A Quick Review of Matrices

A.K.Kapoor
School of physics
University of Hyderabad
Hyderabad 500 046

January 27, 2012

Contents

1 Introduction

We recall a few facts from previous lectures on vector spaces. Every finite dimensional vector V (dimension = N) is isomorphic to C^N. This is demonstrated by selecting a basis $\xi = \{e_1, e_2, \cdots, e_N\}$ in the vector space and writing arbitrary vector x as a linear combination of elements in ξ. The coefficients appearing in the linear combination can be assembled in form of a $N-$ component column vector, to be denoted by $\underset{\sim}{\mathrm{x}}$.
The action of a linear operator on vectors in the vector space is determined once we know the action on the elements of any basis. Thus given a linear operator T, knowledge of vectors in the set $T^\xi = \{Te_1, Te_2, \cdots, Te_N\}$, by linearity, determines action of T on an arbitrary vector. The vectors in T^ξ in turn, like every other vector in V, are completely specified by their expansion coefficients in the basis elements ξ. The action of the linear operator T on an arbitrary vector x can be computed from the knowledge of the coefficients which appear when the vectors in T^ξ are expanded in terms of the vectors in ξ. These coefficients can be arranged conveniently as a matrix $\underset{\sim}{\mathrm{T}}$. Thus Given a basis ξ, we represent

(a) vectors x in $V \longrightarrow N$ component columns $\underset{\sim}{x}$

(b) linear operators T on $V \longrightarrow N * N$ matrices $\underset{\sim}{T}$

If a new basis $u = \{u_1, u_2, \cdots, \}$ u selected we will have the representatives of the vectors and operators will change and we will have

(a) vectors x in $V \longrightarrow N$ component columns $\underset{\sim}{x}$

(b) linear operators T on $V \longrightarrow N * N$ matrices $\underset{\sim}{T}$

If S is a linear operator such that $Se_k = u_k, k = 1, 2, \cdots .N$. Then the representatives of vectors and operators w.r.t. the two basis set are related by

$$\underset{\sim}{x} = \underset{\sim}{S}^{-1}\ \underset{\sim}{x}; \qquad \text{and} \qquad \underset{\sim}{T} = \underset{\sim}{S}^{-1}\ \underset{\sim}{T}\ \underset{\sim}{S}$$

Let us note in passing that the columns representing the elements of a basis are given by

$$\underset{\sim}{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix}; \quad \underset{\sim}{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{bmatrix}; \qquad \underset{\sim}{e}_N = \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix}$$

The above discussion applies to any finite dimensional vector space. The above discussion applies to C^N with V replaced by C^N everywhere because C^N can be regarded as a vector space in its own right. Vectors in C^N are $N-$ component column vectors and linear operators in C^N are $N * N$ matrices. In this context I would like to remark that, although two different matrices are different linear operators in C^N, it is sometimes useful to regard similar matrices as representing the same operator w.r.t two different basis sets.

The Vector Space C^N: The correspondence between the notation for abstract space used so far and C^N is as follows.

Vector in C^N: The vectors in C^N are $N-$component column with complex entries. Thus if x and y are vectors in C^N then

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \cdots \\ \xi_N \end{bmatrix}; \qquad y = \begin{bmatrix} n_1 \\ n_2 \\ \cdots \\ n_N \end{bmatrix}$$

Scalar Product: The scalar product of two vectors x and y in C^N is defined as

$$\begin{aligned}(x,y) \longrightarrow x^\dagger\ y &= \left[\ \xi_1{}^*, \xi_2{}^*, \cdots, \xi_N{}^*\ \right] \begin{bmatrix} n_1 \\ n_2 \\ \cdots \\ n_N \end{bmatrix} \\ &= \xi_1^* n_1 + \xi_2^* n_2 + \cdots + \xi_N^* n_N \\ &= \sum_k \xi_k^* n_k\end{aligned}$$

Linear Operators: If T is an $N * N$ complex matrix then the matrix multiplication, T_x, gives another vector in C^N and it defines a linear operator. In fact the set of all $N * N$ matrices is isomorphic to the set of linear operators on C^N. Many properties of linear operators have been studied for linear operators on finite dimensional vector spaces. They

are all applicable to the $N * N$ complex matrices. However, a study of the properties of complex matrices is useful in its own right. This is what we plan to do in the next few lectures. When ever needed, we shall make use of results from the previous lectures on linear operators. In the following we assume the matrices are complex square $N-$ dimensional. $[T]_{ij}$ will be used to denote an element of a matrix T, appearing in the i^{th} row and j^{th} column.

Hermitian Adjoint: The hermitian adjoint of a linear operator, denoted by $T^\dagger$, satisfies the relation

$$(y, T_x) = (T^\dagger y, x) \qquad \text{for all vectors x and y}$$

This relation can be taken as the definition of the adjoint operator $T^\dagger$ in (finite dimensional) vector spaces. Given an $N * N$ matrix how do we define its hermitian adjoint so that it is consistent with the above definition ? We shall now obtain a relation between $[T^\dagger]_{ij}$ and $[T]_{ij}$ using the above relation. We compute

$$\begin{aligned}(y, Tx) &= y^\dagger(Tx) = \left[\ n_1^*, n_2^*, \cdots, +n_N^*\ \right] [T] \begin{bmatrix} \xi_1 \\ \xi_2 \\ \cdots \\ \xi_N \end{bmatrix} \\ &= \sum n_1^* [T]_{ij} \xi_j\end{aligned}$$

and

$$
\begin{aligned}
= (T^\dagger y, x) &= (x, T^\dagger y)^* \\
&= \text{complex conjugate of} \left[\ \xi_1{}^*, \xi_2{}^*, \cdots, \xi_N{}^*\ \right] [T^\dagger] \begin{bmatrix} n_1 \\ n_2 \\ \cdots \\ n_N \end{bmatrix} \\
&= \left[\ \sum \xi_1{}^* [T^\dagger]_{ij} n_j\ \right]^* = \left[\ \sum \xi_1{}^* [T^\dagger]_{ij} n_j\ \right] \\
&= \sum n_1^* [T^\dagger]_{ij}^* \xi_j = \sum n_j^* [T^\dagger]_{ji}^* \xi_i
\end{aligned}
$$

The requirement $(y, T_x) = (T^\dagger y, x)$ should hold for arbitrary vectors. Hence the adjoint of a matrix T should have its elements $[T^\dagger]_{ij}$ related to the matrix elements of T by $[T^\dagger]_{ij}^* = [T]_{ij}$ which is equivalent to

$$[T^\dagger]_{km} = [T]_{mk}^*$$

Basic Definitions:

1. Given a matrix A, its transpose is the matrix obtained by interchanging rows and columns. The transpose of a matrix A will be denoted by A^T. We thus have $[A^T]_{ij} = [A]_{ji}$.

2. The hermitian adjoint of matrix A, denoted by $A^\dagger$, is obtained by taking the complex conjugate of the transpose of the matrix A. Thus $A^\dagger = (A^T)^*$ and $[A^\dagger]_{ij} = [A]_{ji}^*$

3. A matrix X is symmetric if $X^T = X$, or, $[X]_{ij} = [X]_{ji}$

4. A matrix X is anti-symmetric if $X^T = -X$ or, $[X]_{ij} = -[X]_{ji}$

5. A matrix X is hermitian if $X^\dagger = X$, or, $[X]_{ij} = -[X]_{ji}^*$

6. A matrix X is anti-hermitain if $X^\dagger = -X$, or $[X]_{ij} = -[X]_{ji}^*$

7. A matrix X is orthogonal if $X^T X = I$, or, $X^T = X^{-1}$

8. A matrix X is unitary if $X^\dagger X = I$ or ,$X^\dagger = X^{-1}$

9. A matrix X is normal if X commutes with $X^\dagger$ or ,$X^\dagger X - X^\dagger = 0$

10. A matrix X is nilpotent if $X^r = 0$ for some r.

11. A matrix X is called upper (lower) triangular matrix if all its elements below (above) the main diagonal are zero.

12. A matrix X is invertible (or non-singular) if $\det X \neq 0$

13. A matrix X is A and B are called <u>similar</u> if there exists an invertible matrix X such that $A = XBX^{-1}$

Determinants of University and Orthogonal Matrices: If U is a unitary operator then det $U - exp(i\alpha)$ where α is a real number.
Proof: Taking determinant of $U\ U^\dagger = I$ we get

$$\det(U\ U^\dagger) = 1\ \Rightarrow \det\ U\ \ \det(U)^\dagger = 1 \Rightarrow \det(U^{Tx}) = 1$$

$$\Rightarrow \det U (\det U)^* = 1\ \Rightarrow \det\ U = \exp(i\alpha), \alpha \text{ real}$$

Then an orthogonal matrix is has determinant $= \pm 1$ follows in a similar fashion from $00^T = 1$

The Trace: The trace of a matrix is defined as the sum of its diagonal elements

$$\text{Trace}(A) = \sum_{k=1}^{N} [A]_{kk}$$

$Tr(A)$ or $tr(A)$ are used shorthand notation for the trace of a matrix. The trace of product of (finite dimensional) matrices satisfies the cyclic property:

$$tr(AB) = tr(BA);$$

$$tr(A_1 A_2 A_3 \cdots A_r) = tr(A_r A_1 A_2 \cdots A_{r-1}).$$

We shall give the proof of the cyclic property for product of two matrices.

$$tr(AB) = \sum_k (AB)_{kk} = \sum_k \sum_m (A_{km} B_{mk}) = \sum_m \sum_k (A_{km} B_{mk})$$

where we have interchanged the two summations in the last step. Thus we get

$$tr(AB) = \sum_m (BA)_{mm} = tr(BA)$$

For a product of more than two matrices the result follows from $tr(AB) = tr(BA)$ and identifying $A \longrightarrow A_1$ and $B \longrightarrow A_2 A_3 \cdots, A_r$.
Since the proof makes use of interchange of two summations, the cyclic property is not valid for operators in infinite dimensional vector spaces and for infinite dimensional matrices. Thus the commutation relation for

$$\hat{x}\ \hat{p} - \hat{p}\ \hat{x} = i$$

two operators $\hat{x}$ and $\hat{p}$ can not be satisfied in a finite dimensional vector space.
The cyclic property of the trace impels that the trace of two similar matrices is equal. For let A and B be two similar matrices. Then there exists a non-singular matrix X such that $A = XBX^{-1}$

$$tr(A) = tr(XBX^{-1}) = tr(X^{-1}XA) = tr(B)$$

2 Computation Of Eigenvalues And Eigenvectors

Eigenvalues and Eigenvectors: Let A be $N * N$ matrix we say that λ is an eigenvalue of A if there exists a non zero vector u such that $Au = \lambda u$. The vector u is called eigenvector of A corresponding to the eigenvalue λ.
Note that if u is an eigenvector of matrix all scalar multiple, βu, are also eigenvectors with the same eigenvalue.
Let ν be the number of linearly independent eigenvectors u corresponding to a given eigenvalue λ. If $\nu = 1$ we say that the eigenvalue is non degenerate. If $\nu > 1$ the eigenvalue λ is degenerate and then ν to called the degeneracy of the eigenvalue.

If α be a degenerate eigenvalue of A with degeneracy ν. Let the corresponding eigenvectors be $u_1, u_2, \cdots ., u_\nu$. Then by definition we have

$$Au_1 = \alpha u_1; \quad Au_2 = \alpha u_2; \cdots ; \quad Au_\nu = \alpha u_\nu;$$

If we form a linear combination

$$w = \xi_1 u_1 + \xi_2 u_2 + \cdots + \xi_\nu u_\nu$$

where $\xi' s$ are arbitrary complex numbers, then vector w is also an eigenvector of A with the same eigenvalue $\alpha, Aw = \alpha w$.

The set of all eigenvectors corresponding to a fixed eigenvalue of a matrix is a subspace of C^N. The dimension of this subspace is precisely equal to the degeneracy of the eigenvalue.

To find eigenvalue and eigenvectors we start form the equation $Au = \lambda u$ in the form

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

We shall now discuss example of computation of eigenvalues and eigenvectors for four real matrices.

$\boxed{EXAMPLE - 1}$
We shall at first compute the eigenvalues and eigenvectors of the matrix A matrix A where

$$A = \begin{bmatrix} 3 & -5 & -4 \\ -5 & -6 & -5 \\ -4 & -5 & 3 \end{bmatrix}$$

The eigenvalues are given by $\det(A - \lambda I) = 0$

$$\det \begin{bmatrix} 3-\lambda & -5 & -4 \\ -5 & -6-\lambda & -5 \\ -4 & -5 & 3-\lambda \end{bmatrix} = 0$$

Expanding the determinant gives the characteristic equation

$$\lambda^3 - 93\lambda + 308 = 0$$

Factorizing the left hand side we get

$$(\lambda - 4)(\lambda - 7)(\lambda + 11) = 0$$

Thus the characteristic roots are $\lambda = 4, 7, -11$
To find the eigenvectors we start with $(A - \lambda I)x = 0$, Let x be given by

$$x = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}$$

To determine α, β, γ we use $(A - \lambda I)x = 0$. This equation has the form

$$\begin{bmatrix} 3-\lambda & -5 & -4 \\ -5 & -6-\lambda & -5 \\ -4 & -5 & 3-\lambda \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = 0$$

Eigenvectors for $\lambda = 4$: The equation for the eigenvectors becomes

$$\begin{bmatrix} -1 & -5 & -4 \\ -5 & -10 & -5 \\ -4 & -5 & -1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = 0$$

or

$$-\alpha - 5\beta - 4\gamma = 0 \tag{1}$$
$$-5\alpha - 10\beta - 5\gamma = 0 \tag{2}$$
$$-4\alpha - 5\beta - \gamma = 0 \tag{3}$$

Subtracting (18)from (17)gives

$$3\alpha - 3\gamma = 0; \qquad \therefore \alpha = \gamma$$

Subtracting $\alpha = \gamma$ in (17)gives

$$-5\gamma - 5\beta = 0; \qquad \therefore \beta = -\gamma$$

Thus the eigenvector becomes $x = \gamma \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, where γ is arbitrary. Sometimes it is fixed by demanding $x^\dagger\, x = 1$. We shall take the unknown γ to be equal to 1.

Eigenvectors for $\lambda = 7$: The equation determining the eigenvector takes the form

$$\begin{bmatrix} -4 & -5 & -4 \\ -5 & -13 & -5 \\ -4 & -5 & -4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = 0$$

$$-4\alpha - 5\beta - 4\gamma = 0 \tag{4}$$

$$-5\alpha - 13\beta - 5\gamma = 0 \tag{5}$$

$$-4\alpha - 5\beta - 4\gamma = 0 \tag{6}$$

$4 * (18) - 5 * (19) = 0 \Rightarrow \beta = 0$. Substituting $\beta = 0$ in (17) gives $\alpha = -\gamma$.

Thus we obtain $x = \begin{bmatrix} -\gamma \\ 0 \\ \gamma \end{bmatrix}$ which, on taking $\gamma = -1$, becomes $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$

Eigenvectors for $\lambda = -11$: The eigenvector x is determined from the equation

$$\begin{bmatrix} 14 & -5 & -4 \\ -5 & 5 & -5 \\ -4 & -5 & 14 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = 0$$

$$-14\alpha - 5\beta - 4\gamma = 0 \tag{7}$$

$$-5\alpha - 5\beta - 5\gamma = 0 \tag{8}$$

$$-4\alpha - 5\beta 14\gamma = 0 \tag{9}$$

The equation (18) is $\alpha + \beta + \gamma = 0$ and (17)-(19) implies $\alpha = \gamma$, hence $\beta = 2\gamma$ Thus we obtain $x = \begin{bmatrix} \gamma \\ 2\gamma \\ \gamma \end{bmatrix}$ which, on taking $\gamma = -1$, becomes $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ Thus the final results for the eigenvalues and the eigenvectors are

$$\gamma = 4, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}; \qquad \lambda = \gamma, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}; \qquad \lambda = -11, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$\boxed{EXAMPLE - 2}$

As a second example we take the matrix

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$$

The eigenvalues are given by $\det(A - \lambda I) = 0$.

$$\begin{bmatrix} 5-\lambda & -6 & -6 \\ -1 & 4-\lambda & 2 \\ 3 & -6 & -4-\lambda \end{bmatrix} = 0$$

Expanding we get the characteristic polynomial to be $\lambda^3 - 5\lambda + 8\lambda - 4$ which can be factorized to become $(\lambda - 2)^2(\lambda - 1) = 0$. Thus the characteristic roots are $\lambda = 1$, and 2 (repeated twice).

To find the eigenvectors we again start with $(A - \lambda I)x = 0$.

$$\begin{bmatrix} 5-\lambda & -6 & -6 \\ -1 & 4-\lambda & 2 \\ 3 & -6 & -4-\lambda \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = 0$$

Eigenvectors for $\lambda = 1$: The equation for the eigenvector becomes

$$\begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = 0$$

which is equivalent to the following equations

$$4\alpha - 6\beta - 6\gamma = 0 \tag{10}$$

$$-\alpha + 3\beta + 2\gamma = 0 \tag{11}$$

$$3\alpha - 6\beta - 5\gamma = 0 \tag{12}$$

(17)-(19) implies $\gamma = \alpha$ this together with (18) gives $\gamma = -3\beta$. Taking $\beta = 1$ we get the eigenvector to be $\begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$.

Eigenvectors for $\lambda = 2$

$$\begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = 0$$

This is equivalent to the following three equations

$$3\alpha - 6\beta - 6\gamma = 0 \tag{13}$$

$$-\alpha + 2\beta + 2\gamma = 0 \tag{14}$$

$$3\alpha - 6\beta - 6\gamma = 0 \tag{15}$$

It is seen that the three equations are not independent, and that there is only one independent relation between α, β, γ which we take to be $\alpha = 2\beta + 2\gamma$. With β and γ remaining arbitrary the eigenvector takes the form

$$x = \begin{bmatrix} 2\beta + 2\gamma \\ \beta \\ \gamma \end{bmatrix} \tag{16}$$

There are two ways of proceeding.

1. We write the eigenvector as

$$x = \begin{bmatrix} 2\beta + 2\gamma \\ \beta \\ \gamma \end{bmatrix} = \beta \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

   This clearly shows that an eigenvector for the case $\lambda = 2$ is a linear combination of two linearly independent vectors

$$\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

2. We can take two sets of values for β and γ in such a way as to obtain two linearly independent eigenvectors.

   Thus can takes,$\beta = 1, \gamma = 0$ giving $\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ and next we select $\beta = 0, \gamma = 1$ giving $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$

Example-3

Our next example is

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

The eigenvalues are determined from $\det(A - \lambda I) = 0$ or from

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

The characteristic polynomial is found to be

$$\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$$

or

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

Thus the eigenvalues are $\lambda = 1, 2$. The root $\lambda = 2$ is a double root. The corresponding eigenvectors are obtained by solving for α, β, and γ from

$$\begin{bmatrix} 3-\lambda & 1 & -1 \\ 2 & 2-\lambda & -1 \\ 2 & 2 & -\lambda \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = 0$$

Eigenvectors for $\lambda = 1$

$$\begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = 0$$

This is equivalent to the three equations

$$2\alpha + \beta - \gamma = 0$$

$$2\alpha + \beta - \gamma = 0$$

$$2\alpha + 2\beta - \gamma = 0$$

These equations are solved to give$\gamma = 2\alpha$ and $\beta = 0$. Taking $\alpha = 1$, we find the eigenvector for $\lambda = 1$ to be $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$.

Eigenvectors for $\lambda = 2$.

$$\begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = 0$$

In this case the equations which determine α, β and γ are

$$\alpha + \beta - \gamma = 0 \tag{17}$$

$$2\alpha - \gamma = 0 \tag{18}$$

$$2\alpha + 2\beta - 2\gamma = 0 \tag{19}$$

Subtracting (17) and (19) from (18)implies, respectively, $\alpha = \beta$ and $\gamma = 2\beta$.

Therefore the eigenvector is found to be $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.

$\boxed{EXAMPLE-4}$
In the fourth example we will find that all the three roots of the characteristic polynomial are degenerate and that there is only one linearly independent eigenvector.

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

The eigenvalues are given by $\det(A - \lambda I) = 0$ which takes the form

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

Expanding the determinant gives $(\lambda - 2)^3 = 0$. In this case there is only one root, $\lambda = 2$, of the characteristic polynomial and the root is repeated thrice. To find the eigenvectors we start with $(A - \lambda I)x = 0$. This equation for $\lambda = 2$ is

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = 0$$

These equations give $\beta = 0$ and $\gamma = 0$ and α is not fixed. Thus we get $x = \begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix}$ and there is only one linearly independent eigenvector which can be taken to be $x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

3 Properties Of Eigenvalues And Eigenvectors

In this lecture we discuss with some simple properties of eigenvalues and eigenvectors.

Properties of eigenvalues and eigenvectors

1. If two matrices A and B are similar and λ is an eigenvalue of A then it is also an eigenvalue of the matrix B.

2. If a matrix T is not invertible if $\lambda = 0$ is an eigenvalue of T.

3. If A is a matrix, with $\det A \neq 0$ and if λ is an eigenvalue of A, then λ^-1 is an eigenvalue of A^-1

4. Let A be a matrix and letλ be its eigenvalue. Let $P(A)$ be a polynomial in A. Then $P(\lambda)$ is an eigenvalues of $P(A)$

5. Let A be a triangular matrix. The eigenvalues of A coincide with the elements on the main diagonal.

6. If D is a diagonal matrix, eigenvalues of D are its diagonal elements.

7. Product of all eigenvalues is equal to the determinant of the matrix and the sum of all eigenvalues is equal to the trace of the matrix. If the characteristic polynomial for the matrix has root repeated m times, the corresponding eigenvalue must appear m times in the product or the sum.

8. For a hermitian matrix we have the following results.
   (a) The eigenvalues are real.
   (b) The eigenvectors corresponding to different eigenvalues are orthogonal.

9. For a unitary matrix we have the following results.
   (a) The eigenvalues are pure phase, i.e., of the form $exp(i\alpha)$ where α is a real number.
   (b) The eigenvectors corresponding to different eigenvalues are orthogonal.

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

Proofs:

First Proof:

1. A and B are two similar matrices. Therefore there exists a non-singular matrix X such that

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

The eigenvalues of the matrix are the roots of the characteristic polynomial $\det(A - \lambda I)$. Consider

$$\begin{aligned}\det(A - I) &= \det(XBX^{-1} - \lambda I) \\ &= \det\{X(B - \lambda I)X^{-1}\} \\ &= \det X.\det(B - \lambda I).\det X^{-1} \\ &= \det(B - \lambda I)\end{aligned}$$

The characteristic polynomials for the matrix A and B therefore equal. Hence the two matrices have set of eigenvalues.

Second Proof: Let u be an eigenvector of A with eigenvalue λ. Then $Au = \lambda u$. Substituting for A from $A = XBX^-1$ we get

$$XBX^{-1}u = \lambda u, \quad \text{or} \quad BX^{-1}u = \lambda^{-1}u$$

If we define $v = X^-1u$ we get $Bv = \lambda v$. Therefore λ is an eigenvalue B with eigenvector $(X^{-1}u)$.

2. A matrix T is singular (not invertible) if and only if $\det T = 0$. For a singular matrix T we, therefore, see that the characteristic polynomial $p(\lambda) = \det(T - \lambda I)$ vanishes for $\lambda = 0$. Hence 0 is an eigenvalue of the matrix.

3. Let A be a matrix which is invertible. Then $\det A \neq 0$ and $\lambda = 0$ is not an eigenvalue. Hence if λ_0 is an eigenvalue of A, then $\lambda \neq 0$ and the characteristic polynomial, $p(\lambda) = \det(A - \lambda I)$, for A will vanish.

$$\begin{aligned}
\det(A - \lambda_0 I) = 0 \; &\Leftrightarrow \; \det\{A(I - \lambda_0 A^{-1})\} = 0 \\
&\Leftrightarrow \; \det A \det \lambda_0\{(1/\lambda_0)I - A^{-1}\} = 0 \\
&\Leftrightarrow \; \det\{(1/\lambda_0)I - A^{-1}\} = 0 \\
&\Leftrightarrow \; \det\{A^{-1} - (1/\lambda_0)I\} = 0
\end{aligned}$$

The last result means that $\det(A^{-1} - \lambda I)$. which is the characteristic polynomial for A^{-1}, vanishes when $\lambda = 1/\lambda_0$ showing that $1\lambda_0$ is an eigenvalue of A^{-1}.

4. Let us assume that

$$p(A) = a_0 + a_1 A + a_2 A^2 + \cdots . + a_n A^n$$

Let λ be an eigenvalue of the matrix A and the corresponding eigenvector be u. So that

$$\begin{aligned}
Au &= \lambda u, \\
A^2 u &= A(\lambda u) = \lambda A u = \lambda^2 u \\
A^3 u &= A(\lambda^2 u) = \lambda^2 A u = \lambda^3 u \\
\cdots & \quad \cdots \quad \cdots \\
A^n u &= \lambda^n u \\
(a_0 + a_1 A + a_2 A^2 + \cdots + a_n A^n)u &= (a_0 + a_1 \lambda + a_2 \lambda^2 + \cdots + a_n \lambda^n)u
\end{aligned}$$

(or)

$$P(A)u = P(\lambda)u$$

Thus it has been proved that the vector u is an eigenvector of the matrix $P(A)$ and that $P(\lambda)$ is an eigenvalue of the matrix $P(A)$.

5. Let T be a triangular matrix, upper triangular for definiteness.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ 0 & 0 & \cdots & \cdots \\ 0 & 0 & 0 & a_{nn} \end{bmatrix}$$

The eigenvalues are given by

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

Expanding along the first coloum we get

$$\begin{aligned} \det(A - \lambda I) &= (a_{11} - \lambda) \begin{bmatrix} a_{22} - \lambda & a_{23} & \cdots & a_{2N} \\ 0 & a_{33} - \lambda & \cdots & a_{3N} \\ 0 & 0 & 0 & a_{NN} - \lambda \end{bmatrix} \\ &= (a_{11} - \lambda)(a_{22} - \lambda) \cdots | \cdots | \\ &= (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) \end{aligned}$$

Therefore, the eigenvalues, determined from $\det(A-\lambda I) = 0$, are the elements $a_{11}, a_{22}, \cdots, a_{nn}$ on the main diagonal.

6. This is a special case of the result [5].

7. The proof is obtained by expanding $\det(A - \lambda I) = 0$

8. The result was proved earlier for hermitian operators and is applicable to the hermitian matrices which are hermitian linear operators on complex vector spaces.

9. As the unitary matrices are also unitary operators on C^N the result on the eigenvalues and the eigenvectors follows from the corresponding result for unitary operators on any complex inner product space.

10. Let $u_1, u_2, \cdots, u_r$ be the eigenvectors of a matrix A with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_r$. Let the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_r$ be all distinct. We have to prove that $u_k, k = 1, \cdots n$ are linearly independent. We start with

$$\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + \cdots + \alpha_r u_r = 0 \tag{20}$$

We define $P_1(A) = (A - \lambda_2) \cdots (A - \lambda_r) = \Pi^r_{1 \neq 1}(\lambda_1 - \lambda_k)$. Recall that from result [4] above, $Au = \lambda u$ implies $P(A)u = P(\lambda)u$ where P is a polynomial. Therefore it follows.

$$\begin{array}{lcl} P_1(A)u_1 & = & P_1(\lambda_1)u_1 = (\lambda_1-\lambda_2)(\lambda_1-\lambda_3)\cdots(\lambda_1-\lambda_n)u_1 = P(\lambda_1)u_1 \\ P_1(A)u_2 & = & P_1(\lambda_2)u_2 = (\lambda_2-\lambda_2)(\lambda_2-\lambda_3)\cdots(\lambda_2-\lambda_n)u_2 = 0 \\ P_1(A)u_3 & = & P_1(\lambda_3)u_3 = (\lambda_3-\lambda_2)(\lambda_3-\lambda_3)\cdots(\lambda_3-\lambda_n)u_3 = 0 \\ \cdots & & \cdots \qquad \cdots \\ P_1(A)u_r & = & P_1(\lambda_r)u_r = (\lambda_r-\lambda_2)(\lambda_r-\lambda_3)\cdots(\lambda_r-\lambda_r)u_r = 0 \end{array}$$

Thus we have the result that $P_1(A)u_j = 0$ for $j = 2, \cdots, r$. Multiplying (20) with $P_1(A)$ gives

$$\alpha_1 P_1(\lambda_1)u_1 = 0$$

because all $\lambda_i \neq \lambda_j . P_1(\lambda_1)$ is not zero hence $\alpha_1 = 0$. Similarly we can prove that $\alpha = \cdots = \alpha_n = 0$ by multiplying (20) successively by $P_j(A)$ for $j = 2 \cdots n$, where $P_j(A) = \Pi^n_{1\neq j}(A - \lambda_k)$. This proves that the eigenvectors of a matrix with distinct eigenvalue are linearly independent. The above result [10] implies that if for an $N * N$ matrix the characteristic polynomial has N distinct roots then the matrix has N-linearly independent eigenvectors. We shall give some other sufficient conditions for an $N * N$ matrix to have N linearly independent eigenvectors.

4 Diagonalization

In a finite dimensional vector space, every operator can be represented by a matrix. The values of the elements of the matrix representing an operator will obviously depend on the choice of basis. Let A be an operator in an N dimensional complex vector space. Let A be such that the eigenvectors of A form a basis. Let us denote the eigenvectors by $u_1, u_2, \cdots, u_N$. Then

$$\begin{array}{rcl} Tu_1 = \lambda_1 u_1 & = & \lambda_1 u_1 + 0.u_2 + 0.u_3 + \cdots + 0.u_N \\ Tu_2 = \lambda_2 u_2 & = & 0.u_1 + \lambda u_2 + 0.u_3 + \cdots + 0.u_N \\ Tu_3 = \lambda_3 u_3 & = & 0.u_1 + 0.u_2 + \lambda u_3 + \cdots + 0.u_N \\ \cdots & & \cdots \qquad \cdots \qquad \cdots \\ Tu_N = \lambda_N u_N & = & 0.u_1 + 0.u_2 + 0.u_3 + \cdots + \lambda_N u_N \end{array}$$

We can now construct the matrix representing the operator T w.r.t. the basis $\{u_1, u_2, \cdots, u_N\}$. This is given by the transpose of the matrix of coefficients in the above equations. Thus the operator T is represented by a diagonal matrix:

$$T \longrightarrow \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & & 0 \\ 0 & & \lambda_3 & & 0 \\ 0 & & & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

Definition: Let A be a matrix if there expiates a non singular matrix X such that

$$X^{-1}AX = A,$$

where A is a diagonal matrix, we say that the matrix A can be diagonalized and that the matrix X diagonalizes A.

Recall that whenever two matrices are similar they represent the same operator with respect to two different basis sets. Thus a matrix A can be diagonalized is equivalent to saying that there exists a basis such that the corresponding operator takes a simple diagonal form. From the above discussion it follows that a matrix can be diagonalized if and only if its eigenvectors form a basis set. We shall be concerned here with general conditions under which a matrix has enough linearly independent eigenvectors to get a basis.

We have already seen in the previous lecture that the eigenvectors of a matrix corresponding to distinct eigenvalue are linearly independent. Thus if the characteristic polynomial $p(\lambda) = \det(A - \lambda I)$ has N distinct roots, the matrix will have N linearly independent eigenvector and these will form a basis (because the no. LI vectors =N = dimension of space). Selecting the eigenvectors as the basis will be related to the original form by a similarity transformation. Thus we shall have

$$X^1AX = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ 0 & & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_N$ are the eigenvalues of the matrix A. To summarize we have the following result.

Theorem: A matrix can be diagonalized if all the roots of is characteristic polynomial are distinct.

This converse of this theorem is not always true. There are matrices for which the characteristic polynomial does not have distinct roots but the matrix can be diagonalized. In this connection we state a useful result.

Theorem: If a matrix is normal then it can be diagonalized.

Hermitian, anti-hermitian and unitary matrices are important subsets of set of all normal matrices.
A real symmetric matrix is hermitian and hence can always be diagonalized. A real antisymmetric matrix is anti-hermitian and therefore can be diagonalized. A real orthogonal matrix is unitary and hence normal and therefore can always be diagonalized.

It is easy to give example of matrices which cannot be diagonalized. See the matrices of Example (3) and Example (4) in Lecture-2. The eigenvectors of these matrices do not form a basis and hence they cannot be brought to a diagonal form by a change of basis. The simplest possible forms, other than the diagonal form,to which a matrix can be brought by a change of basis can be classified. These forms are known as Jordan canonical forms. Results on this topic of Jordan canonical forms will not be discussed here.

Suppose a matrix A can be diagonalized. How do we find the matrix X which diagonalizes A($i.e.$satisfies $X^{-1}AX$ = diagonal)?. The method for

Lecture - 4 Diagonalization
doing this will be given by means of an example.
Let is consider the matrix

$$A = \begin{bmatrix} 3 & -5 & -4 \\ -5 & -6 & -5 \\ -4 & -5 & 3 \end{bmatrix}$$

We wish to find the matrix which diagonalizes the above matrix. We first find eigenvalues and eigenvectors. This was done in Lecture -2. The eigenvalues and eigenvectors were found to be

$$\lambda = 4, \quad u_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}; \lambda = 7, \quad u_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}; \lambda = -11, \quad u_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Note that the eigenvalues are all distinct. Thus the eigenvectors form basis. The columns of the matrix X are simply the three eigenvectors. Thus

$$X = \begin{bmatrix} u_1, u_2, u_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 1 & -1 & 3 \end{bmatrix}$$

It is easy to verify that

$$X^{-1}\begin{bmatrix} 3 & -5 & -4 \\ -5 & -6 & -5 \\ -4 & -5 & 3 \end{bmatrix} \quad X = \begin{bmatrix} 4 & & \\ & 7 & \\ & & -11 \end{bmatrix}$$

as it should be. Note that the matrix X is not unique. It is possible to perform operators on X such as

(a) interchange of columns

(b) multiplying X by a non-zero complex number

to get a new matrix $\tilde{X}$ which will also diagonalize the given matrix A.
To summarize the following classes of $N * N$ matrices can be diagonalized.

(a) matrices with N distinct eigenvalues

(b) hermitian matrices

(c) anti hermitian matrices

(d) unitary matrices

(e) real symmetric matrices

(f) real antisymmetric matrices

(g) real orthogonal matrices

(h) normal matrices

What can be said about the matrix X which diagonalizes a given matrix A? In this connection we state the following results.

1 The matrix, which diagonalizes a hermitian matrix, can be chosen to be unitary.

2 For a unitary matrix also the result [1] is true. The matrix which diagonalizes a unitary matrix can be chosen to unitary.

3 A real symmetric matrix is hermitian and as a special cases of [1], it can always be diagonalized by a unitary matrix, Also in this case the matrix which diagonalizes a real symmetric matrix can be chosen to be real orthogonal.

Remark: It is not correct to say that the matrix which diagonalizes a real matrix will always be real. For example, the matrix which diagonalizes the matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

is a complex matrix and cannot be chosen to be real.

Functions Of Matrix: Let A be matrix which can be diagonalized. There exists a matrix X such that

$$X^{-1}AX = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ 0 & & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix} \text{or, } A = X \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ 0 & & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix} X^{-1}$$

where λ_i are the eigenvalues of the matrix A. The given a function $f(.)$, define $f(A)$ so as to satisfy

$$X^{-1}f(A)X = \begin{bmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & & 0 \\ 0 & & \vdots & \vdots \\ 0 & 0 & \cdots & f(\lambda_N) \end{bmatrix}$$

Thus we arrive at the following definition of function, $f(A)$, of matrix A

$$f(A) = X \begin{bmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & & 0 \\ 0 & & \vdots & \vdots \\ 0 & 0 & \cdots & f(\lambda_N) \end{bmatrix} X^{-1}$$

If the function $f(.)$ has a power series expansion (e.g. the exponential function) the series can be used to define the matrix function, $f(A)$, of matrix. For example,

$$\exp(A) = 1 + A + A^2/2! + A^3/3! + \cdots$$

whenever both methods are applicable they give the same answers.

As with the functions of real or complex numbers some functions can be defined only if the matrix satisfies additional conditions. For example, $\log A$ or $\sqrt{A}$ etc. can be defined only for positive definite matrices.

Caley Hamilton Theorem: The Caley Hamilton theorem states that every matrix satisfies its own characteristic equation.

Let $p(\lambda) = \det(A - \lambda I) = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \cdots + \alpha_N\lambda^N$ be the characteristic polynomial of the matrix A. The characteristic equation is $p(\lambda) = 0$. The Caley Hamilton theorem asserts that

$$p(A) = 0$$

(or,)

$$\alpha_0 + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_N A^N = 0$$

is satisfied as a matrix equation.

Proof: Let $B = adj(A - \lambda I)$ denote the adjoint matrix formed the cofactors of the matrix $A - \lambda I$. Then B can be written as

$$B = B_0 + B_1\lambda + B_2\lambda^2 + \cdots + B_{N-1}\lambda^{N-1}$$

where $B_0, B_1, B_2, \cdots, B_{N-1}$ are matrices which do not depend on λ. Note that the elements of B are polynomials in λ of max degree N-1 only. By the definition of adjoint of a matrix we have

$$B(A - \lambda I) = \det(A - AI).I$$

Substituting for B we get

$$(B_0 + B_1\lambda + B_2\lambda^2 + \cdots + B_{N-1}\lambda^{N-1}).(A - \lambda I) = \det(A - AI).I$$

Comparing powers of λ on both sides we get

$$\begin{aligned} B_0 A &= \alpha_0 I \\ B_1 A - B_0 &= \alpha_1 I \\ B_2 A - B_1 &= \alpha_2 I \\ B_3 A - B_2 &= \alpha_3 I \\ \cdots &= \cdots \quad \cdots \\ -B_{N-1} &= \alpha_N I \end{aligned}$$

Multiplying these equations successively on the right by $I, A, A^2, \cdots, A^N$ we and on adding get

$$(B_1 A - B_0) + (B_2 A - B_1)A + (B_3 A - B_2)A^2 + \cdots, -B_{N-1}A^N$$
$$= \alpha_0 + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_N A^N$$

All the terms on the left hand side cancel pairwise and we get the desired result.

$$\alpha_0 + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_N A^N = 0$$