltage gradient and current, respectively, if A has the physical dimensions of electrical resistivity. In general, we say that given one vector, say $\mathbf{x}$, Eq. (1.3) determines another vector $\mathbf{x}'$ through A.

In physics, matrices are also used to represent operators. Now every operator has certain special vectors which are characteristic of the operator such that the operator transforms these vectors essentially into themselves. These special vectors are called **eigenvectors** of the operator. Thus for every matrix A corresponding to an operator A, there exist special vectors such that

$$A\,\mathbf{x} = \lambda\,\mathbf{x} \tag{1.10}$$

where λ is a scalar, which could be real or complex. Thus when A operates on $\mathbf{x}$, it yields a multiple of $\mathbf{x}$ itself.

Eq. (1.10) is referred to as the **eigenvalue equation** for matrix A.

Definition

For a given matrix A, the column vector $\mathbf{x}$ which satisfies the **eigenvalue equation**

$$A\,\mathbf{x} = \lambda\,\mathbf{x}$$

is called an **eigenvector** of A, and the constant λ is called the corresponding **eigenvalue**.

Notice in Eq. (1.10) that the operation of matrix A on a vector $\mathbf{x}$, yields a vector $\mathbf{x}' \equiv A\,\mathbf{x}$ which is parallel to $\mathbf{x}$ and is just a multiple of $\mathbf{x}$. The eigenvalue problem is of particular interest in physics with many applications in classical and quantum mechanics. So let us now consider this problem in detail.

1.3 THE EIGENVALUE PROBLEM

You have just learnt that the equation

$$A\,\mathbf{x} = \lambda\,\mathbf{x} \tag{1.10}$$

is called an eigenvalue equation. In this equation, λ is a scalar which can be real or complex, $\mathbf{x}$ is said to be an eigenvector of A and λ is the corresponding eigenvalue of A. We say that the eigenvector $\mathbf{x}$ belongs to the eigenvalue λ. If $\mathbf{x}$ is identically zero (i.e., all its elements are zero), then we say that $\mathbf{x}$ is a **trivial** eigenvector. If $\mathbf{x}$ is not identically zero (i.e., at least one of its elements is non-zero) we say that $\mathbf{x}$ is a **non-trivial** eigenvector. In physics, only non-trivial eigenvectors are important. Therefore, hereafter, whenever we use the term eigenvector, we shall always mean a non-trivial eigenvector.

Generally, we use the symbol or **u** or **v** for denoting a column vector.

An important example of the eigenvalue equation in physics is that of the Schrödinger equation $H\psi = E\psi$ about which you have read in Blocks 2 and 3 of the physics course PHE-11. Here H is the Hamiltonian of a quantum mechanical system, ψ is the eigenfunction of H and E (the energy) its eigenvalue.

Once we know the eigenvalue equation for a physical system, the next step is obviously to solve it. This means that we should be able to determine the eigenvalues and eigenvectors of a given matrix. This is what you are going to learn now.

1.3.1 Determination of Eigenvalues and Eigenvectors

To solve Eq. (1.10), we rewrite it as

$$A\,\mathbf{x} - \lambda\,I\,\mathbf{x} = 0$$

or

$$(A - \lambda\,I)\,\mathbf{x} = 0 \tag{1.11}$$

Notice that $(A - \lambda)$ does not make sense, since A is a matrix and λ, a scalar. We write

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

which gives

$$A\,\mathbf{x} = \lambda\,I\,\mathbf{x}$$
$$= \lambda\,\mathbf{x}$$

where we have written $\mathbf{x}$ as $I\mathbf{x}$. Here I is the unit matrix of the same dimension as A. For example, if A is a 2×2 matrix, then

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

Similarly, if A is a 3×3, or an $n \times n$ matrix, then I is, respectively, given by

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

or

$$I = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

In compact notation:

$$I_{ij} = 1 \quad \text{if } i = j$$

$$= 0 \quad \text{if } i \neq j$$

Written in the form of Eq. (1.11), the eigenvalue problem reduces to the problem of solving a set of linear homogeneous equations. We now explain this idea with the help of an example.

Example 1

If

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

write the eigenvalue equation as a set of linear homogeneous equations.

Solution

As the matrix is three-dimensional there will be three linear equations. We can obtain these by substituting A and $\mathbf{x}$ in Eq. (1.11) as follows:

$$\left\{ \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \right\} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

or

$$\begin{pmatrix} 2-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

or

$$(2-\lambda)\,x_1 + x_3 = 0 \qquad \text{(A)}$$

$$(1-\lambda)\,x_2 = 0 \qquad \text{(B)}$$

$$x_1 - \lambda\,x_3 = 0 \qquad \text{(C)}$$

Thus we obtain 3 linear homogeneous equation (A), (B) and (C).

You may like to practice this algebra before proceeding further. Try the SAQ given below.

SAQ 2

Spend 5 min

Write the eigenvalue equation $M\mathbf{x} = \lambda\,\mathbf{x}$ as a set of linear homogeneous equations given

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Alternatively, we can write a set of linear equations as an eigenvalue equation involving a matrix.

Example 2

Write the equations

$$(a-\lambda)x_1 + bx_2 = 0$$
$$cx_1 + (d-\lambda)x_2 = 0$$

as an eigenvalue equation.

Solution

The above equations can be written as a product of the following two matrices:

$$\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

or

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right\} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

or

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Hence we can write

$$M\mathbf{x} = \lambda\,\mathbf{x}$$

where $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

Thus, you have seen that the problem of solving the eigenvalue problem is reduced to the problem of solving a set of linear homogeneous equations.

Let us generalise these results to an $n \times n$ matrix. Let A be an $n \times n$ matrix and **x** be represented by an $n \times 1$ column matrix. Then we have

$$(A - \lambda I)\,\mathbf{x} = 0 \tag{1.12a}$$

where I is the $n \times n$ unit matrix with elements

$$I_{ij} = \begin{cases} 1 & \text{if } i = j,\ i, j = 1,2,\ldots,n \\ 0 & \text{if } i \neq j,\ i, j = 1,2,\ldots,n \end{cases} \tag{1.12b}$$

Now Eq. (1.12a) gives rise to n linear homogeneous equations:

$$\left\{\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}\right\} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = 0$$

or

$$\begin{pmatrix} A_{11} - \lambda & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} - \lambda & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = 0 \tag{1.13}$$

or

$$\begin{aligned} (A_{11} - \lambda)x_1 + A_{12}x_2 + \ldots + A_{1n}x_n &= 0 \\ A_{21}x_1 + (A_{22} - \lambda_2)x_2 + \ldots + A_{2n}x_n &= 0 \\ &\vdots \\ A_{n1}x_1 + A_{n2}x_2 + \ldots + (A_{nn} - \lambda)x_n &= 0 \end{aligned} \tag{1.14}$$

Thus to find the eigenvalues λ we have to solve a set of n equations as in Eq. (1.14). We now state a result with which you are familiar from the theory of linear equations (You may like to refer to Units 4 and 5 of Block 2 of the course MTE-04 entitled Elementary Algebra).

The set of n linear homogeneous equations in Eq. (1.14) *has a non-trivial solution if and only if*

$$\det(A-\lambda I) = \begin{vmatrix} A_{11}-\lambda & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22}-\lambda & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn}-\lambda \end{vmatrix} = 0 \qquad (1.15)$$

The determinant given in Eq. (1.15) is called the **secular determinant** and Eq. (1.15) is called the **characteristic equation** for A. You can easily see that this condition comes about as follows.

Consider Eq. (1.12a):

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

If the matrix $(A - \lambda I)$ is non-singular, then its inverse exists and hence on multiplying by $(A - \lambda I)^{-1}$ from the left we get

A matrix is non-singular if its determinant is non-zero. If $\det A = 0$, A is said to be a singular matrix. We write $\det A$ or $|A|$ to denote the determinant.

$$(A-\lambda I)^{-1}(A-\lambda I)\mathbf{x} = 0$$

or

$$I\,\mathbf{x} = 0 \qquad (\because A^{-1}A = I)$$

or

$$\mathbf{x} = 0$$

This solution is trivial since **x** is identically zero. Thus in order that the solution of Eq. (1.12) be non-trivial, we must have $\det(A - \lambda I) = 0$, which is what Eq. (1.15) is. In fact, the condition given by Eq. (1.15) is both necessary and sufficient. For any given matrix A, $\det(A - \lambda I)$ can vanish only for certain values of λ. The values of λ, for which $\det(A - \lambda I) = 0$, are called the eigenvalues of A. To sum up:

To determine the eigenvalues of a given matrix A,

- evaluate $\det(A - \lambda I)$ and
- determine the solutions of the equation: $\det(A - \lambda I) = 0$.

You know how to evaluate the determinant of a given matrix. When the determinant in Eq. (1.15) is evaluated, it would yield an nth degree polynomial in λ, say

$$|A-\lambda I| = c_n\lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0 = 0 \qquad (1.16)$$

This polynomial is called the **characteristic polynomial** of A, and Eq. (1.16) is called the **characteristic equation** or the **secular equation** of the matrix. We conclude that the **eigenvalues of a matrix A are just the roots of the characteristic equation (1.16).** By the Fundamental Theorem of algebra, an nth order polynomial with real coefficients has n roots, which we may denote by $\lambda_1, \lambda_2, \ldots, \lambda_n$. These may be real or complex. We may now conclude that

> An $n \times n$ matrix has r eigenvalues such that, $0 \le r \le n$. However, these eigenvalues need not be distinct. Some or all of them may be equal to one another.

Let us now understand these ideas with the help of an example.

Example 3

Determine the eigenvalue of the matrix

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

Solution

Taking det $(M - \lambda I)$, we obtain

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

$$= (2-\lambda)^2 (1-\lambda)$$

Hence the eigenvalues are given by the roots of the equation

$$\det(M - \lambda I) = (2-\lambda)^2 (1-\lambda) = 0$$

i.e., $\lambda = 2,2$ and 1. Therefore, the eigenvalue 2 occurs twice and 1 once. Altogether there are only 3 eigenvalues as it should be.

The set of all eigenvalues of a matrix is called the **eigenvalue spectrum** of the matrix. If the characteristic equation has r coincident roots, then the **multiplicity** of this eigenvalue is r. In Example 3, the eigenvalue spectrum is (2, 2, 1) and the eigenvalue 2 has a multiplicity two.

You may yourself like to determine the eigenvalues of a given matrix. Attempt the following SAQ.

SAQ 3

Spend 10 min

Determine the eigenvalues of

a) $$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and

b) $$A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}.$$

So far you have learnt how to determine the eigenvalues of a matrix. What about the eigenvectors? How do we determine these? To answer this question we consider the matrix M of SAQ 3 (a) above as an example.

Determining eigenvectors

You have determined the eigenvalues of M to be $+1$ and -1. Let us now determine the eigenvectors corresponding to each of these eigenvalues. Let us first take $\lambda = +1$. The **procedure** is to substitute this value of λ in the set of equations for x_1 and x_2 and solve them. In this case the linear algebraic equations are

$$-\lambda x_1 + x_2 = 0$$

and

$$x_1 - \lambda x_2 = 0$$

For $\lambda = +1$ we get

$$-x_1 + x_2 = 0$$

$$x_1 - x_2 = 0$$

Hence $x_1 = x_2$.

Thus $\mathbf{x} = \begin{pmatrix} x_1 \\ x_1 \end{pmatrix} \equiv x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ with x_1 being an arbitrary constant.

For $\lambda = -1$, we have

$$x_1 + x_2 = 0$$

$$x_1 + x_2 = 0.$$

Hence $x_2 = -x_1$.

Thus $\mathbf{x} = \begin{pmatrix} x_1 \\ -x_1 \end{pmatrix} \equiv x_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ with x_1 being arbitrary.

To summarize, the eigenvalues of M are $\lambda = +1$ and $\lambda = -1$. The eigenvector belonging to $\lambda = +1$ is $x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and that belonging to $\lambda = -1$ is $x_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

As another example let us determine the eigenvalues and eigenvectors of a 3×3 matrix.

Example 4

Determine the eigenvalues and eigenvectors of the matrix

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

Solution

The characteristic polynomial of the matrix A is

$$\begin{vmatrix} -1-\lambda & 1 & 1 \\ 1 & -1-\lambda & 1 \\ 2 & 1 & 2-\lambda \end{vmatrix}$$

$$= -(1+\lambda)[(-1-\lambda)(2-\lambda)-1] - 1[(2-\lambda)-2] + 1[1+(1+\lambda)2]$$

$$= (1+\lambda)^2(2-\lambda) + (1+\lambda) - 2 + \lambda + 2 + 3 + 2\lambda$$

$$= -\lambda^3 + 3\lambda + 2 + 4 + 4\lambda$$

$$= -\lambda^3 + 7\lambda + 6$$

So the characteristic equation is

$$\lambda^3 - 7\lambda - 6 = 0$$

You can verify (by substitution) that its roots are –2, –1 and 3. These are, therefore, the eigenvalues of A.

How about the eigenvectors? For $\lambda = -2$, Eq. (1.3) becomes

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

This yields just two independent equations

$$x_1 + x_2 + x_3 = 0$$

$$2x_1 + x_2 + 4x_3 = 0$$

You can solve these for any two of the variables in terms of the third to obtain

$$x_1 = -3x_3, \qquad x_2 = 2x_3$$

The value of x_3 can be arbitrarily chosen. Let us choose $x_3 = 1$. Then the column vector $\begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}$ or any multiple of it is an eigenvector of the given matrix, belonging to the eigenvalue –2.

You should now be able to determine the eigenvectors belonging to $\lambda = -1$ and $\lambda = 3$. They are $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$, respectively (or multiples of these).

You may now like to calculate the eigenvectors of a matrix. Try it for the matrix A of SAQ 3(b).

SAQ 4

Spend 5 min

Determine the eigenvectors of the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$

Finally, we consider a familiar physical situation and discuss how it can be readily solved as an eigenvalue problem.

Example 5: Free vibrations of a coupled spring-mass system

Consider two masses m_1 and m_2 connected to each other and to fixed (unmoving) supports A and B by springs. You have studied this system in Unit 5 of the course PHE-02 entitled Oscillations and Waves. The force constants (i.e., the force required to produce unit change in length) of each spring are as shown in Fig. 1.3. Let $x_1(t)$ and $x_2(t)$ be the displacements of m_1 and m_2 from their equilibrium positions at time t. Then from Newton's second law we get the equation of motion for longitudinal vibrations of the masses m_1 and m_2 to be the following:

$$m_1 \frac{d^2 x_1}{dt^2} = -k' x_1 + k(x_2 - x_1) = -(k + k')x_1 + kx_2$$

$$m_2 \frac{d^2 x_2}{dt^2} = k(x_1 - x_2) - k' x_2 = kx_1 - (k + k')x_2 \qquad (1.17)$$

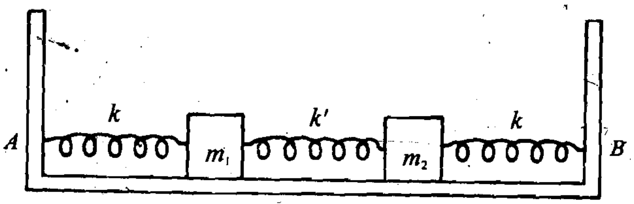

Fig.1.3: Coupled spring-mass system.

We now ask the following question: Do these equations allow **both** the masses to move **simple harmonically** and with **identical** frequencies? For example, can we have

$$x_1(t) = a_1 \sin \omega t, \qquad x_2(t) = a_2 \sin \omega t \qquad (1.18)$$

as a solution of the pairs of Eqs. (1.17)? Thus, we are looking for solutions in which both the masses vibrate simple harmonically with the same frequency. To see whether this is possible, let us substitute this trial solution in Eq. (1.17). The common factor of $\sin \omega t$ gets divided out. Dividing the two equations by $(-m_1)$ and $(-m_2)$, respectively, we get

$$\frac{k+k'}{m_1} a_1 - \frac{k}{m_1} a_2 = \omega^2 a_1$$

$$-\frac{k}{m_2} a_1 + \frac{k+k'}{m_2} a_2 = \omega^2 a_2 \qquad (1.19)$$

Do you recognise that this has the form of an eigenvalue equation? Clearly, this pair of equations can be written as the matrix equation

$$\begin{pmatrix} \frac{k+k'}{m_1} & -\frac{k}{m_1} \\ -\frac{k}{m_2} & \frac{k+k'}{m_2} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \omega^2 \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

or

$$K\mathbf{a} = \omega^2 \mathbf{a} \qquad (1.20a)$$

where

$$K = \begin{pmatrix} \frac{k+k'}{m_1} & -\frac{k}{m_1} \\ -\frac{k}{m_2} & \frac{k+k'}{m_2} \end{pmatrix} \text{ and } \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \qquad (1.20b)$$

This is of the same form as Eq. (1.10). It is thus an **eigenvalue equation** for the given system.

You should note that our system will execute simple harmonic vibrations described by Eq. (1.18), provided Eq. (1.19) or Eq. (1.20a) have non-trivial solutions (i.e., solutions other than the trivial one ($a_1 = a_2 = 0$). To examine whether non-trivial solutions exist in this case, let us simplify matters by considering a special case, taking $m_1 = m_2 = m$. Using this in Eqs. (1.20a and b) and introducing the notations

$$\frac{k}{m} = \omega_a^2, \quad \frac{k+k'}{m} = \omega_b^2 \tag{1.21}$$

we can write Eqs. (1.19) as

$$\begin{aligned} \left(\omega_b^2 - \omega^2\right)a_1 - \omega_a^2 a_2 &= 0 \\ -\omega_a^2 a_1 + \left(\omega_b^2 - \omega^2\right)a_2 &= 0 \end{aligned} \tag{1.22}$$

or

$$\begin{pmatrix} \omega_b^2 - \omega^2 & -\omega_a^2 \\ -\omega_a^2 & \omega_b^2 - \omega^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0$$

or

$$D\,\mathbf{a} = 0$$

where

$$D = \begin{pmatrix} \omega_b^2 - \omega^2 & -\omega_a^2 \\ -\omega_a^2 & \omega_b^2 - \omega^2 \end{pmatrix} \tag{1.23}$$

Now, a non-trivial solution exists if and only if $|D| = 0$. Thus we get the condition

$$\det D = \left(\omega_b^2 - \omega^2\right)^2 - \left(\omega_a^2\right)^2 = 0 \tag{1.24}$$

This condition obviously requires that ω^2 should take one of the following two values:

$$\omega^2 = \omega_b^2 - \omega_a^2 = \frac{k'}{m} = \omega_1^2$$

or (1.25)

$$\omega^2 = \omega_b^2 + \omega_a^2 = \frac{2k+k'}{m} = \omega_2^2$$

Thus the eigenvalue equation $K\mathbf{a} = \omega^2\mathbf{a}$ has non-trivial solutions ($\mathbf{a} \neq 0$) only for two particular values of ω^2, namely, ω_1 and ω_2 given by Eq. (1.25). These values are the **eigenvalues** of the matrix K.

What is the physical significance of the eigenvalues ω_1 and ω_2? Evidently, the significance is that these are the only two frequencies for which both the particles of the system can execute simple harmonic motion governed by Eq. (1.8).

To complete the analysis, we must determine the solutions for a_1 and a_2 when ω equals either of the eigenvalues ω_1 or ω_2. Sticking to the same special case ($m_1 = m_2 = m$) as before, let us set $\omega^2 = \omega_1^2 = k'/m$ in Eq. (1.19). You can easily verify that the first equation leads to

$$a_1 = a_2 \qquad \text{for} \qquad \omega^2 = \omega_1^2 \tag{1.26a}$$

and the second equation tells you exactly the same thing. Repeat the process with $\omega^2 = \omega_2^2$. You will get

$$a_1 = -a_2 \qquad \text{for} \qquad \omega^2 = \omega_2^2 \tag{1.26b}$$

In each case, **only the relation between the elements** of the **vector a is determined**, and not the individual elements. So the value of one of the elements can be arbitrarily chosen, e.g., $a = \frac{1}{\sqrt{2}}$. We can thus write the result as

$$\mathbf{a} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{for} \qquad \omega^2 = \omega_1^2,$$

and

$$\mathbf{a} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad \text{for} \qquad \omega^2 = \omega_2^2.$$

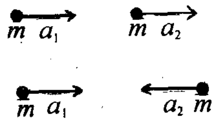

Fig.1.4: Normal modes of vibration for the coupled spring-mass system.

These are the **eigenvectors** of K belonging to the eigenvalues ω_1 and ω_2. These represent the **normal modes** of vibration of the system.

What these eigenvectors tell you is this: When vibrating with frequency ω_1 both the masses move together ($a_1 = a_2$, so $x_1 = x_2$ at all times); when the frequency is ω_2 they move with equal amplitudes but in opposite direction (see Fig. 1.4). Notice that the frequencies ω_1 and ω_2 are the square roots of the eigenvalues and are called the **eigenfrequencies**, **characteristic frequencies**, or **natural frequencies** of the system. The two eigenvectors are also called the two **natural** modes of vibration, "natural" in the sense that they represent the free or natural vibration of the system.

For clarifying these ideas further, you may like to apply them to another physical system.

SAQ 5

Spend 10 min

Consider the three vibrating masses shown in Fig. 1.5. Assume that the motion of the system is one-dimensional and the spring forces obey Hooke's law. Obtain the characteristic frequencies of this system. The equations of motion governing the system are

$$\ddot{x}_1 = -\frac{k}{M}(x_1 - x_2)$$

$$\ddot{x}_2 = -\frac{k}{m}(x_2 - x_1) - \frac{k}{m}(x_2 - x_3)$$

$$\ddot{x}_3 = -\frac{k}{M}(x_3 - x_2)$$

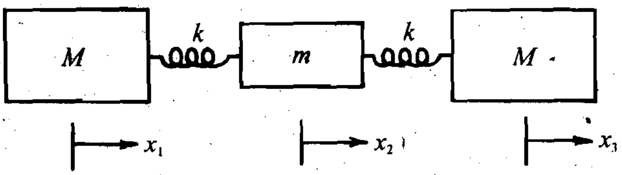

Fig.1.5: Coupled spring-mass oscillations

So far you have acquired the basic understanding of what an eigenvalue equation for a matrix is. You can write the corresponding characteristic equation and calculate the eigenvalues and eigenvectors for simple problems involving 3 × 3 matrices.

We now present some important results concerning eigenvalues and eigenvectors. These will further help you in solving eigenvalue problems.

1.3.2 General Results Concerning Eigenvalues and Eigenvectors

i) **If a matrix is diagonal, its diagonal elements straight away give its eigenvalues.**

A matrix for which only the diagonal elements are non-zero, the rest being zero is called a *diagonal matrix*.

You can verify this easily:

Let

$$M = \begin{pmatrix} M_{11} & 0 & \cdots & 0 \\ 0 & M_{22} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & M_{nn} \end{pmatrix}$$

The characteristic equation is given by

$$|M - \lambda I| = 0 = (M_{11} - \lambda)(M_{22} - \lambda)\cdots(M_{nn} - \lambda)$$

Hence the eigenvalues are M_{11}, M_{22}, ... M_{nn}, which are the diagonal elements of M.

ii) **The coefficients c_i of the characteristic equation (Eq. 1.16) are related to the eigenvalues.**

The eigenvalues of a given $n \times n$ matrix M are determined from the characteristic equation. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the roots of this polynomial equation, we can write Eq. (1.16) as

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

Then to relate the coefficients c_i $(i = 0,1,\ldots n)$ of Eq. (1.16) to the eigenvalues, we write the equation in a factorised form

$$c_n\lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0 = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\ldots(\lambda_n - \lambda) \qquad (1.27a)$$

and match the coefficients of like powers of λ on both sides. For instance, the coefficient of λ^n on the left hand side is c_n, while on the right hand side it is given by $(-1)^n$. Hence,

$$c_n = (-1)^n \qquad (1.27b)$$

Similarly,

$$c_{n-1} = (-1)^{n-1}(\lambda_1 + \lambda_2 + \ldots + \lambda_n) \qquad (1.27c)$$

The last constant c_0 is the co-efficient of λ^0. This is related to $\lambda_1,\ldots,\lambda_n$ as

$$c_0 = \lambda_1\lambda_2\ldots\lambda_n \qquad (1.27d)$$

Using Eqs. (1.27c) and (1.27d) we can arrive at two important results relating the sum and product of eigenvalues to the trace and determinant of M.

iii) **The product of eigenvalues equals the determinant of a matrix.**

To see this let us again write the characteristic equation in the form

$$\begin{vmatrix} M_{11}-\lambda & M_{12} & & M_{1n} \\ M_{21} & M_{22}-\lambda & \cdots & M_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn}-\lambda \end{vmatrix} = \det(M-\lambda I) = c_n\lambda^n + \ldots + c_1\lambda + c_0 \tag{1.28}$$

Regarding this as an identity in λ, if we set $\lambda = 0$, we get

$$\begin{vmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn} \end{vmatrix} = \det M = c_0 \tag{1.29}$$

Combining Eqs. (1.27d) and (1.29), we get

$$\boxed{\lambda_1\lambda_2\ldots\lambda_n = \det M} \tag{1.30}$$

That is, the product of the eigenvalues of a matrix is equal to the determinant of a matrix.

We may deduce from Eq. (1.30) that if M is a singular matrix (i.e., $\det M = 0$), then at least one of its eigenvalues must be zero. The converse is also true.

iv) **The sum of eigenvalues is equal to the trace of a matrix.**

To compute the sum of eigenvalues, we obtain the coefficient of λ^{n-1} from Eq. (1.28). It is given by

$$(-1)^{n-1}\lambda^{n-1}(M_{11}+M_{22}+\ldots+M_{nn}) = c_{n-1}\lambda^{n-1}$$

or

$$c_{n-1} = (-1)^{n-1}(M_{11}+M_{22}+\ldots+M_{nn}) \tag{1.31}$$

Comparing Eq. (1.31) with Eq. (1.27c) we conclude that

$$(\lambda_1+\lambda_2\ldots+\lambda_n) = M_{11}+M_{22}+\ldots+M_{nn} \tag{1.32}$$

$$= \text{Sum of the diagonal elements of } M.$$

The sum of the diagonal elements of a matrix M is called the **trace** of the matrix and is written as Tr M.

Hence,

$$\boxed{\lambda_1 + \lambda_2 + \ldots + \lambda_n = \operatorname{Tr} M} \qquad (1.33)$$

The sum of the eigenvalues equals the trace of the matrix.

You must memorise the results given in Eqs. (1.30) and (1.33). This will help you in checking whether or not your calculations of the eigenvalues of a given matrix are correct.

v) **If a matrix has real elements only, the eigenvalues of M are not necessarily real.**

Take the example $M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. You can calculate its eigenvalues. These are $\lambda = \pm i$. For a general matrix, there is no correlation between the eigenvalues and the elements of the matrix. If the matrix has some symmetry, there may exist a correlation and we shall explore this in the next unit.

vi) **For a *non-singular* matrix M with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, the eigenvalues of its inverse (M^{-1}) are given by**

$$\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_n}.$$

To prove this, we multiply the characteristic equation

$$\det(M - \lambda I) = 0$$

by $\det(M^{-1}/\lambda)$, and use the fact that $\det A \det B = \det AB$. Thus we get

$$0 = \det(M^{-1}/\lambda)\det(M - \lambda I) = \det\left(\frac{M^{-1}M}{\lambda} - \frac{M^{-1}}{\lambda} \times \lambda I\right) = \det\left(\frac{1}{\lambda} - M^{-1}\right)$$

or

$$\det(M^{-1} - \lambda^{-1} I) = 0$$

Hence we may conclude that λ^{-1} is an eigenvalue of M^{-1}, if λ is an eigenvalue of M.

vii) **If λ_1 and λ_2 are two distinct eigenvalues of M, then the corresponding eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are linearly independent.**

Let us first understand what is meant by linear independence of vectors. Two vectors $\mathbf{A}_1$ and $\mathbf{A}_2$ are said to be *linearly dependent* if one is a multiple of the other, i.e., $\mathbf{A}_1 = c\,\mathbf{A}_2$. Conversely, two vectors $\mathbf{A}_1$ and $\mathbf{A}_2$ are said to be **linearly independent** of each other if one is **not** a multiple of the other. In this case, it is impossible to satisfy the equation

$$\alpha\,\mathbf{A}_1 + \beta\,\mathbf{A}_2 = 0 \quad \text{except for } \alpha = 0, \beta = 0.$$

For, if α and β are non zero, this relation would imply

$$\mathbf{A}_2 = \left(-\frac{\alpha}{\beta}\right)\mathbf{A}_1$$

i.e., $\mathbf{A}_2$ is a multiple of $\mathbf{A}_1$.

We can now prove (vii) as follows:

Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be the eigenvectors of M belonging to eigenvalues λ_1 and λ_2:

$$M\mathbf{x}_1 = \lambda_1 \mathbf{x}_1 \tag{1.34a}$$

$$M\mathbf{x}_2 = \lambda_2 \mathbf{x}_2 \tag{1.34b}$$

We are given $\lambda_1 \neq \lambda_2$. We have to show that $\mathbf{x}_1$ is not a scalar multiple of $\mathbf{x}_2$, i.e., if we assume

$$\alpha\, \mathbf{x}_1 + \beta\, \mathbf{x}_2 = 0 \tag{1.34c}$$

then we have to show that $\alpha = 0$ and $\beta = 0$.

Let us assume that Eq. (1.34c) holds and multiply it by M from the left:

$$M(\alpha\, \mathbf{x}_1 + \beta\, \mathbf{x}_2) = 0$$

or

$$\alpha\, M\mathbf{x}_1 + \beta\, M\mathbf{x}_2 = 0$$

or

$$\alpha\, \lambda_1\, \mathbf{x}_1 + \beta\, \lambda_2\, \mathbf{x}_2 = 0 \qquad \text{(From Eqs. (1.34a and b)} \tag{1.34d}$$

From Eq. (1.34c), $\alpha\, \mathbf{x}_1 = -\beta\, \mathbf{x}_2$ and using this in Eq. (1.34d), we get

$$\lambda_1 (-\beta\, \mathbf{x}_2) + \beta\, \lambda_2\, \mathbf{x}_2 = 0$$

or

$$\beta\, (\lambda_2 - \lambda_1)\, \mathbf{x}_2 = 0$$

As $\lambda_2 \neq \lambda_1$, it yields

$$\beta\, \mathbf{x}_2 = 0 \tag{1.34e}$$

Clearly $\mathbf{x}_2$ is a non-trivial eigenvector and hence Eq. (1.34e) can hold only if $\beta = 0$. With $\beta = 0$, Eq. (1.34c) again shows $\alpha\, \mathbf{x}_1 = 0$ and as $\mathbf{x}_1$ is a non-trivial eigenvector, α must be zero. We have thus shown that if $\lambda_1 \neq \lambda_2$, the corresponding eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are not proportional to each other, i.e., they are linearly independent.

The concept of linear independence can be extended to column vectors $\mathbf{u}_1, \ldots \mathbf{u}_k$. These are said to be linearly independent if

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_k\mathbf{u}_k = 0 \quad \text{implies} \quad c_1 = c_2 = \ldots = c_k = 0. \tag{1.35a}$$

In compact notation, we write this result as

$$\sum_{i=1}^{k} c_i\mathbf{u}_i = 0 \quad \text{implies} \quad c_i = 0, \quad \text{for} \quad i = 1, \ldots, k \tag{1.35b}$$

viii) **The number of linearly independent eigenvectors of M belonging to a given eigenvalue λ_1 is $(n - k)$ where n is the dimension of M and k is the rank of the matrix $(M - \lambda_1 I)$.**

We shall not prove this statement.

$f(A)$ and the corresponding eigenvalue is $f(\lambda)$. In particular, cA and $(A - cI)$ have eigenvalues $c\lambda$ and $(\lambda - c)$ respectively, where c is a scalar.

As an example, let us take

$$f(A) = A^3 + 4A^2 + 3A + I.$$

If $A\mathbf{x} = \lambda\mathbf{x}$, it follows that

$$\begin{aligned} f(A)\mathbf{x} &= A^3\mathbf{x} + 4A^2\mathbf{x} + 3A\mathbf{x} + I\mathbf{x} \\ &= \lambda^3\mathbf{x} + 4\lambda^2\mathbf{x} + 3\lambda\mathbf{x} + I\mathbf{x} \\ &= (\lambda^3 + 4\lambda^2 + 3\lambda + I)\mathbf{x} \\ &= f(\lambda)\mathbf{x} \end{aligned}$$

You should end this session with a couple of exercises to check whether you have grasped these ideas.

SAQ 6

Spend 10 min

(a) Calculate the eigenvalues of M^{-1} where

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

(b) Determine the eigenvalues and eigenvectors of the matrix

$$B = \begin{pmatrix} 0 & 1 & 2 \\ -2 & 3 & 4 \\ 1 & -1 & -1 \end{pmatrix}$$

We now come to an important theorem concerning the eigenvalue problem.

1.4 CAYLEY-HAMILTON THEOREM

Cayley-Hamilton theorem is a basic theorem in matrix algebra. It states that

Every matrix satisfies its own characteristic equation.

Before taking up the proof of this theorem, let us illustrate it with the help of an example:

Consider $M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Its characteristic equation is given by

$$\det(M-\lambda I)=\begin{vmatrix}1-\lambda & 2\\ 3 & 4-\lambda\end{vmatrix}=0$$

or

$$\lambda^2-5\lambda-2=0.$$

The Cayley-Hamilton theorem states that M itself obeys the above equation with λ replaced by M, i.e., $M^2-5M-2I=0$. Let us verify this result

$$M^2-5M-2I=\begin{pmatrix}7 & 10\\ 15 & 22\end{pmatrix}-\begin{pmatrix}5 & 10\\ 15 & 20\end{pmatrix}-\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix}$$

$$=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}=0.$$

The product of two matrices is given by

$$\begin{pmatrix}a_1 & b_1\\ a_2 & b_2\end{pmatrix}\begin{pmatrix}A_1 & B_1\\ A_2 & B_2\end{pmatrix}=\begin{pmatrix}a_1A_1+b_1A_2 & a_1B_1+b_1B_2\\ a_2A_1+b_2A_2 & a_2B_1+b_2B_2\end{pmatrix}$$

The Cayley-Hamilton theorem generalises this result and asserts that this is true for every $n\times n$ matrix.

Let us now prove the Cayley-Hamilton theorem. We are giving the proof of this theorem for the sake of interest only. **You will not be examined for this.**

Let the characteristic equation of a matrix A be

$$c_n\lambda^n+c_{n-1}\lambda^{n-1}+\ldots+c_1\lambda+c_0=0.$$

Then according to the Cayley-Hamilton theorem, the matrix A satisfies the equation

$$c_nA^n+c_{n-1}A^{n-1}+\ldots+c_1A+c_0I=0 \qquad (1.36a)$$

or the equation in the factorised form

$$(A-\lambda_1I)(A-\lambda_2I)..(A-\lambda_nI)=0 \qquad (1.36b)$$

Let us define B as the adjoint of $(A-\lambda I)$. To prove Cayley-Hamilton theorem, we make use of the identity

$$(A-\lambda I)B=\det(A-\lambda I)I \qquad (1.37)$$

Since $A-\lambda I$ is a matrix of order n, each element of B will be a polynomial in λ of degree $n-1$. B can, therefore, be written as

$$B(\lambda)=\lambda^{n-1}B_{n-1}+\lambda^{n-2}B_{n-2}+\ldots+B_0 \qquad (1.38)$$

where each B_i is a matrix of order n. Then substituting for B from Eq. (1.38) and $\det(A-\lambda I)$ in Eq. (1.37) we get

$$(A-\lambda I)\left(B_{n-1}\lambda^{n-1}+B_{n-2}\lambda^{n-2}+\ldots+B_0\right)=\lambda^nc_nI+c_{n-1}\lambda^{n-1}I+\ldots+c_1\lambda I+c_0I \qquad (1.39)$$

Cofactor and Adjoint of a matrix

Let A be an $n\times n$ matrix. Consider the square matrix A' of order $n-1$ obtained from A by deleting the i^{th} row and j^{th} column of A. The **cofactor** of the element a_{ij}, denoted by A^{ij}, is defined as

$$A^{ij}=(-1)^{i+j}\det A'$$

where $\det A'$ is called the **minor** of the element a_{ij} in $\det A$. The *matrix of cofactors* is written as

$$A_c=\begin{bmatrix}A^{11} & A^{12} & \ldots & A^{1n}\\ A^{21} & A^{22} & \ldots & A^{2n}\\ \vdots & & \ldots & \\ A^{n1} & A^{n2} & \ldots & A^{nn}\end{bmatrix}$$

Then the transpose of this matrix, obtained by interchanging the rows and columns is called the **adjoint** of the matrix A:

$$A_c^T=\begin{bmatrix}A^{11} & A^{21} & \ldots & A^{n1}\\ A^{12} & A^{22} & \ldots & A^{n2}\\ \vdots & & \ldots & \\ A^{1n} & A^{2n} & \ldots & A^{nn}\end{bmatrix}$$

The coefficient of any power of λ must, of course, be the same on both sides. So we get

$$-B_{n-1} = c_n I$$

$$-B_{n-2} + AB_{n-1} = c_{n-1} I$$

$$\vdots$$

$$-B_0 + AB_1 = c_1 I$$

$$AB_0 = c_0 I$$

Multiplying the first equation on the left by A^n, the second by A^{n-1}, and the last by I, and adding them all, we get

$$\left\{-A^n B_{n-1} + \left(-A^{n-1} B_{n-2} + A^n B_{n-1}\right) + \left(-A^{n-2} B_{n-3} + A^{n-1} B_{n-2}\right) + \ldots + \left(-AB_0 + A^2 B_1\right) + AB_0\right.$$

$$= c_n A^n + c_{n-1} A^{n-1} + \ldots + c_1 A + c_0 I.$$

You should note that the terms on the left hand side cancel out in pairs making the total zero. On the right hand side, we have just the characteristic polynomial with A replacing λ. Hence

$$c_n A^n + c_{n-1} A^{n-1} + \ldots + c_0 I = 0$$

The Cayley-Hamilton theorem is thus proved.

Suppose there are some repeated eigenvalues of a matrix, i.e., suppose that of the eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n$ of A, m_a are equal to each other with a common value λ_a say, m_b are equal to λ_b etc. In that case, the Cayley-Hamilton theorem could be represented as

$$(A - \lambda_a I)^{m_a} (A - \lambda_b I)^{m_b} \ldots = 0 \tag{1.40}$$

Eq. (1.40) clearly follows from Eq. (1.36b). Before studying further, you should apply the Cayley-Hamilton Theorem to a simple example.

SAQ 7

Spend 10 *min*

Verify the Cayley-Hamilton theorem for the rotation matrix.

$$M = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

Application of the Cayley-Hamilton Theorem

An important application of the Cayley-Hamilton theorem is to **compute the inverse of a non-singular matrix**. Let A be a non-singular matrix. By the Cayley-Hamilton theorem, it satisfies Eq. (1.36a). Multiplying by A^{-1} from the left we have

$$c_n A^{n-1} + c_{n-1} A^{n-2} + \ldots + c_1 I + c_0 A^{-1} = 0.$$

From this we get

$$A^{-1} = -\frac{1}{c_0}\left(c_n A^{n-1} + c_{n-1} A^{n-2} + \ldots + c_1 I\right) \tag{1.41}$$

Clearly we must have $c_0 \neq 0$ for A^{-1} to exist. The advantage of this formula is that A^{-1} is expressed as a matrix polynomial in A. As an illustration, let us reconsider the matrix given in Example 4:

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

You recall that the characteristic polynomial of A is given by

$$|A - \lambda I| = (\lambda^3 - 7\lambda - 6) = 0$$

By the Cayley-Hamilton theorem we know that A will obey

$$A^3 - 7A - 6I = 0$$

Hence by Eq. (1.41), we can write

$$A^{-1} = \frac{1}{6}(A^2 - 7I)$$

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

Quite often (not always!) it happens that when A has repeated eigenvalues, it satisfies not only the Cayley-Hamilton theorem but in addition, a simpler equation of lower degree. The knowledge of such an equation is useful in many contexts. We shall devote the last section of the unit to a brief discussion of this aspect.

1.5 MINIMAL EQUATION AND ITS APPLICATIONS

The Cayley-Hamilton theorem shows that any $n \times n$ matrix A satisfies an nth degree equation (1.36a) or (1.36b). In certain cases, when A has repeated eigenvalues it is possible that A satisfies an equation of degree lower than n. Among all such algebraic equations, let the one which is of the lowest degree (say m) be

$$p_m A^m + p_{m-1} A^{m-1} + \ldots + p_0 I = 0 \qquad (1.42)$$

This equation is then called the **minimal equation** of A and the polynomial

$$p(\lambda) = p_m \lambda^m + p_{m-1}\lambda^{m-1} + \ldots + p_0$$

is called the **minimal polynomial** of the matrix A. The possibility of a minimal polynomial of degree $m < n$ arises only if A has atleast one repeated eigenvalue.

The converse of the above result is not true: The existence of repeated eigenvalues does not necessarily imply that $m < n$. Consider, for example, the following three matrices

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

All these matrices have the same characteristic polynomial, $C(\lambda) = (\lambda - 2)^3$. So $(M - 2I)^3 = 0$ for $M = A,B,C$ by the Cayley-Hamilton theorem. But the equations of *lowest degree satisfied* by A, B, C are $(A - 2I) = 0$, $(B - 2I)^2 = 0$, $(C - 2I)^3 = 0$. You can easily verify this. Let us now study a few other applications of the minimal equation.

Applications of the Minimal Equation

1. **To express the inverse of a matrix in terms of powers of the matrix.**

This result is straight forward and we have illustrated it earlier. Suppose the minimal equation of A is

$$p_m A^m + p_{m-1} A^{m-1} + ... + p_0 I = 0 \quad \text{with} \quad p_0 \neq 0. \tag{1.43}$$

Multiplying the equation by A^{-1} and re-arranging we get

$$A^{-1} = -\frac{1}{p_0}\left[p_m A^{m-1} + p_{m-1} A^{m-2} + ... + p_1 I\right] \tag{1.44}$$

Since $m < n$, this method is more efficient.

2. **To simplify functions of a matrix.**

We will illustrate this by the following example. Suppose A is any matrix having the minimal equation $A^2 = I$. Then $A^3 = A$, $A^4 = A^2 = I$, $A^5 = A^3 = A$ etc. Thus, all even powers of A are equal to I and all the odd powers are equal to A. Consider the exponential matrix defined by

$$\exp(i\alpha A) = 1 + i\alpha A + \frac{1}{2!}(i\alpha A)^2 + ... \tag{1.45}$$

For finite dimensional matrices, the series on the right hand side of Eq. (1.45) converges and the resulting expression defines exp ($i\,\alpha\,A$). In our case, on account of the relation $A^2 = I$, the right hand side of Eq. (1.45) can be simplified before being summed:

$$\exp(i\alpha A) = \left(1 - \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} + ...\right)I + i\left(\alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} + ...\right)A$$

The terms in the brackets are the well-known series for cos α and sin α. Hence,

$$\exp(i\alpha A) = (\cos\alpha)I + i(\sin\alpha)A \qquad \text{when} \quad A^2 = I \tag{1.46}$$

In Quantum Mechanics, the effect of a rotation on the states of a particle of spin $\frac{1}{2}$ is represented by a matrix of this kind. If the rotation is through an angle θ about an axis in the direction of the unit vector $\hat{\mathbf{n}}$, the rotation matrix is

$$A = \exp\left[\frac{1}{2} i\theta\sigma.\hat{\mathbf{n}}\right]$$

where $\sigma_x, \sigma_y, \sigma_z$ are the Pauli spin matrices given by:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

You can easily verify that

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I, \; \sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z$$

so that

$$(\sigma.\hat{\mathbf{n}})^2 = I\,\hat{\mathbf{n}}^2 = I$$

Therefore, the reduction of exp $(i\,\alpha\, A)$ given above applies in this case and we have

$$\exp\left(i\frac{\theta}{2}\sigma.\hat{\mathbf{n}}\right) = \left(\cos\frac{\theta}{2}\right)I + i\left(\sin\frac{\theta}{2}\right)(\sigma.\hat{\mathbf{n}})$$

With this we come to the end of this unit. You must have realised that the statements and theorems of this unit do not assume any special symmetry property for the matrices. If a matrix possesses a symmetry, it is possible to make more definitive statements on their eigenvalues and eigenvectors. The next unit is addressed to this problem. Let us now summarise what you have studied in this unit.

1.6 SUMMARY

- The **eigenvalue equation** for a square matrix A is given by

  $$A\,\mathbf{x} = \lambda\,\mathbf{x}$$

  The column matrix **x** which satisfies this equation is called an **eigenvector** of A and the scalar λ is called the corresponding **eigenvalue.**

- The equation

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

  is called the **characteristic equation** for A.

- The **eigenvalues** of a matrix A are given by the solutions of the equation

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

  where $\det(A - \lambda I)$ is called the **secular determinant.**

- An $n \times n$ matrix has r eigenvalues such that $0 \le r \le n$. The set of all eigenvalues of a matrix is called the **eigenvalue spectrum** of the matrix.

- The **eigenvectors** of a matrix corresponding to a set of eigenvalues are obtained by solving the set of linear algebraic equations given by $(A - \lambda I)\,\mathbf{x} = 0$ for each **eigenvalue.**

- For a **diagonal matrix**, the eigenvalues are given by its diagonal elements.

- The **product of eigenvalues** is equal to the determinant of a matrix:

  $$\lambda_1 \lambda_2 \ldots \lambda_n = \det M$$

- The **sum of eigenvalues** is equal to the **trace of a matrix**:

$$\lambda_1 + \lambda_2 + \ldots + \lambda_n = \text{Tr}\, M$$

where Tr M is the sum of the diagonal elements of the matrix.

- For a non-singular matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ the **eigenvalues of its inverse** $\left(M^{-1}\right)$ are given by

$$\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_n}$$

- The eigenvectors corresponding to distinct eigenvalues of a matrix are **linearly independent**.

- If **x** is an eigenvector of a matrix A, then it is also an eigenvector of $f(A)$ where f is any function of A, with the eigenvalue $f(\lambda)$.

- **Cayley-Hamilton Theorem**: Every matrix satisfies its own characteristic equation:

$$c_n A^n + c_{n-1} A^{n-1} + \ldots + c_1 A + c_0 I = 0$$

From the Cayley-Hamilton theorem, the **inverse of a matrix** is given by

$$A^{-1} = -\frac{1}{c_0}\left(c_n A^{n-1} + c_{n-1} A^{n-2} + \ldots + c_1 I\right)$$

- When an $n \times n$ matrix A has repeated eigenvalues, A can satisfy an equation of a lower degree (say m) than n. Such an equation is called the **minimal equation** for A:

$$p_m A^m + p_{m-1} A^{m-1} + \ldots + p_0 I = 0$$

The polynomial $p(\lambda) = p_m \lambda^m + p_{m-1} \lambda^{m-1} + \ldots + p_0$ is called the **minimal polynomial** of A.

1.7 TERMINAL QUESTIONS

Spend 30 *min*

1. Determine the eigenvalues and eigenvectors for the matrix

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

2. Obtain the inverse of matrix

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

3. If A is such that $A^3 = A$, show that

$$\cos(\alpha A) = I - (1 - \cos\alpha)A^2$$

$$\sin(\alpha A) = A\sin\alpha$$

1.8 SOLUTIONS AND ANSWERS

Self-assessment Questions

1. $C = BA$

$$= \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

$$= \begin{pmatrix} \cos\phi\cos\theta - \sin\phi\sin\theta & \cos\phi\sin\theta + \sin\phi\cos\theta \\ -\sin\phi\cos\theta - \cos\phi\sin\theta & -\sin\phi\sin\theta + \cos\phi\cos\theta \end{pmatrix}$$

$$= \begin{pmatrix} \cos(\theta+\phi) & \sin(\theta+\phi) \\ -\sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}$$

Rule of Matrix Multiplication:

If $C = AB$

$$(C)_{ij} = \sum_{k=1}^{n} (A)_{ik}(B)_{kj}$$

where C, A and B are square matrices of order n. Thus

$$C_{11} = A_{11}B_{11} + A_{12}B_{12} + A_{13}B_{31} + \ldots + A_{1n}B_{n1}$$

and so on.

Alternatively you would need to follow the method given below:

$$x'' = x'\cos\phi + y'\sin\phi$$

$$y'' = -x'\sin\phi + y'\cos\phi$$

$$x' = x\cos\theta + y\sin\theta$$

$$y' = -x\sin\theta + y\cos\theta$$

Substituting for x' and y' we obtain

$$x'' = (x\cos\theta + y\sin\theta)\cos\phi + (-x\sin\theta + y\cos\theta)\sin\phi$$

$$y'' = -(x\cos\theta + y\sin\theta)\sin\phi + (-x\sin\theta + y\cos\theta)\cos\phi$$

or

$$x'' = x(\cos\theta\cos\phi - \sin\theta\sin\phi) + y(\sin\theta\cos\phi + \cos\theta\sin\phi)$$

$$y'' = -x(\cos\theta\sin\phi + \sin\theta\cos\phi) + y(-\sin\theta\sin\phi + \cos\theta\cos\phi)$$

or

$$x'' = x\cos(\theta+\phi) + y\sin(\theta+\phi)$$

$$y'' = -x\sin(\theta+\phi) + y\cos(\theta+\phi)$$

or

$$C = \begin{pmatrix} \cos(\theta+\phi) & \sin(\theta+\phi) \\ -\sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}$$

which is a far lengthier method and prone to computational errors.

2. $$M\mathbf{x} = \lambda\,\mathbf{x}$$

or

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

or

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

or

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

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

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

3. (a) The secular equation is

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

or

$$\lambda^2 - 1 = 0$$

or

$$\lambda = \pm 1$$

The eigenvalues of M are $+1$ and -1.

(b) The secular equation is

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

or

$$(1-\lambda)\,(2-\lambda)-6=0$$

or

$$2+\lambda^2-3\lambda-6=0$$

or

$$\lambda^2-3\lambda-4=0$$

or

$$\lambda^2-4\lambda+\lambda-4=0$$

or

$$(\lambda-4)\,(\lambda+1)=0$$

Thus the eigenvalues of A are +4 and −1.

4. The eigenvalues of A are $\lambda = +4$ and $\lambda = -1$. For $\lambda = +4$, Eq. (1.13) becomes

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

or

$$-3x_1+2x_2=0$$

$$3x_1-2x_2=0$$

This yields

$$3x_1=2x_2$$

or

$$x_2=\frac{3}{2}x_1$$

The eigenvector corresponding to $\lambda = +4$ is

$$x_1\begin{pmatrix} 1 \\ 3/2 \end{pmatrix}$$

For $\lambda = -1$, we have

$$\begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

or

$$2x_1 + 2x_2 = 0$$

$$3x_1 + 3x_2 = 0$$

which yields $x_1 = -x_2$. Thus the eigenvector corresponding to $\lambda = -1$ is $x_1\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

5. Let the characteristic frequency at which all three masses vibrate be ω. Since the system obeys Hooke's law, we can assume the solutions to be

$$x_i = x_{i0}e^{i\omega t}, \qquad i = 1,2,3.$$

Substituting these solutions into the equations of motion we get the following eigenvalue equation after dividing out the common factor $e^{i\omega t}$

$$\begin{pmatrix} \frac{k}{M} & -\frac{k}{M} & 0 \\ -\frac{k}{m} & \frac{2k}{m} & -\frac{k}{m} \\ 0 & -\frac{k}{M} & \frac{k}{M} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = +\omega^2\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

The secular equation is

$$\begin{vmatrix} \frac{k}{M}-\omega^2 & -\frac{k}{M} & 0 \\ -\frac{k}{m} & \frac{2k}{m}-\omega^2 & -\frac{k}{m} \\ 0 & -\frac{k}{M} & \frac{k}{M}-\omega^2 \end{vmatrix} = 0$$

This yields

$$\omega^2\left(\frac{k}{M}-\omega^2\right)\left(\omega^2 - \frac{2k}{m} - \frac{k}{M}\right) = 0$$

The characteristic frequencies are

$$\omega_1 = 0, \ \omega_2 = \sqrt{\frac{k}{M}} \quad \text{and} \quad \omega_3 = \sqrt{\frac{k}{M} + \frac{2k}{m}}.$$

Notice that these are all real.

6. (a) The characteristic equation of the matrix M is

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

or

$$\lambda^3 - 7\lambda - 6 = 0$$

This has roots $\lambda_1 = -2$, $\lambda_2 = -1$ and $\lambda_3 = 3$. These are therefore the eigenvalue of M. The eigenvalues of M^{-1} are, therefore

$$\lambda_1' = -\frac{1}{2}, \quad \lambda_2' = -1 \quad \text{and} \quad \lambda_3' = \frac{1}{3}$$

since M is a non-singular matrix.

(b) The characteristic equation of B is

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

The eigenvalues are $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 1$.

For $\lambda = 0$, we get the trivial solution $\mathbf{x} = \begin{pmatrix} 0 \\ 0. \\ 0 \end{pmatrix}$

For $\lambda = 1$, the eigenvalue equation $(B - \lambda I)\,\mathbf{x} = 0$ becomes

$$\begin{pmatrix} -1 & 1 & 2 \\ -2 & 2 & 4 \\ 1 & -1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

All three rows of the matrix yield the same equation

$$x_1 - x_2 - 2x_3 = 0$$

This equation permits any two of three parameters x_1, x_2, x_3 to be arbitrarily chosen; the third then gets determined by these two. We can take, for instance, $x_1 = x_2 + 2x_3$, with x_2 and x_3 left arbitrary. Then eigenvectors of the matrix B which belong to the eigenvalue 1 can be written as

$$\begin{pmatrix} x_2 + 2x_3 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

7. The characteristic equation of M is

$$\begin{vmatrix} \cos\theta - \lambda & \sin\theta \\ -\sin\theta & \cos\theta - \lambda \end{vmatrix} = 0$$

or

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

or

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

Thus M should satisfy the equation

$$M^2 - (2\cos\theta)M + (\cos^2\theta + \sin^2\theta)I = 0$$

$$\begin{pmatrix} \cos^2\theta - \sin^2\theta & 2\cos\theta\sin\theta \\ -2\cos\theta\sin\theta & \cos^2\theta - \sin^2\theta \end{pmatrix} - \begin{pmatrix} 2\cos^2\theta & 2\cos\theta\sin\theta \\ -2\cos\theta\sin\theta & 2\cos^2\theta \end{pmatrix}$$

$$+ \begin{pmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{pmatrix}$$

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

Thus, the Cayley-Hamilton theorem is verified.

Terminal Questions

1. The characteristic equation of B is

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

or

$$\lambda^3 = 0$$

The eigenvalues are $\lambda_1 = 0$, $\lambda_2 = 0$ and $\lambda_3 = 0$. The eigenvectors are given by the equation

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

This equation yields

$$x_2 = 0, \quad x_3 = 0$$

Hence

$$\mathbf{x} = \begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix}$$

2. The characteristic equation of A is

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

or

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

or

$$\lambda^4 - 2\lambda^2 + 1 = 0$$

The inverse of A is given by

$$A^{-1} = -1\left(A^3 - 2A\right)$$
$$= -A\left(A^2 - 2I\right)$$

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

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = I \qquad (*)$$

$$A^2 - 2I = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$-A\left(A^2 - 2I\right) = -\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Thus

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

Notice from (*) that A satisfies the equation $A^2 = I$ which yields $A^{-1} = A$.

3. $$\cos(\alpha A) = \left(I - \frac{\alpha^2 A^2}{2!} + \frac{\alpha^4 A^4}{4!} - \frac{\alpha^6 A^6}{6!} + \ldots \right)$$

$$= \left(I - \frac{\alpha^2 A^2}{2!} + \frac{\alpha^4 A^2}{4!} - \frac{\alpha^6 A^2}{6!} + \ldots \right)$$

since $A^3 = A$.

$$\cos(\alpha A) = I - A^2 + A^2 \left(I - \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} - \frac{\alpha^6}{6!} + \ldots \right)$$

$$= I - A^2 + A^2 \cos\alpha$$

$$= I - (1 - \cos\alpha) A^2$$

$$\sin(\alpha A) = \alpha A - \frac{\alpha^3 A^3}{3!} + \frac{\alpha^5 A^5}{5!} + \ldots$$

$$= \alpha A - \frac{\alpha^3 A}{3!} + \frac{\alpha^5 A}{5!} - \ldots \qquad (\text{since } A^3 = A)$$

$$= A \left(\alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} - \ldots \right)$$

$$= A \sin\alpha.$$