Lecture Notes: Matrices

Semester I Core Course I
University of Calcutta

Dr. P. Mandal

Department of Physics
St. Paul's Cathedral Mission College

July 19, 2019

Contents

1

Preliminary Topics

Syllabus: (a) Addition and Multiplication of Matrices. Null Matrices. Diagonal, Scalar and Unit Matrices. Transpose of a Matrix. Symmetric and Skew-Symmetric Matrices. Conjugate of a Matrix. Hermitian and Skew- Hermitian Matrices. Singular and Non-Singular matrices. Orthogonal and Unitary Matrices. Trace of a Matrix. Inner Product.

(b) Eigenvalues and Eigenvectors. Cayley-Hamiliton Theorem. Diagonalization of Matrices. Solutions of Coupled Linear Ordinary Differential Equations. Functions of a Matrix.

1.1 Definitions: Some Terminologies

1. **Matrix:** A matrix is an array of numbers which may be complex. The array contains $m \times n$ numbers in m rows and n columns. $m \times n$ is the order of the matrix.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & \dots & a_{mn} \end{pmatrix}$$

   The element which belongs to the i-th row and j-th column is denoted as a_{ij}.

2. **Row matrix:** A row matrix contains only one row i.e. the numbers are arrayed in a single row. The order of such matrix is thus $1 \times n$.

$$R = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \end{pmatrix}$$

3. **Column matrix:** A column matrix contains only one column i.e. the numbers are arrayed in a single column. The order of such matrix is thus $m \times 1$.

$$C = \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ \dots \\ a_{m1} \end{pmatrix}$$

4. **Null matrix:** If all the elements of a matrix are zero, it is called a null matrix i.e. for a null matrix $a_{ij} = 0$ for all i, j.

5. **Square matrix:** If the number of rows (m) and the number of columns (n) of a matrix are equal, it is called a square matrix. The matrix S below is a square matrix of order 3.

$$S = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

   The elements a_{ii} are called the diagonal elements.

6. **Diagonal matrix:** If all the off-diagonal elements of a square matrix are zero but at least one diagonal element is non-zero, it is called a diagonal matrix i.e. for a diagonal matrix $a_{ij} \neq 0$ for all $i \neq j$ but $a_{ii} = 0$ for at least one i.

$$D = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$$

   D is a diagonal matrix of order 3.

7. **Unit or Identity matrix:** If all the diagonal elements of a diagonal matrix are equal to identity, it is called a unit or identity matrix. For a unit matrix $a_{ij} = 0$ for all $i \neq j$ and $a_{ii} = 1$ for all i i.e. $a_{ij} = \delta_{ij}$.

   $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is a diagonal matrix of order 3.

8. **Equal matrices:** Two matrices of same order are said to be equal if the elements of one matrix are equal to the corresponding elements of other matrix i.e. matrices A and B are equal if $a_{ij} = b_{ij}$ for all i, j.

1.2 Matrix Algebra

1. **Addition & Subtraction:** Addition or subtraction of two matrices A and B of same order is defined as $C = A \pm B$ where $c_{ij} = a_{ij} \pm b_{ij}$ for all i and j. For example,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \text{and } B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$$

$$\therefore A \pm B = \begin{pmatrix} a_{11} \pm b_{11} & a_{12} \pm b_{12} & a_{13} \pm b_{13} \\ a_{21} \pm b_{21} & a_{22} \pm b_{22} & a_{23} \pm b_{23} \end{pmatrix}$$

   Properties:

   - Matrix addition is commutative i.e. $A + B = B + A$.
   - Matrix addition is associative i.e. $A + (B + C) = (A + B) + C$

2. **Multiplication:**

   (a) Multiplication by a scalar: Multiplying a matrix by a number or scalar (say k) means multiplying each element by the number i.e. $kA = Ak = k(a_{ij})$.

   (b) Matrix product: The product of two matrices A and B is defined as $C = AB$, where

$$c_{ij} = \sum_{k=1}^{s} a_{ik} b_{kj}$$

For the product to be defined, matrices A and B must be conformable i.e. if $A = (a)_{ij}$ is an $m \times n$ matrix, then $B = (b)_{ij}$ must be an $n \times s$ matrix. The product matrix C is an $m \times s$ matrix. Consider, for example,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$$

Therefore, by definition, the product matrix

$$C = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix}$$

Consider the following system of equations:

$$\begin{aligned} 3x - y + 2z &= -3 \\ 2x + 3y - z &= 2 \\ x - 2y + z &= 5 \end{aligned} \tag{1.1}$$

The set of equations can be written in matrix form as

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

Properties:

- Matrix multiplication is, in general, not commutative i.e. $AB \neq BA$.
- Matrix multiplication is associative i.e. $A(BC) = (AB)C$
- Matrix multiplication follows the distributive law i.e. $A(B + C) = AB + AC$.

   (c) Direct product: The direct product is defined for general matrices. Given an $n \times n$ matrix A and an $m \times m$ matrix B, the direct product of A and B is an $nm \times nm$ matrix, and is defined by $C = A \otimes B$ where $C_{ik,jl} = a_{ij} b_{kl}$. If

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$C = A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix}$$

3. **Commutator of matrices:** Commutator of two square matrices A and B of same order is defined as $[A, B] = AB - BA$. In general, $AB \neq BA$ and hence $[A, B] \neq 0$. For example,

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

Note that

$$AB = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 4 & 6 \end{pmatrix} \text{ but } BA = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 8 \end{pmatrix}$$

If the commutator $[A, B] = AB - BA = 0$ i.e. $AB = BA$, the matrices are said to be commutative.

Properties:

- $[A, B] = -[B, A]$
- $[A, B + C] = [A, B] + [A, C]$
- $[A, BC] = [A, B]C + B[A, C]$
- $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$

The anticommutator of two matrices is defined as $\{A, B\} = AB + BA$.

4. **Power of a matrix:** For any positive integer n, the power of a square matrix A is defined as $A^n = AA...A$ (n times) i.e. $A^2 = AA$, $A^3 = AAA$. In particular, $A^0 = I$, the unit matrix[1].

5. **Function of a matrix:** The function of a matrix maps a matrix to another matrix. For example, consider a matrix function $f(A) = 3A^2 - 2A + I$ where I is the unit matrix of same order. A more fancy example is $f(A) = \sum_k a_k A^k$, where a_k are scalar coefficients. Another common series is defined by

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

There are several techniques for lifting a real function to a square matrix function. If the real function $f(x)$ has the Taylor series expansion

$$f(x) = f(0) + f'(0)x + f''(0)\frac{x^2}{2!} + ...$$

then a matrix function $f(A)$ can be defined by substituting x by a matrix A i.e.

$$f(A) = f(0)I + f'(0)A + f''(0)\frac{A^2}{2!} + ...$$

where the powers become matrix powers, the additions become matrix sums and the multiplications become scaling operations.

[1] $A^{-n} = A^{-1}A^{-1}...A^{-1}$ (n times) is defined if A is a nonsingular matrix

6. **Transpose of a matrix:** For any matrix A, the transpose matrix A^T is obtained by interchanging corresponding rows and columns of A, i.e. if $A = (a_{ij})$, $A^T = (a_{ji})$. For example,

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

Properties:

- $(A^T)^T = A$.
- $(A + B)^T = A^T + B^T$.
- $(AB)^T = B^T A^T$.

7. **Complex Conjugate:** For any matrix A, the complex conjugate matrix A^* is formed by taking the complex conjugate of each element of A, i.e. if $A = (a_{ij})$, $A^* = (a^*_{ij})$ for all i and j. For example,

$$A = \begin{pmatrix} 1 & 2+i \\ 3-i & i \end{pmatrix} \Rightarrow A^* = \begin{pmatrix} 1 & 2-i \\ 3+i & -i \end{pmatrix}$$

Obviously, $(A^*)^* = A$.

8. **Hermitian Conjugate:** For an arbitrary matrix A, the Hermitian conjugate matrix $A^\dagger$ is obtained by taking the complex conjugate of the matrix, and then the transpose of the complex conjugate matrix i.e. if $A = (a_{ij})$, $A^\dagger = (a^*_{ji})$ for all i and j. For example,

$$\begin{aligned} A &= \begin{pmatrix} 2+3i & 4-3i \\ 4i & 3 \end{pmatrix} \\ \Rightarrow A^\dagger &= \begin{pmatrix} 2-3i & 4+3i \\ -4i & 3 \end{pmatrix}^T = \begin{pmatrix} 2-3i & -4i \\ 4+3i & 3 \end{pmatrix} \end{aligned}$$

Properties:

- $(A^\dagger)^\dagger = A$.
- $(A + B)^\dagger = A^\dagger + B^\dagger$.
- $(AB)^\dagger = B^\dagger A^\dagger$.

9. **Trace of a matrix:** The trace of a square matrix is defined as the sum of its diagonal elements i.e. $\mathrm{Tr}(A) = \sum_i a_{ii}$. For example,

$$A = \begin{pmatrix} 2 & i \\ 0 & 3 \end{pmatrix} \Rightarrow \mathtt{Tr}(A) = 2 + 3 = 5$$

Properties:

- $\texttt{Tr}(A) = \texttt{Tr}(A^T)$.
- $\texttt{Tr}(A+B) = \texttt{Tr}(A) + \texttt{Tr}(B)$.
- $\texttt{Tr}$(AB)=$\texttt{Tr}$(BA).

10. **Determinant of a matrix:** The determinant of a square matrix A is defined as the determinant having same array as that of the matrix and is generally denoted as $|A|$ or $\texttt{det}(A)$. For example, the determinant of the matrix $A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$ is $|A| = \begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix} = 5$. If the determinant of a matrix is zero i.e. $|A| = 0$, A is called a singular matrix.

    Properties:

    - $|AB| = |BA| = |A||B|$
    - $|A^T| = |A|$

11. **Cofactor matrix:**

    The cofactor matrix is defined as $A^c = A^{ij}$. For example,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \Rightarrow A^c = \begin{pmatrix} A^{11} & A^{12} & A^{13} \\ A^{21} & A^{22} & A^{23} \\ A^{31} & A^{32} & A^{33} \end{pmatrix}$$

    where

$$A^{11} = (-1)^{1+1} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad A^{12} = (-1)^{1+2} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \quad A^{13} = (-1)^{1+3} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$A^{21} = (-1)^{2+1} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}, \quad A^{22} = (-1)^{2+2} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}, \quad A^{23} = (-1)^{2+3} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

$$A^{31} = (-1)^{3+1} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}, \quad A^{32} = (-1)^{3+2} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}, \quad A^{33} = (-1)^{3+3} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

    The cofactor matrix of $A = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}$ is $A^c = \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}$

12. **Adjoint of a matrix:**

    The adjoint of a matrix is defined as the transpose of its cofactor matrix i.e. $\texttt{adj}(A) = A^{cT}$. For example, consider the matrix $A = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}$. The cofactor matrix $A^c = \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}$.

    Hence, the adjoint matrix $\texttt{adj}(A) = \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}^T = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$.

13. **Inverse of a matrix:**

For a given nonsingular matrix A, there exists another matrix B such that $AB = BA = I$. Matrix B is called the inverse matrix of A ($B = A^{-1}$). For example, consider the following matrices:

$$A = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \quad \& \quad B = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$$

Note that,

$$AB = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Similarly,

$$BA = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

$\therefore B = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} = A^{-1}$, the inverse matrix of A.

The inverse matrix can be found by using the relation

$$A^{-1} = \frac{\texttt{adj}(A)}{|A|} \tag{1.2}$$

Problem 1: *Find out the inverse matrix of* $A = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}$.

Solution: Adjoint matrix of A i.e. $\texttt{adj}(A) = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$. The determinant of the matrix i.e. $|A| = \begin{vmatrix} 3 & 1 \\ 0 & 2 \end{vmatrix} = 6$

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

Properties:

- $(A^{-1})^{-1} = A$.
- $(A+B)^{-1} = A^{-1} + B^{-1}$.
- $(AB)^{-1} = B^{-1}A^{-1}$.

14. **Derivative of a matrix:** The derivative of a matrix with respect to a variable say, x is equal to the derivative of each element with respect x separately. For example,

$$\frac{d}{dx} \begin{pmatrix} x & x^2 & 1 \\ e^x & 0 & 2x^3 \end{pmatrix} = \begin{pmatrix} 1 & 2x & 0 \\ e^x & 0 & 6x^2 \end{pmatrix}$$

15. **Integral of a matrix:** The integral of a matrix with respect to a variable say, x is equal to the integral of each element with respect x separately. For example,

$$\int \begin{pmatrix} x & 3x^2 & 1 \\ e^x & 0 & 2x^3 \end{pmatrix} = \begin{pmatrix} \frac{x^2}{2} & x^3 & x \\ e^x & c & \frac{x^4}{2} \end{pmatrix} + \begin{pmatrix} c_1 & c_2 & c_3 \\ c_4 & 0 & c_5 \end{pmatrix}$$

1.3 Special Square Matrices

1. **Singular and non-singular matrices:** If the determinant of a matrix is zero, it is called singular i.e. for a singular matrix A, $|A| = 0$. Consider the following matrix for example:
$$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$$
As seen, $|A| = -1 + 1 = 0$. Therefore, it is a singular matrix.

   For a non-singular matrix, $|A| \neq 0$.

2. **Symmetric and skew-symmetric matrices:** A matrix is said to be symmetric if the transpose matrix is equal to the matrix itself i.e. for a symmetric matrix A, $A^T = A$. For example, consider the following matrix:
$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = A$$
If $A^T = -A$, the matrix A is called anti-symmetric or skew-symmetric. For an example, consider the matrix below:
$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -A$$

Problem 2: *Diagonal elements of a skew-symmetric matrix are zero.*

Solution: For a skew-symmetric matrix A, $A^T = -A$. In terms of the ij-th element, $a_{ij} = -a_{ji}$.
Now, for the diagonal elements $i = j$.
Therefore, $a_{ii} = -a_{ii}$ or, $a_{ii} = 0$ for all i.

Problem 3: *Any square matrix can be uniquely written as the sum of a symmetric matrix and a skew-symmetric matrix.*
Solution: Let A is a square matrix.
$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T) = P + Q$$
$$\text{Now, } P^T = \frac{1}{2}(A + A^T)^T = \frac{1}{2}\{A^T + (A^T)^T\} = \frac{1}{2}(A^T + A) = P$$
$$Q^T = \frac{1}{2}(A - A^T)^T = \frac{1}{2}\{A^T - (A^T)^T\} = \frac{1}{2}(A^T - A) = -Q$$
i.e. P is a symmetric matrix and Q is a skew-symmetric matrix. So any square matrix can be written as the sum of a symmetric matrix and a skew-symmetric matrix. To prove the representation unique, we assume $A = R + S$ where R is a symmetric matrix and S is a skew-symmetric matrix i.e. $R^T = R$ and $S^T = -S$.
$$A^T = R^T + S^T = R - S$$
$$\Rightarrow R = \frac{1}{2}(A + A^T), S = \frac{1}{2}(A - A^T)$$

---

3. **Hermitian and skew-Hermitian matrices:** A matrix is said to be Hermitian if the Hermitian conjugate matrix is equal to the matrix itself i.e. for a Hermitian matrix H, $H^\dagger = H$. For example, consider the following matrix:

$$\begin{aligned} H &= \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \\ \Rightarrow H^\dagger &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}^T = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = H \end{aligned}$$

If $H^\dagger = -H$, the matrix H is called anti-Hermitian or skew-Hermitian. For an example, consider the matrix below:

$$\begin{aligned} H &= \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \\ \Rightarrow H^\dagger &= \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}^T = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = -H \end{aligned}$$

Problem 4: *For an arbitrary matrix A, show that $A + A^\dagger$ and $i(A - A^\dagger)$ are both Hermitian.*

Solution: A matrix H is Hermitian if $H^\dagger = H$. Now,

$$\begin{aligned} \left(A + A^\dagger\right)^\dagger &= A^\dagger + (A^\dagger)^\dagger \\ &= A^\dagger + A \end{aligned}$$

Therefore, $A + A^\dagger$ is Hermitian. Similarly,

$$\begin{aligned} \left[i\left(A - A^\dagger\right)\right]^\dagger &= -i\left[A^\dagger + (A^\dagger)^\dagger\right] \\ &= -i\left(A^\dagger - A\right) \\ &= i(A - A^\dagger) \end{aligned}$$

Therefore, $i(A - A^\dagger)$ is also Hermitian.

---

4. **Orthogonal matrix:** For a unitary matrix O, $OO^T = O^TO = I$, the identity matrix. Consider the following example.

$$\begin{aligned} O &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow O^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \therefore\ OO^T &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I = O^TO \end{aligned}$$

So, O is a unitary matrix.

Problem 5: *Show that the determinant of an orthogonal matrix is* ± 1.

Solution: For an orthogonal matrix O, $OO^T = O^TO = I$, the identity matrix. Now,

$$\begin{aligned} |OO^T| &= |I| = 1 \\ \Rightarrow |O||O^T| &= 1 \\ \Rightarrow |O|^2 &= 1 \quad (\because |O^T| = |O|) \\ \Rightarrow |O| &= \pm 1. \end{aligned}$$

Problem 6: *Show that the inverse of an orthogonal matrix is equal to its transpose i.e.* $O^{-1} = O^T$.
Solution: For an orthogonal matrix O, $OO^T = O^TO = I$. Since $|O| = \pm 1$, the inverse matrix O^{-1} exists. Now,

$$\begin{aligned} O^{-1}OO^T &= O^{-1}I \\ \Rightarrow O^{-1} &= O^T \quad (\because O^{-1}O = I) \end{aligned}$$

---

5. **Unitary matrix:** For a unitary matrix U, $UU^\dagger = U^\dagger U = I$, the identity matrix. Consider the following example.

$$\begin{aligned} U &= \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \Rightarrow U^\dagger = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \\ \therefore\ UU^\dagger &= \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I = U^\dagger U \end{aligned}$$

Hence, U is a unitary matrix.

Problem 7: *Show that the inverse of a unitary matrix is equal to its Hermitian conjugate i.e.* $U^{-1} = U^\dagger$.
Solution: For a unitary matrix U, $UU^\dagger = U^\dagger U = I$. If U^{-1} is the inverse matrix of U, $U^{-1}U = I$. Now,

$$\begin{aligned} U^{-1}UU^\dagger &= U^{-1}I \\ \Rightarrow U^{-1} &= U^\dagger \end{aligned}$$

---

6. **Self-adjoint matrix:** If the transpose of the cofactor matrix i.e. the adjoint of any arbitrary matrix is equal to the matrix itself, it is called a self-adjoint matrix i.e. for a self-adjoint matrix $\texttt{adj}(A) = A$. For example,

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

$$\texttt{adj}(A) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}^T = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = A$$

1.4 Eigenvalue Problems

Every square matrix A satisfies a relation

$$AX = \lambda X \tag{1.3}$$

where λ is a scalar (real or complex) and X is a column matrix. Eq. (1.3) is called the eigenvalue equation of matrix A with eigenvalue λ and eigenvector X. If A is a square matrix of order n, X is a column matrix of order $n \times 1$.

From eq. (1.3), $(AX - \lambda I)X = 0$. In terms of the elements of the matrices A and X,

$$\begin{pmatrix} a_{11}-\lambda & a_{12} & ... & ... & a_{1n} \\ a_{21} & a_{22}-\lambda & ... & ... & a_{2n} \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & ... & a_{nn}-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ ... \\ ... \\ x_n \end{pmatrix} = 0 \tag{1.4}$$

$$\begin{aligned} \Rightarrow (a_{11}-\lambda)x_1 + a_{12}x_2 + ... + a_{1n}x_n &= 0 \\ a_{21}x_1 + (a_{22}-\lambda)x_2 + ... + a_{2n}x_n &= 0 \\ .... &= 0 \\ .... &= 0 \\ a_{n1}x_1 + a_{n2}x_2 + ... + (a_{nn}-\lambda)x_n &= 0 \end{aligned}$$

Thus we have a set of n number of linear homogeneous equations. Non-trivial solution exists if the determinant of the coefficients vanishes, i.e.

$$\begin{aligned} D(\lambda) &= \begin{vmatrix} a_{11}-\lambda & a_{12} & ... & ... & a_{1n} \\ a_{21} & a_{22}-\lambda & ... & ... & a_{2n} \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & ... & a_{nn}-\lambda \end{vmatrix} = 0 \\ \Rightarrow |A-\lambda I| &= 0 \end{aligned} \tag{1.5}$$

$D(\lambda)$ is a polynomial of degree n. It is called the characteristics polynomial of the given matrix A.

$D(\lambda) = |A - \lambda I| = 0$ (eq. 1.5) is the characteristic equation of the matrix A. The equation has n roots i.e. n number of possible values of λ - say $\lambda_1, \lambda_2, ..., \lambda_n$ (some of them may be equal). Thus we conclude that a matrix of order n has n number of eigenvalues.

The polynomial $D(\lambda)$ of degree n can be expressed as

$$D(\lambda) = |A - \lambda I| = c_0 + c_1\lambda + c_2\lambda^2 + ... + c_{n-1}\lambda^{n-1} + c_n\lambda^n \tag{1.6}$$

which implies that $c_0 = |A|$.

As $\lambda_1, \lambda_2, ..., \lambda_n$ are the roots of the characteristic equation (eq. 1.5),

$$D(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda).....(\lambda_n - \lambda) \tag{1.7}$$

By speculation of eq. 1.6 and eq. 1.7

$$\lambda_1\lambda_2....\lambda_n = c_0 = |A| \tag{1.8}$$

Thus the product of the eigenvalues of a matrix is equal to its determinant.

Similary, by inspection of eq. 1.6 and eq. 1.7 (equating the coefficients of λ^{n-1}) we find

$$\begin{aligned} c_{n-1} = (-1)^{n-1}(a_{11} + a_{22} + ... + a_{nn}) &= (-1)^{n-1}(\lambda_1 + \lambda_2 + ... + \lambda_n) \\ \Rightarrow \lambda_1 + \lambda_2 + ... + \lambda_n &= a_{11} + a_{22} + ... + a_{nn} = \mathtt{Tr}(A) \end{aligned} \tag{1.9}$$

Thus the sum of the eigenvalues is equal to the trace of the matrix.

Problem 8: *Find the trace and determinant of the matrix* $A = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}$ *and hence determine its eigenvalues.*

Solution: The trace of the matrix is the sum of its diagonal elements i.e. $\mathtt{Tr}(A) = 2 - 2 = 0$.

The determinant of the matrix is

$$|A| = \begin{vmatrix} 2 & -1 \\ 3 & -2 \end{vmatrix} = -4 + 3 = -1$$

If λ_1 and λ_2 are the eigenvalues of the matrix, by eq. 1.8 and eq. 1.9 we have

$$\begin{aligned} \lambda_1 + \lambda_2 &= \mathtt{Tr}(A) = 0 \\ \lambda_1\lambda_2 &= |A| = -1 \end{aligned}$$

Solving these equations, we find the eigenvalues as $\lambda_1 = -1, \lambda_2 = 1$.

How to determine the eigenvalues and the normalized eigenvectors of a matrix? Let us understand with the following examples.

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

The eigenvalue equation of the matrix is $AX = \lambda X$ or, $(A - \lambda I)X = 0$ where λ is the eigenvalue and X is the corresponding eigenvector. The characteristic equation is $|A - \lambda I| = 0$ i.e.

$$\begin{aligned} \begin{vmatrix} 2-\lambda & -1 \\ 3 & -2-\lambda \end{vmatrix} &= 0 \\ \Rightarrow (2-\lambda)(2+\lambda) + 3 &= 0 \\ \Rightarrow \lambda^2 &= 1 \\ \mathtt{or,}\ \lambda &= \pm 1 \end{aligned}$$

Thus the eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 1$.

M: 8902442561

Let, X_1 is the eigenvector of A which corresponds to the eigenvalue $\lambda_1 = -1$. From the eigenvalue equation $AX_1 = \lambda_1 X_1$, we have

$$\begin{aligned}
(A - \lambda_1 I)X_1 &= 0 \\
\Rightarrow \begin{pmatrix} 3 & -1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= 0 \\
\Rightarrow 3x_1 - x_2 &= 0 \\
\texttt{or, } x_2 &= 3x_1
\end{aligned}$$

If $x_1 = a$, $x_2 = 3a$ where a is an arbitrary number ($\neq 0$). The eigenvector of the given matrix corresponding to the eigenvalue $\lambda_1 = -1$ is $X_1 = \begin{pmatrix} a \\ 3a \end{pmatrix}$. In normalized form, $X_{1n} = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

Similarly, let us consider X_2 as the eigenvector of A corresponding to the eigenvalue $\lambda_2 = 1$. From the eigenvalue equation $AX_2 = \lambda_2 X_2$, we have

$$\begin{aligned}
(A - \lambda_2 I)X_2 &= 0 \\
\Rightarrow \begin{pmatrix} 1 & -1 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= 0 \\
\Rightarrow x_1 - x_2 &= 0 \\
\texttt{or, } x_1 &= x_2
\end{aligned}$$

If $x_1 = b$, $x_2 = b$ where b is another arbitrary number ($\neq 0$). Therefore, the eigenvector of the given matrix corresponding to the eigenvalue $\lambda_2 = 1$ is $X_2 = \begin{pmatrix} b \\ b \end{pmatrix}$. In normalized form, $X_{2n} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

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

The eigenvalue equation of the matrix is $AX = \lambda X$ or, $(A - \lambda I)X = 0$ where λ is the eigenvalue and X is the corresponding eigenvector. The characteristic equation is $|A - \lambda I| = 0$ i.e.

$$\begin{aligned}
\begin{vmatrix} -\lambda & 0 \\ 1 & -\lambda \end{vmatrix} &= 0 \\
\Rightarrow \lambda^2 + 1 &= 0 \\
\texttt{or, } \lambda &= \pm i
\end{aligned}$$

Thus the eigenvalues are $\lambda_1 = -i$ and $\lambda_2 = i$.

Let, X_1 is the eigenvector of A which corresponds to the eigenvalue $\lambda_1 = -i$. From the eigenvalue equation $AX_1 = \lambda_1 X_1$, we have

$$\begin{aligned}
(A - \lambda_1 I)X_1 &= 0 \\
\Rightarrow \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= 0 \\
\Rightarrow ix_1 - x_2 &= 0 \\
\texttt{or, } x_2 &= ix_1
\end{aligned}$$

If $x_1 = a$, $x_2 = ia$ where a is an arbitrary number ($\neq 0$). The eigenvector of the given matrix corresponding to the eigenvalue $\lambda_1 = -1$ is $X_1 = \begin{pmatrix} a \\ ia \end{pmatrix}$. In normalized form, $X_{1n} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$.

Similarly, let us consider X_2 as the eigenvector of A corresponding to the eigenvalue $\lambda_2 = i$. From the eigenvalue equation $AX_2 = \lambda_2 X_2$, we have

$$\begin{aligned} (A - \lambda_2 I)X_2 &= 0 \\ \Rightarrow \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= 0 \\ \Rightarrow x_1 - ix_2 &= 0 \\ \texttt{or,}\ x_1 &= ix_2 \end{aligned}$$

If $x_2 = b$, $x_1 = ib$ where b is another arbitrary number ($\neq 0$). Therefore, the eigenvector of the given matrix corresponding to the eigenvalue $\lambda_2 = i$ is $X_2 = \begin{pmatrix} ib \\ b \end{pmatrix}$. In normalized form, $X_{2n} = \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 1 \end{pmatrix}$.

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

The eigenvalue equation of the matrix is $AX = \lambda X$ or, $(A - \lambda I)X = 0$ where λ is the eigenvalue and X is the corresponding eigenvector. The characteristic equation is $|A - \lambda I| = 0$ i.e.

$$\begin{aligned} \begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & 2 \\ 0 & 2 & -\lambda \end{vmatrix} &= 0 \\ \Rightarrow (1-\lambda)(\lambda^2 - 4) &= 0 \\ \texttt{or,}\ \lambda &= 1, \pm 2 \end{aligned}$$

Thus the eigenvalues are $\lambda_1 = -2$, $\lambda_2 = 1$ and $\lambda_3 = 2$.

Let, X_1 is the eigenvector of A which corresponds to the eigenvalue $\lambda_1 = -2$. From the eigenvalue equation $AX_1 = \lambda_1 X_1$, we have

$$\begin{aligned} (A - \lambda_1 I)X_1 &= 0 \\ \Rightarrow \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= 0 \\ \Rightarrow 3x_1 &= 0 \quad \& \quad x_2 + x_3 = 0 \\ \texttt{or,}\ x_1 &= 0 \quad \& \quad x_2 = -x_3 \end{aligned}$$

If $x_3 = a$, $x_2 = -a$ where a is an arbitrary number ($\neq 0$). The eigenvector of the given matrix corresponding to the eigenvalue $\lambda_1 = -2$ is $X_1 = \begin{pmatrix} 0 \\ -a \\ a \end{pmatrix}$. In normalized form,

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

Similarly, let us consider X_2 as the eigenvector of A corresponding to the eigenvalue $\lambda_2 = 1$. From the eigenvalue equation $AX_2 = \lambda_2 X_2$, we have

$$\begin{aligned} (A - \lambda_2 I)X_2 &= 0 \\ \Rightarrow \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 2 \\ 0 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= 0 \\ \Rightarrow -x_2 + 2x_3 &= 0 \quad \& \quad 2x_2 - x_3 = 0 \\ \texttt{or,}\ x_2 = x_3 &= 0 \end{aligned}$$

Let $x_1 = b$, an arbitrary number ($\neq 0$). The eigenvector of the given matrix corresponding to the eigenvalue $\lambda_2 = 1$ is $X_2 = \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix}$. In normalized form, $X_{2n} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

If X_3 is the eigenvector of A corresponding to the eigenvalue $\lambda_3 = 2$, from the eigenvalue equation $AX_3 = \lambda_3 X_3$ we have

$$\begin{aligned} (A - \lambda_3 I)X_3 &= 0 \\ \Rightarrow \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 2 \\ 0 & 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= 0 \\ \Rightarrow x_1 &= 0 \quad \& \quad x_2 = x_3 \end{aligned}$$

Let $x_2 = x_3 = c$, where c is an arbitrary number ($\neq 0$). The eigenvector of the given matrix corresponding to the eigenvalue $\lambda_3 = 2$ is $X_3 = \begin{pmatrix} 0 \\ c \\ c \end{pmatrix}$. In normalized form, $X_{3n} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

1.4.1 Corollaries

1. *Eigenvalues of a diagonal matrix are equal to its diagonal elements.*

   Proof: Consider a diagonal matrix of order n:

$$\begin{pmatrix} a_{11} & 0 & \ldots & \ldots & 0 \\ 0 & a_{22} & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & a_{nn} \end{pmatrix}$$

The characteristic equation of the matrix is

$$\begin{aligned}
|D-\lambda I| &= 0\\
\Rightarrow \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & \dots & a_{1n}\\ a_{21} & a_{22}-\lambda & \dots & \dots & a_{2n}\\ \dots & \dots & \dots & \dots & \dots\\ \dots & \dots & \dots & \dots & \dots\\ a_{n1} & a_{n2} & \dots & \dots & a_{nn}-\lambda \end{vmatrix} &= 0\\
\Rightarrow (a_{11}-\lambda)(a_{22}-\lambda)...(a_{nn}-\lambda) &= 0
\end{aligned}$$

i.e. $\lambda = a_{11}, a_{22}, ..., a_{nn}$, the diagonal elements of the matrix.

2. *At least one eigenvalue of a singular matrix is zero.*

   Proof: Consider a singular matrix A i.e. $|A| = 0$.
   If λ_1, λ_2, $\lambda_3, ...$ are the eigenvalues of the matrix A, the product of the eigenvalues must be equal to the determinant of A (eq. 1.8) i.e.

   $$\lambda_1.\lambda_2.\lambda_3... = |A| = 0$$

   Therefore, at least one of the eigenvalues must be zero.

3. *If λ is the eigenvalue of a non-singular matrix A, the eigenvalue of A^{-1} is $1/\lambda$ corresponding to a given eigenvector.*

   Proof: Let λ and λ' are respectively the eigenvalues of a non-singular matrix A and its inverse matrix A^{-1} corresponding to the same eigenvector X. The eigenvalue equations are

   $$\begin{aligned} AX &= \lambda X\\ A^{-1}X &= \lambda' X \end{aligned}$$

   Now multiplying the first equation by A^{-1} from left

   $$\begin{aligned} A^{-1}AX &= \lambda A^{-1}X\\ \Rightarrow X &= \lambda\lambda' X\\ \text{or, } (1-\lambda\lambda')X &= 0 \end{aligned}$$

   Since X is the eigenvector ($X \neq 0$), $\lambda' = 1/\lambda$ i.e. the eigenvalues of the inverse matrix are the reciprocal of the eigenvalues of the original matrix.

4. *Eigenvalues of a unitary matrix are of unit magnitude.*

   Proof: Consider a unitary matrix U having an eigenvalue λ corresponding to an eigenvector X. The eigenvalue equation is

   $$UX = \lambda X \tag{1.10}$$

   Taking the Hermitian conjugate of eq. 1.10

   $$\begin{aligned} (UX)^\dagger &= (\lambda X)^\dagger\\ \Rightarrow X^\dagger U^\dagger &= \lambda^* X^\dagger \end{aligned} \tag{1.11}$$

Multiplying eq. (1.11) by eq. (1.10) from left

$$\begin{aligned}
X^\dagger U^\dagger U X &= \lambda^* \lambda X^\dagger X \\
\Rightarrow X^\dagger X &= \lambda^* \lambda X^\dagger X \quad (\because U \text{ is unitary, } U^\dagger U = I) \\
\text{or, } (1-|\lambda|^2) X^\dagger X &= 0 \\
\Rightarrow 1-|\lambda|^2 &= 0 \quad (\because X^\dagger X \neq 0) \\
\Rightarrow |\lambda|^2 &= 1
\end{aligned}$$

5. *The eigenvalues of a Hermitian matrix are real and the eigenvectors corresponding to different eigenvalues are orthogonal.*

   Proof: Let us consider a Hermitian matrix H having an eigenvalue λ corresponding to an eigenvector X. The eigenvalue equation is

$$HX = \lambda X \tag{1.12}$$

Multiplying eq.(1.12) by $X^\dagger$ from left

$$X^\dagger H X = \lambda X^\dagger X \tag{1.13}$$

Taking the hermitian conjugation of eq. (1.12),

$$\begin{aligned}
(HX)^\dagger &= (\lambda X)^\dagger \\
\Rightarrow X^\dagger H^\dagger &= \lambda^* X^\dagger \\
\text{or, } X^\dagger H &= \lambda^* X^\dagger \quad (\because H \text{ is Hermitian, } H^\dagger = H) \qquad (1.14)
\end{aligned}$$

Multiplying eq.(1.14) by X from right

$$X^\dagger H X = \lambda^* X^\dagger X \tag{1.15}$$

Comparing eq.(1.13) and eq.(1.15),

$$\begin{aligned}
\lambda X^\dagger X &= \lambda^* X^\dagger X \\
\Rightarrow (\lambda - \lambda^*) X^\dagger X &= 0 \\
\Rightarrow \lambda^* &= \lambda \quad (\because X^\dagger X \neq 0)
\end{aligned}$$

Thus the eigenvalues of a Hermitian matrix are real.

Now consider two distinct eigenvalues λ_1 and λ_2 of the Hermitian matrix H corresponding to the eigenvectors X_1 and X_2 respectively. $\lambda_1^* = \lambda_1$, $\lambda_2^* = \lambda_2$ and $\lambda_1 \neq \lambda_2$. The eigenvalue equations are

$$\begin{aligned}
HX_1 &= \lambda_1 X_1 \qquad (1.16) \\
HX_2 &= \lambda_2 X_2 \qquad (1.17)
\end{aligned}$$

Taking the hermitian conjugation of eq. (1.17),

$$\begin{aligned}
(HX_2)^\dagger &= (\lambda X_2)^\dagger \\
\Rightarrow X_2^\dagger H &= \lambda_2 X_2^\dagger \quad (\because H^\dagger = H \;\&\; \lambda_2^* = \lambda_2) \qquad (1.18)
\end{aligned}$$

Multiplying eq.(1.18) by X_1 from right

$$X_2^\dagger H X_1 = \lambda_2 X_2^\dagger X_1 \tag{1.19}$$

Multiplying eq.(1.16) by $X_2^\dagger$ from left

$$X_2^\dagger H X_1 = \lambda_1 X_2^\dagger X_1 \tag{1.20}$$

Comparing eq. 1.19 and eq. 1.20,

$$\begin{aligned} \lambda_1 X_2^\dagger X_1 &= \lambda_2 X_2^\dagger X_1 \\ \Rightarrow (\lambda_1 - \lambda_2) X_2^\dagger X_1 &= 0 \\ \Rightarrow X_2^\dagger X_1 &= 0 \quad (\because \lambda_2 \neq \lambda_1) \end{aligned}$$

Thus X_1 and X_2 are orthogonal.

6. *If two matrices commute, they will have simultaneous eigenfunction.*

   Proof: Let two matrices A and B commute i.e. $AB = BA$.
   If X is an eigenvector of A and λ is the associated eigenvalue, $AX = \lambda X$.
   Multiplying by B from left,

$$\begin{aligned} BAX &= \lambda BX \\ \text{or, } ABX &= \lambda BX \quad (\because AB = BA) \\ \Rightarrow A(BX) &= \lambda(BX) \end{aligned}$$

   Thus BX is another eigenfunction of A for the same eigenvalue λ. BX is therefore, a scalar multiple of X i.e.

$$BX = \mu X \tag{1.21}$$

   This is the eigenvalue equation of matrix B with eigenvalue μ and associated eigenfunction X. Thus X is the simultaneous eigenfunction for the matrices A and B.

1.5 Cayley-Hamilton Theorem

Cayley-Hamilton theorem states that every square matrix satisfies its own characteristic equation. Let us consider a matrix A of order n. If λ is the eigenvalue of A, the characteristic equation is $D(\lambda) = |A - \lambda I| = 0$. We rewrite eq. 1.6 as

$$D(\lambda) = c_0 + c_1\lambda + c_2\lambda^2 + ... + c_n\lambda^n = \sum_{i=0}^{n} c_i \lambda^i \tag{1.22}$$

is a polynomial of order n. The Cayley-Hamilton theorem states that substituting the matrix A for λ in this polynomial (eq. 1.22) results in the null matrix i.e.

$$D(A) = \sum_{i=0}^{n} c_i A^i = 0 \tag{1.23}$$

The theorem can be verified with the following example.

Consider a matrix $A = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}$. The characteristic equation of the matrix is

$$\begin{aligned} D(\lambda) &= \begin{vmatrix} 2-\lambda & -1 \\ 3 & -2-\lambda \end{vmatrix} = 0 \\ \Rightarrow D(\lambda) &= \lambda^2 - 1 = 0 \end{aligned} \tag{1.24}$$

By Cayley-Hamilton theorem, the characteristic eq. 1.24 will be satisfied by the matrix A i.e. $D(A) = A^2 - I = 0$ or, $A^2 = I$. Now,

$$A^2 = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \ \texttt{(q.e.d.)}$$

Cayley-Hamilton theorem is often used to determine the inverse of a matrix. Given the characteristic equation (eq. 1.22) of a matrix A, Cayley-Hamilton theorem implies

$$D(A) = \sum_{i=0}^{n} c_i A^i = c_0 I + c_1 A + c_2 A^2 + ... + c_n A^n = 0 \tag{1.25}$$

Multiplying eq. 1.25 by A^{-1},

$$\begin{aligned} c_0 I A^{-1} + c_1 A A^{-1} + c_2 A^2 A^{-1} + ... + c_n A^n A^{-1} &= 0 \\ \texttt{or,}\ c_0 A^{-1} + c_1 I + c_2 A + ... + c_n A^{n-1} &= 0 \\ \texttt{or,}\ c_0 A^{-1} &= -(c_1 I + c_2 A + ... + c_n A^{n-1}) \\ \Rightarrow A^{-1} = -\frac{1}{c_0}(c_1 I + c_2 A + ... + c_n A^{n-1}) &= -\frac{1}{c_0}\sum_{i=1}^{n} c_i A^{i-1} \end{aligned}$$

Problem 9: *Determine the inverse of the matrix* $A = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}$ *by using the Cayley-Hamilton theorem.*

Solution: The characteristic equation of the matrix is

$$\begin{aligned} \begin{vmatrix} 2-\lambda & 0 & 1 \\ 1 & 1-\lambda & 2 \\ 0 & 1 & 1-\lambda \end{vmatrix} &= 0 \\ \Rightarrow (2-\lambda)\{(1-\lambda)^2 - 2\} + 1 &= 0 \\ \Rightarrow \lambda^3 - 4\lambda^2 + 3\lambda + 1 &= 0 \end{aligned} \tag{1.26}$$

By Cayley-Hamilton theorem, eq. 1.26 will be satisfied by the matrix A itself i.e.

$$A^3 - 4A^2 + 3A + I = 0 \tag{1.27}$$

Multiplying eq. 1.27 by A^{-1},

$$\begin{aligned} A^3 A^{-1} - 4A^2 A^{-1} + 3AA^{-1} + IA^{-1} &= 0 \\ \Rightarrow A^{-1} &= -(A^2 - 4A + 3I) \end{aligned}$$

Now, $A^2 = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 1 & 3 \\ 3 & 3 & 5 \\ 1 & 2 & 3 \end{pmatrix}$

$$\begin{aligned} \therefore A^{-1} = -A^2 + 4A - 3I &= -\begin{pmatrix} 4 & 1 & 3 \\ 3 & 3 & 5 \\ 1 & 2 & 3 \end{pmatrix} + 4\begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} - 3\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 & 1 \\ 1 & 2 & 3 \\ -1 & 2 & 2 \end{pmatrix} \end{aligned}$$

1.6 Diagonalization of Matrices

A diagonal matrix corresponding to a square matrix is a matrix of same order having its diagonal elements as the eigenvalues of the original matrix and all other elements are zero. For example, consider

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & ... & a_{1n} \\ a_{21} & a_{22} & ... & ... & a_{2n} \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & ... & a_{nn} \end{pmatrix}$$

and the eigenvalues of A are λ_1, λ_2, ... λ_n. Therefore, the diagonal matrix of A is

$$D = \begin{pmatrix} \lambda_1 & 0 & ... & ... & 0 \\ 0 & \lambda_2 & ... & ... & 0 \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & ... & \lambda_n \end{pmatrix}$$

If A has n number of linearly independent eigenvectors, a matrix S can be found such that $S^{-1}AS = D$, the diagonal matrix. The matrix S is called the diagonalizing matrix.

Let, X_1, X_2, ... X_n are the linearly independent eigenvectors of A. Thus the diagonalizing matrix

$$S = (X_1 X_2 ... X_n) = \begin{pmatrix} x_{11} & x_{12} & ... & ... & x_{1n} \\ x_{21} & x_{22} & ... & ... & x_{2n} \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ x_{n1} & x_{n2} & ... & ... & x_{nn} \end{pmatrix}$$

where we denote the eigenvectors X_i by the column matrices having elements x_{1i}, x_{2i}, ..., x_{ni}.

Note that the diagonalizing matrix S is not unique as we could arrange the eigenvectors X_1, X_2, ... X_n in any order to construct it. The following steps may be followed to diagonalize a matrix:

- Find the eigenvalues of original matrix.
- Find out corresponding eigenvectors. The eigenvectors must be linearly independent. Otherwise, the matrix will not be diagonalizable.
- Construct the diagonalizing matrix S with its column elements as the linearly independent eigenvectors.
- Determine the inverse matrix S^{-1}.
- The matrix $D = S^{-1}AS$ is the diagonal matrix with λ_1, λ_2, ... λ_n as its successive diagonal elements, where λ_i is the eigenvalue corresponding to the eigenvector X_i.

Problem 10: *Diagonalize the matrix*

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

Solution: Note that the eigenvalues of A are $\lambda_1 = -1$, $\lambda_2 = 1$. Corresponding eigenvectors are $X_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $X_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ respectively. The eigenvectors are linearly independent[2].

Thus the diagonalizing matrix $S = \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix}$

The inverse matrix $S^{-1} = -\frac{1}{2}\begin{pmatrix} 1 & -1 \\ -3 & 1 \end{pmatrix}$

Therefore, the diagonal matrix

$$D = S^{-1}AS = -\frac{1}{2}\begin{pmatrix} 1 & -1 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

1.6.1 Corollaries

1. *Diagonalizing matrix of a real symmetric matrix is orthogonal.*

   Proof: Let us consider a symmetric matrix A i.e. $A^T = A$. If λ_i are the eigenvalues of A and S is the diagonalizing matrix,

$$\begin{aligned} S^{-1}AS &= D = \mathtt{diag}(\lambda_1, \lambda_2, ...\lambda_n) \\ \Rightarrow (S^{-1}AS)^T &= D^T \\ \Rightarrow S^TA^T(S^{-1})^T &= D \\ \Rightarrow S^TA(S^{-1})^T &= S^{-1}AS \\ \Rightarrow S^T &= S^{-1} \\ \Rightarrow S^TS &= I \end{aligned}$$

   i.e. S is an orthogonal matrix.

[2]The task is left for the readers

2. *Diagonalizing matrix of a Hermitian matrix is unitary.*

   Proof: Let us consider a Hermitian matrix H i.e. $H^\dagger = H$. If D is the diagonal matrix and S is the corresponding diagonalizing matrix,

$$\begin{aligned} S^{-1}HS &= D \\ \Rightarrow (S^{-1}HS)^\dagger &= D^\dagger \\ \Rightarrow S^\dagger H^\dagger (S^{-1})^\dagger &= D \\ \Rightarrow S^\dagger H (S^{-1})^\dagger &= S^{-1}HS \\ \Rightarrow S^\dagger &= S^{-1} \\ \Rightarrow S^\dagger S &= I \end{aligned}$$

   i.e. S is unitary.

1.7 Similarity Transformation

Consider a square matrix A of order n and a non-singular matrix S such that $S^{-1}AS = B$, another square matrix of same order as A. Matrix B is similar to A and the transformation from A to B through the relation $S^{-1}AS = B$ is called similarity transformation. Diagonalization is a special type of similarity transformation.

Problem 11: *Eigenvalues of a matrix remain invariant under similarity transformation*

Solution: Consider a similarity transformation $S^{-1}AS = B$. If λ is the eigenvalue of B, the characteristic equation is $|B - \lambda I| = 0$ i.e.

$$\begin{aligned} |S^{-1}AS - \lambda I| &= 0 \\ \Rightarrow |S^{-1}AS - S^{-1}\lambda IS| &= 0 \\ \Rightarrow |S^{-1}(A - \lambda I)S| &= 0 \\ \Rightarrow |S^{-1}||A - \lambda I||S| &= 0 \\ \Rightarrow |S^{-1}S||A - \lambda I| &= 0 \\ \Rightarrow |A - \lambda I| &= 0 \end{aligned}$$

which is the characteristic equation of the original matrix A with same eigenvalue λ. Thus the eigenvalues remain invariant under similarity transformation.

1.8 Unitary Transformation

The similarity transformation may be done by a unitary matrix U. The transformation $U^{-1}AU = B$ is called unitary transformation. Since for a unitary matrix U, $U^{-1} = U^\dagger$; the unitary transformation may be defined as $B = U^\dagger AU$.

Problem 12: *A Hermitian matrix remains Hermitian under unitary transformation.*

Solution: Let A is a Hermitian matrix i.e. $A^\dagger = A$.

The unitary transformation matrix $B = U^{-1}AU = U^\dagger AU$ where U is a unitary matrix ($UU^\dagger = I$ & $U^{-1} = U^\dagger$).

$$\begin{aligned} B &= U^\dagger AU \\ \Rightarrow B^\dagger &= (U^\dagger AU)^\dagger \\ &= U^\dagger A^\dagger (U^\dagger)^\dagger \\ &= U^\dagger AU \quad (\because A^\dagger = A) \\ &= B \end{aligned}$$

Thus B is Hermitian if A is Hermitian.

Problem 13: *The norm of a matrix remains unchanged under the unitary transformation.*

Solution: Consider a matrix A and its unitary transformation matrix $B = U^{-1}AU$, where U is a unitary matrix i.e. $UU^\dagger = U^\dagger U = I$ or $U^{-1} = U^\dagger$. Now,

$$B^\dagger = (U^{-1}AU)^\dagger = (U^\dagger AU)^\dagger = U^\dagger A^\dagger U$$

Multiplying the above equation by $B = U^\dagger AU$ from right

$$\begin{aligned} B^\dagger B &= U^\dagger A^\dagger UU^\dagger AU = U^\dagger A^\dagger AU \\ \Rightarrow |B^\dagger B| &= |U^\dagger||A^\dagger A||U| = |U^\dagger U||A^\dagger A| = |A^\dagger A| \end{aligned}$$

Thus the norm of the matrix remains invariant under the unitary transformation.

1.9 Evaluating Power of a Matrix

Consider diagonalization of a matrix A by the matrix S: $S^{-1}AS = D$ or, $A = SDS^{-1}$. For a function $f(A)$ of matrix A, we have

$$f(A) = Sf(D)S^{-1}, \tag{1.28}$$

where $f(D)$ is similar function of D. Thus from eq. 1.28, for any power A^n of matrix A

$$A^n = SD^nS^{-1} \tag{1.29}$$

Problem 14: $A = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}$. *Find* A^{50}.

Solution: Refer to Problem 10. The diagonal matrix $D = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ and the diagonalising matrix $S = \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix}$. The inverse matrix $S^{-1} = -\frac{1}{2}\begin{pmatrix} 1 & -1 \\ -3 & 1 \end{pmatrix}$

$\therefore$ By eq. 1.29,

$$\begin{aligned}
A^{50} = SD^{50}S^{-1} &= -\frac{1}{2}\begin{pmatrix} 1 & -1 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}^{50}\begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} \\
&= -\frac{1}{2}\begin{pmatrix} 1 & -1 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}$$

1.10 Solutions of Linear Coupled First Order Ordinary Differential Equations

Consider the following pair of linear coupled first order differential equations:

$$\begin{aligned}
y_1'(t) = a_{11}y_1(t) + a_{12}y_2(t) \\
y_2'(t) = a_{21}y_1(t) + a_{22}y_2(t)
\end{aligned}$$

The equations, in matrix form, can be represented as

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \texttt{ i.e. } Y' = AY$$

Let the boundary conditions are $y_1 = c_1$ and $y_2 = c_2$ i.e. $Y(0) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$

Steps:

- Determine the eigenvalues $\lambda_1, \lambda_2, ...$ of the matrix A.
- Find out the corresponding eigenvectors $X_1, X_2, ...$
- The solutions of the coupled equations can be written as $Y(t) = \sum_i a_i e^{\lambda_i t} X_i$, where a_i are arbitrary constants.
- Applying the boundary conditions a_i can be determined and exact solution is obtained.

Let us consider the following set of equations:

$$\begin{aligned}
y_1' &= 2y_1 + 3y_2 \\
y_2' &= 4y_1 + y_2
\end{aligned}$$

The given initial conditions are $y_1(0) = 2, y_2(0) = 1$. The equations, in matrix form, can be represented as

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \texttt{ i.e. } Y' = AY$$

where $A = \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix}$. The eigenvalues of the matrix A are $\lambda_1 = -2$ and $\lambda_2 = 5$. Corresponding eigenvectors are $X_1 = \begin{pmatrix} 3 \\ -4 \end{pmatrix}$ and $X_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ respectively.

Thus, the general solutions are

$$\begin{aligned}
Y(t) &= \sum_i a_i e^{\lambda_i t} X_i \\
\Rightarrow \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} &= a_1 e^{-2t} \begin{pmatrix} 3 \\ -4 \end{pmatrix} + a_2 e^{5t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
\texttt{or,}\ y_1(t) &= 3a_1 e^{-2t} + a_2 e^{5t} \\
\texttt{and}\ y_2(t) &= -4a_1 e^{-2t} + a_2 e^{5t}
\end{aligned}$$

Applying the initial conditions,

$$\begin{aligned}
y_1(0) &= 3a_1 + a_2 = 2 \\
y_2(0) &= -4a_1 + a_2 = 1
\end{aligned}$$

Solving these equations, we have $a_1 = 1$ and $a_2 = -1$. Hence, the exact solutions are

$$\begin{aligned}
y_1(t) &= 3e^{-2t} - e^{5t} \\
y_2(t) &= -4e^{-2t} - e^{5t}
\end{aligned}$$

Bibliography

[1] T. L. Chow, *Mathematical Methods for Physicists: A Concise Introduction* Cambridge University Press (2000)

[2] Charlie Harper, *Introduction to Mathematical Physics* Prentice Hall India (1978)

[3] B. S. Rajput, *Mathematical Physics* Pragati Prakashan (2014)