UNIT 3 DIAGONALIZATION OF MATRICES

Structure

3.1 INTRODUCTION

In the preceding two Units, you have learnt how to solve the eigenvalue problems for matrices, with particular reference to special matrices like the hermitian and unitary matrices. In Unit 2, you have also learnt that the eigenvalue problem becomes very easy to solve if the matrix involved is a diagonal matrix. Therefore, in this unit we shall discuss the **diagonalization** of a matrix, i.e., its reduction to a diagonal form. In order to reduce any matrix to a diagonal form, we make use of **similarity transformations.** So to begin with in Sec. 3.2 we shall explain what is meant by a similarity transformation. Then you will learn how to diagonalize a matrix. You will also learn under what conditions a matrix is diagonalizable. Finally, you will study some important applications of matrix diagonalization, especially to quadratic forms and physical situations in Sec. 3.3.

With this we come to an end on our discussion of matrices which focussed essentially on solving the eigenvalue problem in physics. In the next unit we will introduce you to another important mathematical construct – the tensor which is used in many areas of classical physics, and the theories of relativity.

Objectives

After studying this unit, you should be able to:

- define similarity, unitary and orthogonal transformations;
- reduce a matrix to its diagonal form;
- identify diagonalizable matrices; and
- solve problems involving diagonalization of matrices.

3.2 MATRIX DIAGONALIZATION

In Unit 2, you have learnt that the eigenvalues of a diagonal matrix are simply its diagonal elements. Now suppose we are able to recast a typical eigenvalue problem $A\mathbf{x} = \lambda\mathbf{x}$, so that A is transformed into a diagonal matrix. Then our calculations would become very simple. You may well ask: Is this possible? Yes, it is possible through what is known as a **similarity transformation**. So as a first step we will explain what this transformation is.

3.2.1 Similarity Transformations

Recall that, in physics, many a times, we introduce a change of variables to rewrite the original problem in a simpler form whose solution is easier to obtain than that of the original problem. This is precisely what we are going to do now. Suppose we have to solve the eigenvalue problem

$$A\mathbf{x} = \lambda\mathbf{x} \tag{3.1}$$

where A is a square matrix of order n and $\mathbf{x}$ the unknown eigenvector. We now introduce a change of variables or a transformation

$$\mathbf{x} = P\,\mathbf{y} \tag{3.2}$$

such that the inverse of P exists. Then the eigenvalue problem (3.1) takes the form

$$AP\,\mathbf{y} = \lambda\, P\,\mathbf{y}$$

Multiplying by P^{-1} from the left, we get

$$P^{-1}AP\,\mathbf{y} = \lambda\mathbf{y} \tag{3.3a}$$

or

$$B\,\mathbf{y} = \lambda\mathbf{y} \tag{3.3b}$$

where

$$B = P^{-1}AP \tag{3.4a}$$

Now, if for an eigenvalue problem, we are able to determine P such that the matrix $B = P^{-1}A\,P$ is diagonal, and the eigenvalues of B and A are the same, then we can obtain the eigenvalues of A very easily. This is possible under a **similarity transformation**.

By definition, the matrices A and B of Eq. (3.4a) are said to be related by a **similarity transformation** and B is said to be **similar** to A if there exists an invertible matrix P such that Eq. (3.4a) is satisfied.

The inverse transformation of Eq. (3.4a) is

$$A = PBP^{-1} \tag{3.4b}$$

This is also a similarity transformation.

What does this mean? This means that the matrices A and B transform the vectors in the same manner but in different coordinate systems. If the matrix of transformation from $\mathbf{x}$ to $\lambda\mathbf{x}$ in a given coordinate system (say S) is A, then the matrix of transformation for the **same two** vectors $\mathbf{x}$ and $\lambda\mathbf{x}$ in the coordinate system $\mathbf{y} = P\mathbf{x}$ (say S') is $B = P^{-1}A\,P$.

If B is similar to A, we write $B \sim A$.

Now we list some basic properties of similar matrices:

1. **A matrix is similar to itself.**

$$A \sim A \tag{3.5a}$$

2. **A is similar to B if and only if (iff) B is similar to A.**

$$A \sim B \quad \text{iff} \quad B \sim A \tag{3.5b}$$

3. **If A is similar to B and B is similar to C then A is similar to C.**

$$A \sim B \text{ and } B \sim C \text{ then } A \sim C \tag{3.5c}$$

An important property of similar matrices involves their eigenvalues.

4. **Similar matrices have the same characteristic polynomial and hence exactly the same eigenvalues, including multiplicities.**

To prove this property, suppose $A \sim B$, i.e.,

$$B = P\,A\,P^{-1}$$

for some invertible matrix P. It follows that

$$B-\lambda I = PAP^{-1}-\lambda I$$

$$= PAP^{-1}-\lambda PP^{-1}$$

$$= PAP^{-1}-P(\lambda I)P^{-1}$$

$$= P(A-\lambda I)P^{-1}$$

Hence

$$\det(B-\lambda I) = \det P \det(A-\lambda I)\det P^{-1}$$

But

$$\det(P)\det\left(P^{-1}\right) = \det\left(PP^{-1}\right) = \det(I) = 1$$

Therefore

$$\det(B-\lambda I) = \det\left(A-\lambda I\right) \tag{3.6}$$

Note, however, that two matrices may have the same characteristic polynomial and yet might not be similar.

5. **The eigenvectors of similar matrices are, generally, not the same but related by a coordinate transformation.**

   For example, the eigenvectors $\mathbf{x}$ of A are related to eigenvectors $\mathbf{y}$ by Eq. (3.2).

   If, in particular, the matrix P with which a similarity transformation is performed happens to be unitary, then the transformation is called a **unitary transformation**.

   Further, if P is a real, orthogonal matrix, the transformation is called an **orthogonal transformation**. We thus have, by definition,

Similarity transformation

$$B = P^{-1}AP,$$

or

$$A = PBP^{-1}, \qquad \text{if } P^{-1} \text{ exists}$$

Unitary transformation

$$B = P^{-1}AP, \text{ if } P^{-1} = P^{\dagger} \tag{3.7a}$$

Orthogonal transformation

$$B = P^{-1}AP, \text{ if } P^{-1} = P^{T} \tag{3.7b}$$

So our problem of solving Eq. (3.1) is now converted into determining a matrix P such that the matrix $P^{-1}AP$ is diagonal since the diagonal elements of $P^{-1}AP$ are the eigenvalues of A.

You may now like to attempt an SAQ based on the ideas presented thus far.

SAQ 1

Spend 10 *min*

(a) Show that $A \sim B$ given

$$A = \begin{pmatrix} 1 & -3 \\ 2 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 2 \\ -3 & 1 \end{pmatrix} \text{ and } P = P^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

(b) Prove Eqs. (3.5b) and (3.5c).

The process of obtaining $P^{-1}AP$, given the transformation matrix P, such that it is of a diagonal form is called the diagonalization of a matrix. You will now learn how to calculate such a matrix P and reduce a given matrix to its diagonal form.

3.2.2 Reducing a Matrix to a Diagonal Form

Let A be an $n \times n$ matrix with elements a_{ij}. Let A have n distinct eigenvalues λ_i and n linearly independent eigenvectors $\mathbf{x}_i$ corresponding to λ_i. Thus

$$A\,\mathbf{x}_i = \lambda_i \mathbf{x}_i \tag{3.8a}$$

Let us denote the eigenvectors $\mathbf{x}_i$ by column matrices of the form $(x_{1i}, x_{2i}, x_{3i}, \ldots, x_{ni})^T$. Then the eigenvalue equation (3.8a) can be written in the matrix form as

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{ni} \end{pmatrix} = \lambda_i \begin{pmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{ni} \end{pmatrix}, \quad i = 1, \ldots, n \tag{3.8b}$$

Thus, there are n equations of the form (3.8b) corresponding to each λ_i and $\mathbf{x}_i$.

We can write the j^{th} equation in (3.8b) as

$$\sum_{k=1}^{n} a_{jk} x_{ki} = \lambda_i\, x_{ji} \tag{3.8c}$$

Let us now construct a matrix P of order $n \times n$ whose columns are the vectors $\mathbf{x}_i$. Thus

$$P = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_n)^T$$

Since each vector $\mathbf{x}_i$ can be written as a column vector of the form of Eq. (3.8b), we have

$$P = \begin{pmatrix} x_{11} & x_{1i} & \cdots & x_{1n} \\ x_{21} & x_{2i} & \cdots & x_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{ni} & \cdots & x_{nn} \end{pmatrix} \tag{3.9a}$$

or

$$(P)_{ji} = x_{ji} \tag{3.9b}$$

We now make use of the following result: If the vectors $\mathbf{x}_i$ are linearly independent, the matrix P is non-singular and P^{-1} exists.

Then we can prove that

The matrix $P^{-1}AP$ is diagonal and its diagonal elements are the eigenvalues of A.

To prove this statement, we define a diagonal matrix M whose diagonal elements are λ_i.

$$M = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & & & \cdots & \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{3.10}$$

We have to prove that

$$P^{-1}AP = M \tag{3.11a}$$

Multiplying from the left by P, we get

$$AP = PM \tag{3.11b}$$

To prove Eq. (3.11b) we consider the ji^{th} element of the left hand side

$$(AP)_{ji} = \sum_{k=1}^{n} (A)_{jk} (P)_{ki}$$

$$= \sum_{k=1}^{n} a_{jk}\, x_{ki}\,, \qquad \text{using Eq. (3.9b).}$$

We can now use Eq. (3.8c) to write

$$(AP)_{ji} = \lambda_i\, x_{ji} \tag{3.11c}$$

Similarly we can write

$$(PM)_{ji} = \sum_{k=1}^{n} (P)_{jk} (M)_{ki}$$

$$= \sum_{k=1}^{n} x_{jk}\, \lambda_i\, \delta_{ki} \quad , \text{ using Eq. (3.9b).}$$

Since M is a diagonal matrix

$$(M)_{ij} = \lambda_i \delta_{ij}$$

or

$$(PM)_{ji} = \lambda_i x_{ji} \tag{3.11d}$$

$$\equiv (AP)_{ji}$$

So we have proved Eq. (3.11b) and hence Eq. (3.11a). In the process we have found the method of diagonalizing a matrix. We can sum it up as follows:

Diagnonalization of a matrix A of order n

- Given an $n \times n$ matrix A having n distinct eigenvalues λ_i and n linearly independent eigenvectors $\mathbf{x}_i$:

$$A\,\mathbf{x}_i = \lambda_i\,\mathbf{x}_i,$$

construct an $n \times n$ matrix P whose columns are the eigenvectors $\mathbf{x}_i$:

$$P = (\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n)$$

$$(P)_{ji} = x_{ji}$$

- The matrix $P^{-1}AP$ is similar to A. It is a **diagonal** matrix whose diagonal elements are the eigenvalues of A.

When we obtain $P^{-1}AP$ given A, we say that we have **diagonalized** A by a similarity transformation.

We would like to explain these ideas further with the help of a simple example.

Example 1

Diagonalize the matrix

$$A = \begin{pmatrix} -17 & 30 \\ -10 & 18 \end{pmatrix}$$

Solution

Using the methods studied in Unit 2, you can verify that the two distinct eigenvalues of A are $\lambda_1 = -2$ and $\lambda_2 = 3$ and the corresponding eigenvectors are

$$\mathbf{x}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

We now construct a matrix $P = (\mathbf{x}_1, \mathbf{x}_2)$:

$$P = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$$

You can determine its inverse to be

$$P^{-1} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$$

Then we have

$$P^{-1}AP = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} -17 & 30 \\ -10 & 18 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} -4 & 9 \\ -2 & 6 \end{pmatrix}$$

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

Thus $P^{-1}AP$ is a diagonal matrix whose diagonal elements are the eigenvalues of A.

To understand this method better you should also solve the following SAQ.

SAQ 2

Spend 15 *min*

Using the process outlined in Sec. 3.2.2, diagonalize the matrix

$$M = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}$$

In Unit 2, you have learnt about real, symmetric and hermitian matrices. The diagonalization of such matrices finds important applications about which you will study in the next section. Here we would like to state some important results (without proof) concerning diagonalization of these matrices.

1. **A hermitian matrix can be diagonalised via a unitary transformation.**
2. **A real, symmetric matrix can be diagonalised by a real, orthogonal matrix.**

Recall that an $n \times n$ real, symmetric matrix has a set of n orthonormal eigenvectors. Thus the matrix P of eigenvectors is orthogonal and then the calculation of P^{-1} becomes very simple: $P^{-1} = P^T$. Let us demonstrate this idea with the help of an example.

Example 2: Diagonalization of (a) a hermitian and (b) a real, symmetric matrix.

(a) The matrix

$$A = \begin{pmatrix} 2 & 3-3i \\ 3+3i & 5 \end{pmatrix}$$

is a hermitian matrix. You can verify that its eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 8$ and the corresponding orthonormal eigenvectors are

$$\mathbf{x}_1 = \frac{1}{\sqrt{3}}\begin{pmatrix} -1+i \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1+i \end{pmatrix}$$

Thus, the diagnoalizing matrix U can be written as

$$U = \frac{1}{\sqrt{3}}\begin{pmatrix} -1+i & 1 \\ 1 & 1+i \end{pmatrix}$$

Now you can easily see that U is a unitary matrix. Therefore $U^{-1} = U^\dagger$:

$$U^{-1} = \frac{1}{\sqrt{3}}\begin{pmatrix} -1-i & 1 \\ 1 & 1-i \end{pmatrix}$$

and

$$U^{-1}AU \equiv U^{\dagger}AU = \frac{1}{\sqrt{3}}\begin{pmatrix} -1-i & 1 \\ 1 & 1-i \end{pmatrix}\begin{pmatrix} 2 & 3-3i \\ 3+3i & 5 \end{pmatrix}\frac{1}{\sqrt{3}}\begin{pmatrix} -1+i & 1 \\ 1 & 1+i \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 8 \end{pmatrix}$$

which is a diagonal matrix with the eigenvalues as its diagonal elements.

(b) As a second example, let us consider the matrix

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

which is a real, symmetric matrix. You can easily verify that it has eigenvalues $\lambda_1 = \lambda_2 = -3$ and $\lambda_3 = 6$ with the corresponding orthonormal set of eigenvectors:

$$\mathbf{x}_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{x}_2 = \frac{1}{3\sqrt{5}}\begin{pmatrix} 4 \\ 2 \\ 5 \end{pmatrix}, \quad \mathbf{x}_3 = \frac{1}{3}\begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$$

The diagonalizing matrix P is therefore given by

$$P = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \begin{pmatrix} \frac{-1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} & \frac{2}{3} \\ \frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} & \frac{-1}{3} \\ 0 & \frac{5}{3\sqrt{5}} & \frac{2}{3} \end{pmatrix}$$

Note that P is orthogonal since $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ are orthogonal and

$$P^{-1}AP = P^{T}AP = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

All this while, did you notice that it is rather easy to find the matrix $B = P^{-1}AP$? We only need to solve the characteristic equation of A. Then B is a matrix with the diagonal elements given by the eigenvalues of A, with all other elements being zero **provided A can be diagonalized**. So it is not always necessary to determine P and do the lengthy calculation of finding $P^{-1}AP$. You may now ask: Is it possible to know whether A can be diagonalized or not? This is the question we are going to address now.

3.2.3 Diagonalizable Matrices

As we have seen from the definition in Sec.3.2.1, first and foremost, a matrix A can be diagonalized if A is similar to a diagonal matrix.

We now state without proof, the situations in which a matrix is diagonalizable.

1. **An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.**

If A has n linearly independent eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n$ and $P = (\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n)$ then $P^{-1}AP = B$ is a diagonal matrix. The j^{th} diagonal element of B is equal to the j^{th} eigenvalue of A.

Now a **sufficient condition** for A to have n linearly independent vectors is that it has n distinct eigenvalues. Thus we have the following result.

2. **If an $n \times n$ matrix has n distinct eigenvalues, then it is diagonalizable**, since in this case it will possess n linearly independent eigenvectors. In the cases where an eigenvalue has multiplicity r, and there are still n linearly independent eigenvectors, the matrix is diagonalizable.

3. **Every symmetric matrix A is diagonalizable, whether or not its eigenvalues are distinct**. The matrix P for which $P^{-1}AP$ is a diagonal matrix, can be chosen to be orthogonal.

4. **Every hermitian matrix can be diagonalized by a unitary matrix.**

5. **A matrix of order n is diagonalizable** (or has a set of n linearly independent eigenvectors) **if and only if its minimal polynomial has distinct roots.** In other words, an $n \times n$ matrix is diagonalizable if and only if its minimal polynomial is of the form

$$P(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \ldots (\lambda_l - \lambda)$$

where λ_i $(i = 1, 2, \ldots l)$ are the distinct eigenvalues of the matrix.

Let us understand these ideas with the help of an example involving a non-diagonalizable matrix.

Example 3: Nondiagonalizable matrix.

Determine whether the matrix

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

can be diagonalized.

Solution

The characteristic polynomial of A is

$$\det|A - \lambda I| = \begin{vmatrix} 3-\lambda & 1 & -1 \\ 2 & 2-\lambda & -1 \\ 2 & 2 & -\lambda \end{vmatrix} = (1-\lambda)(2-\lambda)^2$$

The eigenvalues are 1,2,2. To determine the minimal polynomial, consider the product

$$(A - \lambda_1 I)(A - \lambda_2 I) = (A - I)\,(A - 2I)$$

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

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

This is not a zero matrix.

Now from the Cayley Hamilton theorem $(A-I)(A-2I)^2 = 0$. Therefore, the minimal polynomial is $(1-\lambda)(2-\lambda)^2$. It has a double root for $\lambda=2$ and therefore, by statement 5 above, the matrix A is not diagonalizable.

Let us now try to obtain the eigenvectors of A. For $\lambda=1$, you can determine the eigenvector to be

$$\mathbf{x}_1 = (1 \;\; 0 \;\; 2)^T$$

For the double root $\lambda=2$, we have

$$(A-2I)\mathbf{x} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

The solution is $x_2 = x_1$, $x_3 = 2x_1$ which gives the second eigenvector as

$$\mathbf{x}_2 = (1 \;\; 1 \;\; 2)^T$$

This is the only linearly independent vector associated with $\lambda=2$ which has multiplicity 2. Thus for the matrix A of order 3, there exist only 2 linearly independent eigenvectors. Hence the matrix is nondiagonalizable.

Would you like to do a numerical calculation to verify the result obtained in Example 3? Try the following SAQ.

SAQ 3

Spend 10 min

(a) Take a vector $\mathbf{x}_3$ linearly independent of $\mathbf{x}_1$ and $\mathbf{x}_2$:

$$\mathbf{x}_3 = (0 \;\; 0 \;\; 1)^T.$$

Construct the matrix $P = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ and calculate $P^{-1}AP$. Is it a diagonal matrix?

(b) Apply statements 1 to 5 to ascertain whether the following matrices are diagonalizable or not?

(i) $A = \begin{pmatrix} 2 & -3 \\ 0 & 0 \end{pmatrix}$ (ii) $B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$

(iii) $C = \begin{pmatrix} 1 & 1-i \\ 1+i & 0 \end{pmatrix}$ (iv) $D = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$

So far you have learnt the method of reducing a matrix to its diagonal form. You have seen that the process becomes particularly simple for matrices that are diagonalizable.Let us now consider some applications of matrix diagonalization.

3.3 APPLICATIONS OF MATRIX DIAGONALIZATION

In this section you are going to study certain specific applications of matrix diagonalization to quadratic equations, vibrational problems and mechanics. Let us begin by studying quadratic forms.

3.3.1 Quadratic Forms

We come across many quadratic forms in physics. The simplest example of such a form is the expression of energy of a linear harmonic oscillator. You know that the expression for energy is given by

$$E = \frac{1}{2} m \dot{x}^2 + \frac{1}{2} k x^2 \tag{3.12}$$

where the symbols have their usual meaning.

Notice that the first term in Eq. (3.12) is quadratic in velocity and the second term is quadratic in displacement. So E is a homogeneous polynomial of degree '2' in $\dot{x}$ and x. Thus you can appreciate why we call E a quadratic form. Like Eq. (3.12) any homogeneous polynomial of the second degree in any number of variables is called a **quadratic form**. Thus,

$$2x^2 + 4xy + 7y^2, \; x^2 + 2y^2 + z^2 - xy + 3yz - 2zx$$

are quadratic forms in 2 and 3 variables, respectively. By definition, a general quadratic form is given by

$$\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_i x_j$$

which is in n variables $x_1, x_2, ..., x_n$.

Both i and j range from 1 to n. Note that the square terms $x_1^2, x_2^2, ..., x_n^2$ correspond to $i = j$, and the corresponding diagonal coefficients are $a_{11}, a_{22}, ..., a_{nn}$.

The coefficient of each product term $x_i x_j \, (i \neq j)$ is $(a_{ij} + a_{ji})$ as it arises out of $x_i \times x_j$ as well as $x_j \times x_i$. For example, in Eq. (3.12) $a_{11} = m/2, \; a_{12} = a_{21} = 0, \; a_{22} = k/2$.

By now you must be wondering what the connection between matrices and quadratic forms is. Let us study that now.

Let us define another set of numbers b_{ij} such that

$$b_{ij} = b_{ji} = \frac{1}{2}(a_{ij} + a_{ji}), \qquad \text{for } i \neq j \tag{3.13}$$

In matrix notation we can write this as

$$B = B^T = \frac{1}{2}(A + A^T)$$

where

$$B = (b_{ij})$$

$$A = (a_{ij})$$

Since $B = B^T$, B is a symmetric matrix.

Then we can write the general quadratic form as

$$\sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}x_i x_j = \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ij}x_i x_j$$

Thus using Eq. (3.13) **every quadratic form can be so written that the matrix $B = [b_{ij}]$ of its coefficients is symmetric.**

For example, for the coefficients a_{ij} of Eq. (3.12),

$$B = \begin{pmatrix} m/2 & 0 \\ 0 & k/2 \end{pmatrix}$$

Let us consider the quadratic form

$$2x^2 + 4xy + 7y^2 = 2x^2 + 2xy + 2yx + 7y^2,$$

Notice that here we have expressed $4xy$ as a sum of two parts. Then you can see that

$$a_{11} = 2,\ a_{12} = 2,\ a_{21} = 2,\ a_{22} = 7.$$

Thus

$$B = \begin{pmatrix} 2 & 2 \\ 2 & 7 \end{pmatrix}$$

The quadratic form

$$x^2 + 2y^2 + z^2 - xy + 3yz - 2zx$$

$$= x^2 + 2y^2 + z^2 + \left(-\frac{1}{2}\right)xy + \left(-\frac{1}{2}\right)yx + \frac{3}{2}yz + \frac{3}{2}yz + (-1)zx + (-1)xz$$

Therefore, $a_{11} = 1,\ a_{12} = -1/2,\ a_{13} = -1,\ a_{21} = -1/2,\ a_{22} = 2,\ a_{23} = 3/2,$

$$a_{31} = -1,\ a_{32} = 3/2,\ a_{33} = 1$$

Thus

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

Further, **a quadratic form can be expressed as a product of matrices.**

Consider the quadratic form $F(x_1, x_2) = b_{11}x_1^2 + 2b_{12}x_1x_2 + b_{22}x_2^2$. Let us write

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \qquad \text{where } b_{12} = b_{21}.$$

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \text{and} \qquad X^T = (x_1, x_2)$$

Then you can verify that $F(x_1, x_2) \equiv X^T B X$.

$$X^T BX = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} b_{11}x_1 + b_{12}x_2 \\ b_{21}x_1 + b_{22}x_2 \end{pmatrix}$$

$$= x_1(b_{11}x_1 + b_{12}x_2) + x_2(b_{21}x_1 + b_{22}x_2)$$

$$= b_{11}x_1^2 + 2b_{12}x_1x_2 + b_{22}x_2^2 \qquad (\because b_{12} = b_{21})$$

So, $X^T BX$ is a product of matrices X^T, B and X representing a quadratic form. We can generalise this result for an $(n \times n)$ matrix B and the corresponding $(n \times 1)$ matrix X:

$$F(x_1, x_2, ..., x_n) \equiv X^T BX$$

where

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}, \qquad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

You may now like to do an SAQ on quadratic forms to get familiar with them.

SAQ 4

Spend 10 min

a) Express the given quadratic form as a product of matrices

$$ax^2 + 2hxy + bxy$$

b) Write down the quadratic form corresponding to the following symmetric matrix:

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

You now know that a quadratic form can be expressed as

$$S = X^T AX \tag{3.14}$$

where X is an $(n \times 1)$ column matrix and A an $(n \times n)$ symmetric matrix. Now recall from Sec. 3.2.3 that a real, symmetric matrix can be diagonalized by an orthogonal matrix. We can apply this result to simplify quadratic forms by getting rid of the terms containing cross products. Let us see how. Consider the linear transformation

$$X = BY$$

where B is an $(n \times n)$ orthogonal matrix and Y is an $(n \times 1)$ column matrix.

It will transform Eq. (3.14) as

$$S = (BY)^T A\ BY$$

$$= Y^T B^T A\ BY$$

Now, if B is an orthogonal matrix formed by n eigenvectors of A as its columns, then $B^T AB = D$, is a diagonal $(n \times n)$ matrix. In such an event, S is a new quadratic form given by

$$S = Y^T D Y \qquad (3.15)$$

Thus, if $Y = (y_1, y_2, ..., y_n)^T$ and if the eigenvalues of A are $\lambda_1, \lambda_2, ..., \lambda_n$, then Eq. (3.15) becomes

$$S = (y_1, y_2, ..., y_n) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & & \\ 0 & & & \lambda_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

or

$$S = \lambda_1 y_1^2 + \lambda_2 y_2^2 + ... + \lambda_n y_n^2 \qquad (3.16)$$

Thus, **in the transformed form S has only squared terms and no cross products.**

We can use this result to identify conic sections whose equations have been expressed in a quadratic form.

Example 4: Identification of a conic section from a given quadratic form.

Identify the conic section whose equation is

$$5x^2 - 2xy + 5y^2 = 4.$$

Solution

We can rewrite this equation in the form

$$X^T A X = 4$$

where $X = \begin{pmatrix} x \\ y \end{pmatrix}$ and $A = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}$.

You should verify this result before studying further. A is symmetric and hence diagonalizable. Therefore, to identify the conic, we need to obtain only the eigenvalues of A:

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

or

$$(5-\lambda)^2 - 1 = 0$$

or

$$\lambda^2 - 10\lambda + 24 = (\lambda - 6)(\lambda - 4) = 0$$

Thus $\lambda_1 = 6$, and $\lambda_2 = 4$. So we can write the quadratic form as

$$6x'^2 + 4y'^2 = 4$$

or

$$\frac{x'^2}{2/3} + y'^2 = 1$$

This is the equation of an ellipse.

You may now like to solve a similar SAQ to get some practice.

SAQ 5

Spend 10 *min*

For the quadratic equation

$$5x^2 - 4xy + 2y^2 = 4$$

write down the matrix of coefficients and diagonalize it. Recast it in new variables and identify the conic section it represents.

Let us now study the applications of diagonalization in physics.

3.3.2 Applications in Physics

Matrix diagonalization finds many applications in physics. Let us begin by studying the very instructive example of Inertia Matrix.

Moment of Inertia

You know that angular momentum **L** is defined as

$$\mathbf{L} = I\,\omega \tag{3.17}$$

where ω is the angular velocity vector, and I is the moment of inertia. For a rigid body R (Fig. 3.1), I is expressed as a (3 × 3) matrix, called the ***Inertia Matrix***.

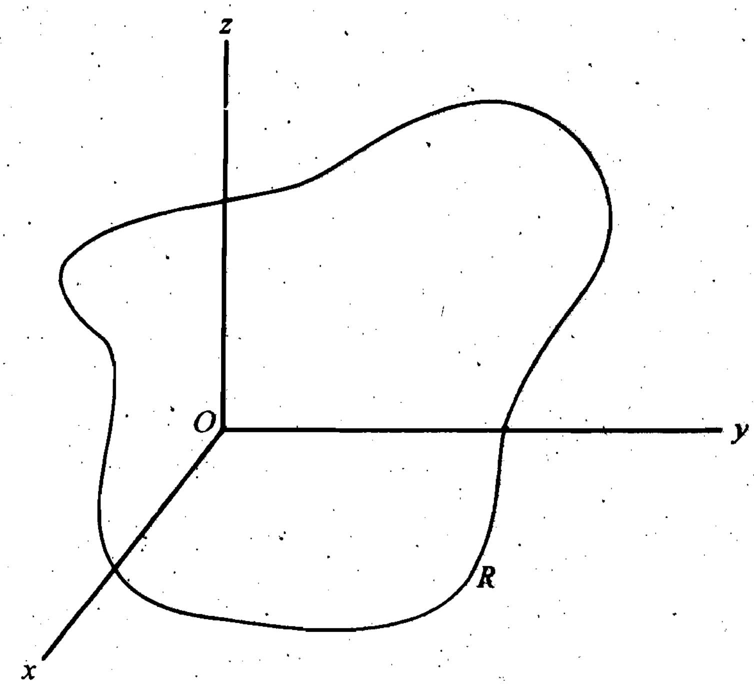

Fig.3.1: Moment of inertia of a body about the x, y, z axes

In this case Eq. (3.17) takes the form

$$\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \quad (3.18)$$

where

$$I_{xx} = \int_R \left(y^2 + z^2\right) dm \quad (3.19a)$$

$$I_{yy} = \int_R \left(x^2 + z^2\right) dm \quad (3.19b)$$

$$I_{zz} = \int_R \left(x^2 + y^2\right) dm \quad (3.19c)$$

$$I_{xy} = I_{yx} = -\int_R x\, y\, dm \quad (3.19d)$$

$$I_{xz} = I_{zx} = -\int_R x\, z\, dm \quad (3.19e)$$

$$I_{yz} = I_{zy} = -\int_R y\, z\, dm \quad (3.19f)$$

I_{xx}, I_{yy} and I_{zz} are known as the **moments of inertia** of R about the x, y, z axes, respectively. I_{xy}, I_{xz}, I_{yz} are the **products of inertia** of R and dm is the element of mass. Notice that I is a symmetric matrix, and hence it can be diagonalized. The new axes, say x', y', z' chosen to diagonalise I, are called the **principal axes** (Fig. 3.2).

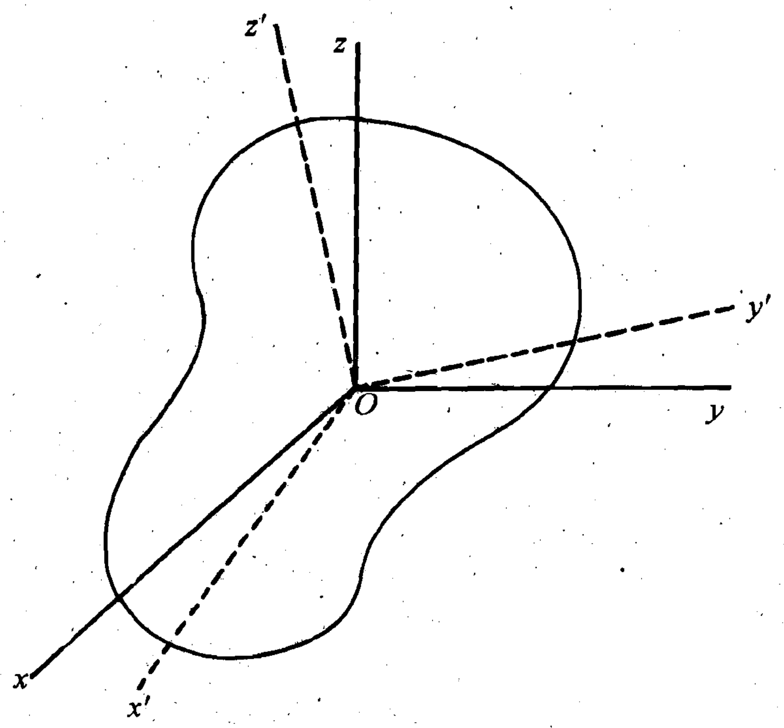

Fig.3.2: The unprimed axes represent the original rectangular co-ordinate system. The primed axes are the principal axes.

Thus if $\mathbf{x} = P\,\mathbf{x}'$, where $\mathbf{x} = (x, y, z)^T$ and $\mathbf{x}' = (x', y', z')^T$ then

$$\mathbf{L} = P\,\mathbf{L}' \text{ and } \boldsymbol{\omega} = P\boldsymbol{\omega}'$$

so that $\mathbf{L} = I\boldsymbol{\omega}$ takes the form

Do not confuse the matrix I here with the identity matrix.

$$P\,\mathbf{L}' = I\,P\boldsymbol{\omega}'$$

or

$$\mathbf{L}' = P^{-1}I\,P\boldsymbol{\omega}'$$

where $P^{-1}I\,P = I'$ is diagonal. Its diagonal elements $I_{x'x'}, I_{y'y'}$ and $I_{z'z'}$ are called **principal inertias**.

This transformation to principal axes can in fact be considered as a rotation in space and is thus an example of orthogonal transformation.

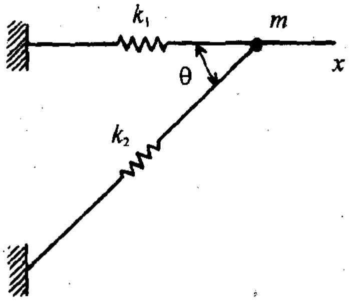

Fig.3.3: Spring-mass system lying on a frictionless table in a horizontal plane as seen from above.

One of the most important applications of matrix diagonalization occurs in vibration problems in physics, for example, vibrations of strings in musical instruments, vibrations of mechanical systems, of radio waves, or of electric currents and voltages in a tuned radio, and so on. Here we consider a coupled spring-mass system (Fig. 3.3).

Spring-mass system

For small oscillations of this system, the differential equations governing the motion of the system are given by

$$m\ddot{x} = -\left(k_1 + \frac{k_2}{2}\right)x - \frac{k_2}{2}y$$

$$m\ddot{y} = -\frac{k_2}{2}x - \frac{k_2}{2}y$$

We can solve these equations by following the procedure of Example 4 in Unit 1. Alternatively we can use the procedure of diagonalization. Using the matrix notations

$$\mathbf{x} \equiv \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{and} \quad A \equiv \begin{pmatrix} \dfrac{2k_1 + k_2}{2m} & \dfrac{k_2}{2m} \\ \dfrac{k_2}{2m} & \dfrac{k_2}{2m} \end{pmatrix}$$

We can express the problem compactly as

$$\ddot{\mathbf{x}} + A\mathbf{x} = 0$$

As you can see, A is not diagonal. But it is symmetric, and hence diagonalizable.

Now introducing a change of variables

$$\mathbf{x} = P\,\mathbf{y}, \quad \text{where} \quad \mathbf{y} = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

we get

$$P\,\ddot{\mathbf{y}} + A\,P\,\mathbf{y} = 0$$

If P^{-1} exists then

$$\ddot{\mathbf{y}} + P^{-1}A\,P\,\mathbf{y} = 0$$

Since $P^{-1}AP$ is a diagonal matrix, the solutions for **y** are easily determined as follows:

You can determine the eigenvalues of A to be

$$\lambda = \frac{(k_1 + k_2) \pm \sqrt{k_1^2 + k_2^2}}{2m}$$

For the sake of simplicity in writing, let us assign some numerical values to m, k_1 and k_2:

$$m = 1, \quad k_1 = 3 \quad \text{and} \quad k_2 = 4$$

Then

$$\lambda_1 = 1, \quad \mathbf{y}_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \lambda_2 = 6, \quad \mathbf{y}_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Therefore

$$P = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}$$

and

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}$$

The coupled equations in **x** are thus uncoupled in **y** as follows:

$$\ddot{x}' + x' = 0$$

$$\ddot{y}' + 6y' = 0$$

The general solutions are

$$x' = A_1 \sin(t + \phi_1)$$

$$y' = A_2 \sin\left(\sqrt{6}\,t + \phi_2\right)$$

where A_1, A_2, ϕ_1, and ϕ_2 are constants of integration. To write the solution in the original variables, we use the transformation $\mathbf{x} = P\,\mathbf{y}$ to get

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

which yields

$$x(t) = \frac{A_1}{\sqrt{5}} \sin(t + \phi_1) + \frac{2A_2}{\sqrt{5}} \sin\left(\sqrt{6}\, t + \phi_2\right)$$

$$y(t) = -\frac{2A_1}{\sqrt{5}} \sin(t + \phi_1) + \frac{A_2}{\sqrt{5}} \sin\left(\sqrt{6}\, t + \phi_2\right)$$

This example must have given you a flavour of how much the diagonalization of a matrix simplifies our calculations.

You may like to end this unit by attempting the following SAQ.

SAQ 6

Spend 5 min

The Hamiltonian of a one-dimensional harmonic oscillator in its matrix representation is given by

$$H = \hbar \omega_0 \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \cdots & \\ \frac{1}{2} & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & \cdots & \\ \frac{1}{2} & \frac{1}{2} & \frac{5}{2} & \frac{1}{2} & \cdots & \\ \vdots & & & & & \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \cdots & \left(n + \frac{1}{2}\right) \\ \vdots & & & & & \end{pmatrix}.$$

with eigenvalues $\hbar \omega_0 \left(n + \frac{1}{2}\right)$, $n = 0, 1, 2, \ldots$ Reduce H to a diagonal form.

Let us now summarise what you have studied in this unit.

3.4 SUMMARY

- Two matrices A and B are said to be related by a **similarity transformation** if there exists an invertible matrix P such that

  $$B = P^{-1} A P \quad \text{or} \quad A = PBP^{-1}$$

  If P is a unitary matrix, the transformation is called a **unitary transformation**. If P is a real, symmetric matrix, the transformation is called an **orthogonal transformation**.

- If the transforming matrix P is constructed with the orthonormal eigenvectors of the matrix A as its column vectors then the matrix $B \left(\equiv P^{-1} AP\right)$ is a **diagonal matrix**, with the eigenvalues of A as its diagonal elements.

- A matrix is **diagonalizable** only if it has n linearly independent eigenvectors. If it is a symmetric or hermitian matrix, then it can always be diagonalized.

- A hermitian matrix can be diagonalized via a **unitary transformation.**

- A real, symmetric matrix can be diagonalized via an **orthogonal transformation.**

- Matrix diagonalization finds applications in many areas, e.g., in reducing quadratic forms, vibration problems, mechanics and quantum mechanics.

3.5 TERMINAL QUESTIONS

Spend 40 min

1. Diagonalize the matrix

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

2. Express the following quadratic forms as products of matrices:

   (a) $ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy$

   (b) $x_1x_2 + x_2x_3 + x_3x_1 + x_1x_4 + x_2x_4 + x_3x_4$

3. If $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, calculate A^8.

4. The moment of inertia matrix of a square of side 6 units is given by (see Example 2, Unit 2)

$$I = \begin{pmatrix} 12 & -9 & 0 \\ -9 & 12 & 0 \\ 0 & 0 & 24 \end{pmatrix}$$

Obtain the diagonalizing matrix and hence reduce I to its diagonal form. What is the physical significance of this result?

3.6 SOLUTIONS AND ANSWERS

Self-assessment Questions

1. (a) We have to show that

$$B = P^{-1} A P$$

Now

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

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

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

$$= B$$

Hence proved.

(b) We have to show that if $B \sim A$ then $A \sim B$ and if $A \sim B$ then $B \sim A$. Now if $B \sim A$ then there exists an invertible matrix P such that

$$B = P^{-1} A P$$

Multiplying from the left by P and from the right by P^{-1} we get

$$P B P^{-1} = A$$

which by Eq. (3.4b) is also a similarity transformation. If $A \sim B$ then

$$A = P^{-1} B P$$

Again multiplying by P from the left and P^{-1} from the right we get

$$P A P^{-1} = B$$

which is again a similarity transformation.

If $A \sim B$ and $B \sim C$, then we must have invertible matrices P and Q such that

$$A = P^{-1} B P \quad \text{and} \quad B = Q^{-1} C Q$$

Substituting for B we get

$$A = P^{-1} Q^{-1} C Q P$$

Let us define $R = Q P$. Then $R^{-1} = P^{-1} Q^{-1}$. Thus we have

$$A = R^{-1} C R$$

which is a similarity transformation.

2. You can verify that the eigenvalues of M are $\lambda_1 = 6$ and $\lambda_2 = 1$ and the corresponding eigenvectors are:

$$\mathbf{x}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{x}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Thus P is given by

$$P = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}$$

and P^{-1} can be determined to be

$$P^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}$$

Thus

$$P^{-1} M P = \frac{1}{5} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$= \frac{1}{5}\begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -12 & 1 \\ 6 & 2 \end{pmatrix}$$

$$= \frac{1}{5}\begin{pmatrix} 30 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$$

which is a diagonal matrix.

3. (a) $$P = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix}$$

You can determine the inverse of P to be

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

Hence

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

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

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

Thus $P^{-1} A\ P$ is not a diagonal matrix. Whatever components we choose for $\mathbf{x}_3$, such that $\mathbf{x}_3$ is linearly independent of $\mathbf{x}_1$ and $\mathbf{x}_2$, the matrix P will **not** reduce A to a diagonal form. There exists no matrix P for which $P^{-1} A\ P$ is diagonal.

(b) (i) The eigenvalues of A are 2 and 0. Since there are 2 distinct eigenvalues for this 2×2 matrix, it is diagonalizable.

(ii) B is a symmetric matrix, hence it is diagonalizable.

(iii) C is a 2×2 hermitian matrix with distinct eigenvalues $\lambda_1=2$, $\lambda_2=-1$, hence it is diagonalizable.

(iv) D has only one distinct eigenvalue and it is not symmetric. Hence it is not diagonalizable.

4. (a)
$$ax^2+2hxy+by^2$$
$$=ax^2+hxy+hyx+by^2$$
$$=(x\ \ y)\begin{pmatrix} a & h \\ h & b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

(b) The required quadratic form would be

$$(x\ \ y\ \ z)\begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 3 \\ 3 & 3 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$=(x\ \ y\ \ z)\begin{pmatrix} x+2y+3z \\ 2x+3z \\ 3x+3y+z \end{pmatrix}$$

$$=x(x+2y+3z)+y(2x+3z)+z(3x+3y+z)$$
$$=x^2+z^2+6yz+6zx+4xy$$

5. We can write the quadratic equation in the form

$$(x\quad y)\begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=4$$

The matrix of coefficients is $\begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}$. It is symmetric with eigenvalues $\lambda_1=6$ and $\lambda_2=1$. Thus the matrix can be reduced to its diagonal form $\begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$. The equation in the new variables is

$$6x'^2+y'^2=4$$

or

$$\frac{x'^2}{2/3}+\frac{y'^2}{4}=1.$$

This is the equation of an ellipse.

6. Since H is symmetric, it is diagonalizable with its eigenvalues as the diagonal elements. Hence, in its diagonal form

$$H = \hbar\,\omega_0 \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & \cdots \\ 0 & \frac{3}{2} & 0 & 0 & \cdots \\ 0 & 0 & \frac{5}{2} & 0 & \cdots \\ \vdots & & & & \\ 0 & 0 & 0 & 0 & \left(n+\frac{1}{2}\right) \end{pmatrix}$$

since the eigenvalues of H are given to be $\hbar\,\omega_0\left(n+\frac{1}{2}\right)$.

Terminal Questions

1. The characteristic equation of S is

$$\begin{vmatrix} 1-\lambda & -6 & -4 \\ 0 & 4-\lambda & 2 \\ 0 & -6 & -3-\lambda \end{vmatrix} = 0$$

$\therefore \quad (1-\lambda)\ \{(\lambda-4)(\lambda+3)+12\}=0$

or $\quad (\lambda-1)\ (\lambda^2-\lambda)=0$

or $\quad \lambda(\lambda-1)\ (\lambda-1)=0$

So the roots are 0, 1, 1.

You can verify that $(2\ \ -1\ \ 2)^T$, $(1\ \ 0\ \ 0)^T$ and $(0\ \ 2\ \ -3)^T$ are the eigenvectors corresponding to $\lambda = 0$, $\lambda = 1$ and $\lambda = 1$, respectively.

Thus the diagonalizing matrix is

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

And

$$P^{-1} = \frac{\text{Adj}\,P}{\det P} = \begin{pmatrix} 0 & 3 & 2 \\ 1 & -6 & -4 \\ 0 & 2 & 1 \end{pmatrix}$$

The diagonalized matrix can be calculated as follows:

$$B = P^{-1} S P = \begin{pmatrix} 0 & 3 & 2 \\ 1 & -6 & -4 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -6 & -4 \\ 0 & 4 & 2 \\ 0 & -6 & -3 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ -1 & 0 & 2 \\ 2 & 0 & -3 \end{pmatrix}$$

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

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

2. (a) $ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy$

$$= ax^2 + by^2 + cz^2 + hxy + hyx + fyz + fzy + gzx + gxz$$

$$= (x \quad y \quad z) \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

(b) The given expression can be written as

$$= 0.x_1^2 + 0.x_2^2 + 0.x_3^2 + 0.x_4^2$$

$$+ \frac{1}{2}(x_1x_2 + x_2x_1) + \frac{1}{2}(x_1x_3 + x_3x_1) + \frac{1}{2}(x_1x_4 + x_4x_1)$$

$$+ \frac{1}{2}(x_2x_3 + x_3x_2) + \frac{1}{2}(x_2x_4 + x_4x_2) + \frac{1}{2}(x_3x_4 + x_4x_3)$$

$$= (x_1 \quad x_2 \quad x_3 \quad x_4) \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

3. Suppose B is similar to the matrix A:

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

$$\therefore \quad PB = PP^{-1} AP = AP$$

and

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

Now

$$A^2 = PBP^{-1}PBP^{-1}$$

$$= PB^2P^{-1}$$

$$A^3 = PB^2P^{-1}PBP^{-1}$$

$$= PB^3P^{-1}$$

$$\vdots$$

$$A^8 = PB^8P^{-1}$$

Since B^8 can be a diagonal matrix, the problem reduces to calculating P, P^{-1} and obtaining PB^8P^{-1}.

The characteristic equation of A is

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0 \qquad \text{i.e.} (\lambda-1)(\lambda-3) = 0$$

or $\lambda = 1, 3$. For $\lambda = 1$, the characteristic vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ satisfies

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

$$\therefore \quad 2x_1 + x_2 = x_1 \qquad \text{and} \qquad x_1 + 2x_2 = x_2$$

or $\quad x_1 = -x_2$

And for $\lambda = 3$, we have

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\therefore \quad 2x_1 + x_2 = 3x_1 \qquad \text{and} \qquad x_1 + 2x_2 = 3x_2$$

or $\quad x_1 = x_2$

Thus we obtain two eigenvectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

The normalised eigenvectors are $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Now let us write the diagonalizing matrix P and the diagonal matrix B:

$$P = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

Since A is symmetric, the diagonal matrix is

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

And since

$$P^T = P,$$

$$\therefore \quad P^{-1} = P$$

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

Now

$$B^8 = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}^8 = \begin{pmatrix} 3^8 & 0 \\ 0 & 1^8 \end{pmatrix} = \begin{pmatrix} 6561 & 0 \\ 0 & 1 \end{pmatrix}$$

Therefore,

$$A^8 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 6561 & 0 \\ 0 & 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 6561 & 6561 \\ 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 6562 & 6560 \\ 6560 & 6562 \end{pmatrix} = \begin{pmatrix} 3281 & 3280 \\ 3280 & 3281 \end{pmatrix}$$

4. The inertia matrix is

$$\begin{pmatrix} 12 & -9 & 0 \\ -9 & 12 & 0 \\ 0 & 0 & 24 \end{pmatrix} \text{ which is symmetric.}$$

$$= 3\begin{pmatrix} 4 & -3 & 0 \\ -3 & 4 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

Let us now diagonalize the matrix. For that, we shall leave out the common factor 3.

The characteristic equation is

$$\begin{vmatrix} 4-\lambda & -3 & 0 \\ -3 & 4-\lambda & 0 \\ 0 & 0 & 8-\lambda \end{vmatrix} = 0$$

which yields 3 roots; $\lambda = 8, 7, 1$.

Determination of eigenvectors

For $\lambda = 8$

$$\begin{pmatrix} 4 & -3 & 0 \\ -3 & 4 & 0 \\ 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 8 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$\therefore \quad 4x_1 - 3x_2 = 8x_1; \quad -3x_1 + 4x_2 = 8x_2; \quad 8x_3 = 8x_3$$

Thus

$$\frac{x_2}{x_1} = \frac{-4}{3}, \ \frac{x_2}{x_1} = \frac{-3}{4}$$ which are satisfied only for $x_1 = x_2 = 0$

and x_3 can be any finite quantity. Thus we can write the eigenvector as $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ which is normalised.

For $\lambda = 7$,

$$\begin{pmatrix} 4 & -3 & 0 \\ -3 & 4 & 0 \\ 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 7 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$\therefore \quad 4x_1 - 3x_2 = 7x_1; \quad -3x_1 + 4x_2 = 7x_2; \quad 8x_3 = 7x_3$$

i.e.

$$x_1 = -x_2 \text{ and } x_3 = 0$$

and the eigenvector can be written as $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ which after normalisation becomes

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

For $\lambda = 1$,

$$\begin{pmatrix} 4 & -3 & 0 \\ -3 & 4 & 0 \\ 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$\therefore \quad 4x_1 - 3x_2 = x_1; \quad -3x_1 + 4x_2 = x_2; \quad 8x_3 = x_3$$

i.e.

$$x_2 = x_1 \quad \text{and} \quad x_3 = 0$$

and the eigenvector can be written as $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ which after normalisation becomes $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

You can easily verify that the eigenvectors are mutually orthogonal.

The diagonalizing matrix is then

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

which is a symmetric matrix.

The moment of inertia matrix is diagonalized as follows:

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}\begin{pmatrix} 4 & -3 & 0 \\ -3 & 4 & 0 \\ 0 & 0 & 8 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}\begin{pmatrix} 1 & 7 & 0 \\ 1 & -7 & 0 \\ 0 & 0 & 8\sqrt{2} \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 16 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

You would remember that we had left out the common factor 3 and thus the diagonal matrix would be

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 21 & 0 \\ 0 & 0 & 24 \end{pmatrix}$$

The eigenvectors

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix},\ \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}\ \text{and}\ \begin{pmatrix}0\\0\\1\end{pmatrix}$$

expressed in terms of the basis $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ take the shapes mentioned below:

$$\frac{1}{\sqrt{2}}\left(\hat{\mathbf{i}}+\hat{\mathbf{j}}\right),\ \frac{1}{\sqrt{2}}\left(\hat{\mathbf{i}}-\hat{\mathbf{j}}\right)\ \text{and}\ \hat{\mathbf{k}}$$

which can be identified as the two diagonals of the given square and an axis perpendicular to the plane of the square and passing through the point of intersection of the diagonals.

The physical significance of principal axis transformation is that with reference to these axes, $\mathbf{L}$ is parallel to $\boldsymbol{\omega}$, otherwise $\mathbf{L}$ and $\boldsymbol{\omega}$ are not parallel.