DEPARTMENT OF PHYSICS
INDIAN INSTITUTE OF TECHNOLOGY, MADRAS

PH1020 Physics II **April 2020**

Eigenvalues, Eigenvectors and Diagonalizing Matrices

Suresh Govindarajan, Department of Physics, IIT Madras

In this note we will focus on the properties of linear operators acting on an LVS $\mathbb{V}$ of dimension n. Let $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ be a basis for $\mathbb{V}$. We begin with a recap of ideas from Lecture 6.

1. To each linear operator T, we associated a matrix $\mathbf{T}$ in the following fashion.
$$T \cdot \mathcal{B} := (T(\hat{e}_1), T(\hat{e}_2), \ldots, T(\hat{e}_n)) =: \mathcal{B} \cdot \mathbf{T} \ .$$
The i-th column of the matrix $\mathbf{T}$ is the column vector $T(\hat{e}_i)$ in the basis $\mathcal{B}$.

   Example 1: Let $\mathbb{V} = \mathcal{P}_2(x)$, the space of degree ≤ 2 polynomial in x. Consider the monomial basis: $\mathcal{B} = (1, x, x^2)$. Then the matrix for the linear operator d/dx is obtained as follows.
$$\frac{d}{dx} \cdot \overbrace{(1, x, x^2)}^{\mathcal{B}} = (0, 1, 2x) = \overbrace{(1, x, x^2)}^{\mathcal{B}} \cdot \overbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}}^{\text{The matrix for } d/dx}$$

2. Let $\mathbb{V}$ be an inner product space with an ON basis (in Dirac notation) $\mathcal{B} = (|\hat{e}_1\rangle, |\hat{e}_2\rangle, \ldots, |\hat{e}_n\rangle)$. Then, the matrix for T is given by
$$\mathbf{T} = (T_{ab}) = \begin{pmatrix} T_{11} & T_{12} & \cdots & T_{1n} \\ T_{21} & T_{22} & \cdots & T_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{n1} & T_{n2} & \cdots & T_{nn} \end{pmatrix} \quad \text{where} \quad T_{ab} = \langle \hat{e}_a | \ T \ | \hat{e}_b \rangle = \langle \hat{e}_a | T(\hat{e}_b) \rangle \ .$$

3. Under a change of basis:
$$\mathcal{B}' = (|\hat{e}_1'\rangle, |\hat{e}_2'\rangle, \ldots, |\hat{e}_n'\rangle) = \mathcal{B} \cdot S \ ,$$
the matrix $\mathbf{T}$ changes to $\mathbf{T}' = S \cdot \mathbf{T} \cdot S^{-1}$. As the two bases are ON, then S is unitary i.e., $S^\dagger = S^{-1}$ or $S \cdot S^\dagger = I$.

In physics, we are interested in properties of linear operators that do **not** depend on the choice of basis.

- Can we identify properties of the matrix $\mathbf{T}$ that are basis independent?
- Can we use the freedom in choosing the basis to bring the matrix $\mathbf{T}$ to a simple form?

We will address these two questions next.

1. Eigenvalues and Eigenvectors of Linear Operators

Definition: A vector $|\mathbf{v}\rangle \in \mathbb{V}$ is an **eigenvector** of the linear operator T with **eigenvalue** λ if

$$T\,|\mathbf{v}\rangle = \lambda\,|\mathbf{v}\rangle\ . \tag{1}$$

In the basis $\mathcal{B}$ this becomes the matrix equation

$$\mathbf{T}\cdot\mathbf{v} = \lambda\,\mathbf{v} \quad \text{where} \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}\ . \tag{2}$$

Suppose, we find a solution to the above equation, then it is easy to see that in another basis, one has $\mathbf{T}'\cdot\mathbf{v}' = \lambda\,\mathbf{v}'$ with $\mathbf{v}' = S\cdot\mathbf{v}$. The eigenvalue is clearly **basis independent** even though the form of the eigenvector changed.

Example 1: Let R denote the matrix of the linear operator on $\mathbb{R}^3$ corresponding to a rotation by an angle θ about the z-axis. The matrix of the operator in the standard Cartesian basis is

$$\mathbf{R} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}\ .$$

It is easy to see that $\mathbf{v} = (0,0,1)^T$ is an eigenvector of $\mathbf{R}$ with eigenvalue 1. This says that **all** vectors parallel to the z-axis remains unchanged under the rotation. More generally, an arbitrary rotation in $\mathbb{R}^3$ is a rotation by an angle θ about an axis, $\hat{n}$ which is an eigenvector of the rotation with eigenvalue unity.

Given a linear operator T, we wish to determine all solutions to Eq. (2). Note that both the eigenvalue and eigenvector are undetermined at the moment and hence the RHS of the equation is not fixed. With this in mind, we rewrite Eq. (2) as

$$\mathbf{T}_\lambda\cdot\mathbf{v} := (\mathbf{T} - \lambda\,I)\cdot\mathbf{v} = 0\ . \tag{3}$$

where I is the $n\times n$ identity matrix. When $\mathbf{T}_\lambda$ is an invertible matrix, then by multiplying the equation by its inverse, $\mathbf{T}_\lambda^{-1}$, we see that the only solution is $\mathbf{v} = 0$. This is called a **trivial** solution. We conclude that in order that the equation admits **non-trivial** solutions, we need $\det(\mathbf{T}_\lambda) = 0$ – such a matrix is called **singular**. We thus obtain a condition

$$P_T(\lambda) = \det(\mathbf{T}_\lambda) = \det(\mathbf{T} - \lambda\,I) = 0\ . \tag{4}$$

1. $P_T(\lambda)$ is a polynomial of degree n in λ.

2. One can see that the characteristic polynomial is basis independent.

$$\det(\mathbf{T}'-\lambda I) = \det\left(S^{-1}\cdot(\mathbf{T}-\lambda I)\cdot S\right) = \det(S^{-1})\det(\mathbf{T}-\lambda I)\det(S) = \det(\mathbf{T}-\lambda I)\ .$$

   Above we have used two properties of the determinant. For any three matrices A, B, C, one has $\det(ABC) = \det(A)\det(B)\det(C)$ and for invertible A, $\det(A)\det(A^{-1}) = 1$. It is called the **characteristic** polynomial of the linear operator T.

3. Setting $\lambda = 0$, we obtain $p_T(0) = \det(T)$. More generally, one has

$$p_T(\lambda) = (-1)^n\,\lambda^n + \cdots - \mathrm{Tr}(T)\,\lambda + \det(T)\ ,$$

   where Tr stands for the trace of a matrix – this is defined to be the sum of all elements on the diagonal.

Solutions of Eq. (4). the characteristic equation, are the only possible eigenvalues of the linear operator T.

Example 2: The characteristic equation of the rotation matrix is

$$p_R(\lambda) = -\lambda^3 + (1+2\cos\theta)\lambda^2 - (1+2\cos\theta)\,\lambda + 1 = (1-\lambda)(\lambda^2 - 2\cos\theta + 1)\ .$$

It has one real eigenvalue 1 and two eigenvalues $e^{+i\theta}$ and $e^{-i\theta}$ which is generically complex except when $\theta = 0, \pi$. Thus, we see that the rotation matrix generically has only one **real** eigenvector. However, if treat it as a linear operator on a complex LVS, then we solve Eq. (2) to obtain eigenvectors for the complex eigenvalues. We obtain the following eigensystem for R:

$$\left[\left(e^{i\theta},\ \tfrac{1}{\sqrt{2}}(1,i,0)^T\right)\ ,\quad \left(e^{-i\theta},\ \tfrac{1}{\sqrt{2}}(1,-i,0)^T\right)\ ,\quad \left(1,(0,0,1)^T\right)\right]$$

The eigenvectors have been normalized to have unit norm.

Example 3: Consider the upper triangular matrix

$$\mathbf{T} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}\ .$$

Then $p_T(\lambda) = (\lambda-3)^2$[2]. It has eigenvalue 3 which occurs with **multiplicity** two. However, it has only **one** eigenvector $\mathbf{v} = (1,0)^T$. This is different from what happened with the rotation matrix where we found eigenvectors for the complex eigenvalues by working with a complex LVS. Here the eigenvalues are real and going to the complex field doesn't help us.

2. Normal matrices

The general expectation is that a linear operator will have as many (linearly independent) eigenvectors as eigenvalues. However, Example 3 clearly shows that the expectation is incorrect. However, there is a large class of matrices for which this expectation holds. These are called **normal** operators.

Definition: A linear operator/matrix T is **normal** if

$$T\,T^\dagger = T^\dagger\,T \quad \text{or} \quad \mathbf{T}\cdot\mathbf{T}^\dagger = \mathbf{T}^\dagger\cdot\mathbf{T}\ . \tag{5}$$

We say that T commutes with $T^\dagger$ – the order in which they act doesn't matter.

Exercise: Show that unitary and hermitian matrices are normal.

Properties of normal operators: Let T be a normal operator.

1. The characteristic equation of T and its adjoint are related by complex conjugation. So if λ is an eigenvalue of T, then λ^* is an eigenvalue of $T^\dagger$.

2. We can say something stronger. If $|v\rangle$ is an eigenvector of T with eigenvalue λ, then it is also an eigenvector of $T^\dagger$ with eigenvalue λ^*. Define

   $$T_\lambda := (T - \lambda I)\ . \tag{6}$$

   This is also normal if T is normal. Suppose that $T_\lambda|\mathbf{v}\rangle = 0$, Hence,

   $$||(T_\lambda)^\dagger|\mathbf{v}\rangle||^2 = \langle\mathbf{v}|T_\lambda(T_\lambda)^\dagger|\mathbf{v}\rangle = \langle\mathbf{v}|(T_\lambda)^\dagger\ T_\lambda|\mathbf{v}\rangle = 0.$$

   In obtaining the second equality above, we make use of the normality of T. Since the only vector which has zero norm is the zero vector, we get

   $$(T_\lambda)^\dagger|\mathbf{v}\rangle = 0 \quad \text{or} \quad T^\dagger|\mathbf{v}\rangle = \lambda^*|\mathbf{v}\rangle\ ,$$

   since $T_\lambda^\dagger = (T^\dagger - \lambda^* I)$.

3. If T is hermitian, then it follows that $\lambda = \lambda^*$ and hence all eigenvalues are real. This is very important in quantum mechanics, where we will see that physical quantities are represented by hermitian operators.

4. Let $|v\rangle$ and $|w\rangle$ be eigenvectors of T with eigenvalues λ and $\mu \neq \lambda$ respectively. Then, the two eigenvectors are orthogonal to each other, i.e., $\langle v|w\rangle = 0$.

   $$\langle\mathbf{v}|T|\mathbf{w}\rangle = \mu\ \langle\mathbf{v}|\mathbf{w}\rangle \quad \text{and} \quad \Big(\langle\mathbf{w}|T^\dagger|\mathbf{v}\rangle = \lambda^*\langle\mathbf{w}|\mathbf{v}\rangle\Big)^*$$

   Since $\big(\langle\mathbf{w}|T^\dagger|\mathbf{v}\rangle\big)^* = \langle\mathbf{v}|T|\mathbf{w}\rangle$, after comparing the two expressions above, we get

   $$\mu\ \langle\mathbf{v}|\mathbf{w}\rangle = \lambda\ \langle\mathbf{v}|\mathbf{w}\rangle \implies \boxed{\langle\mathbf{v}|\mathbf{w}\rangle = 0} \text{ since } \mu \neq \lambda\ .$$

3. Diagonalizing a Normal Operator

We will need a couple of definitions. Let M be a linear operator on $\mathbb{V}$.

Definition: The kernel or null space of M is the subspace of $\mathbb{V}$ consisting of all vectors such that $M|\mathbf{v}\rangle = 0$. This subspace is referred to as $\text{Ker}(M)$.

Exercise: Determine the kernel of the operator given in Exercise 1.

Definition: Two subspaces $\mathbb{W}_1$ and $\mathbb{W}_2$ are **disjoint** if $\mathbb{W}_1 \cap \mathbb{W}_2 = (\mathbf{0})$.

Let T be a normal operator. We will show that there exists an ON basis where the matrix of T is diagonal. Let $(\lambda_1, \lambda_2, \ldots, \lambda_r)$ be the set of **distinct** eigenvalues with **multiplicities** $(m_1, m_2, \ldots, m_r)$. The characteristic polynomial is

$$P_T(\lambda) = \prod_{i=1}^{r} (\lambda_i - \lambda)^{m_i} \ .$$

For example, a hermitian operator P satisfying $P^2 = P$ will have two distinct eigenvalues i.e., $0, 1$.[1] The characteristic polynomial will be of the form $(-1)^n \lambda^{m_0} (\lambda - 1)^{m_1}$. The vector space $\mathbb{V}$ can be decomposed as follows:

$$\mathbb{V} = \text{Ker}(T_{\lambda_1}) \oplus \text{Ker}(T_{\lambda_2}) + \cdots + \text{Ker}(T_{\lambda_r}) \ .$$

where T_λ is defined in Eq. (6) and the $\oplus$ ('direct sum') indicates that each of the subspaces are **disjoint** from the other subspaces. This follows from property 4 of normal operators. For a fixed λ_i, the dimension of $\text{Ker}(T_{\lambda_i})$ is equal to the multiplicity m_i of the eigenvalue λ_i.

Suppose, the multiplicity is greater than 1, then we can use the Gram-Schmidt orthogonalization procedure to obtain an orthonormal basis of vectors for $\text{Ker}(T_{\lambda_i})$. Thus, given a normal operator T, we obtain a **natural** orthonomal basis for $\mathbb{V}$. Call this ON basis $|u_a\rangle$ $(a = 1, \ldots, n)$. Form the $n \times n$ matrix U whose columns are these n eigenvectors. The orthonormality of the basis implies that U is unitary (check!). It is also easy to see that

$$\mathbf{T} \cdot U = U \cdot \mathbf{D} \text{ where } \mathbf{D} = \text{Diag}(\lambda_1, \ldots, \lambda_k) \ .$$

Thus, we see that $\mathbf{D} = U^\dagger \cdot \mathbf{T} \cdot U$ – thus the diagonal form is the diagonal matrix of eigenvalues and U is the matrix of eigenvectors.

Example 4: Consider the matrix

$$\mathbf{T} = \begin{pmatrix} 1 & 1+i \\ 1-i & 2 \end{pmatrix}$$

[1] Such operators are called **projection** operators. They act as identity on the subspace with eigenvalue 1 and project out the null eigenspace.

Let us verify the properties mentioned above for this matrix.

1. The characteristic equation $\det(T - \lambda I) = 0$ is: $\lambda^2 - 3\lambda = 0$. Thus the eigenvalues are 0 and 3 both occur with multiplicity one.

2. The eigenvector for $\lambda = 0$ is $|v\rangle = (-1 - i, 1)^T$ and for $\lambda = 3$ is $|w\rangle = (1 + i, 2)^T$. One can verify that $\langle v|w\rangle = 0$.

3. We thus obtain one-dimensional subspaces corresponding to each eigenvalue. $\mathrm{Ker}(T_0) = \mathrm{Span}(|v\rangle)$ and $\mathrm{Ker}(T_3) = \mathrm{Span}(|w\rangle)$

4. Since the two eigenvectors are not normalised, we define normalised eigenvectors as follows. Let $|u_1\rangle = |v\rangle/\sqrt{3}$ and $|u_2\rangle = |w\rangle/\sqrt{6}$ leading to the unitary matrix

$$U = \begin{pmatrix} \frac{-1-i}{\sqrt{3}} & \frac{1+i}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{pmatrix}, \text{ and } \boxed{T \cdot U = U \cdot \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix}}.$$

Exercise: Repeat this exericise for the rotation matrix $\mathbf{R}$ defined in Example 2.

Exercise: Consider the matrix

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

Diagonalize the above matrix after verifying that it is hermitian.

4. Concluding Remarks

Linear operators play an important role in science and engineering. They provide a general framework that has a wide range of applications. Our interest in quantum mechanics will be to hermitian operators which will correspond to observables in quantum mechanics. Further applications of linear maps occurs in quantum computing and quantum information, a very modern application of quantum mechanics. Machine Learning (ML, in short) is another area that needs linear algebra (For instance, see Examples of Linear Algebra in ML).

A wonderful visual discussion on ideas in Linear Algebra appear in this youtube series by 3blue1brown titled The Essence of Linear Algebra. I recommend this to students who like to visualize ideas.

A lot of the discussion here will also appear in your current Mathematics course. See the lectures on this topic by Prof. Jayanthan.

There is a popular course given by Prof. Gilbert Strang on Linear Algebra that appears on MIT's OpenCourseWare.

What is the expression for an arbitrary rotation matrix in $\mathbb{R}^3$? Prof. Balakrishnan derives it in an interesting (magical?) fashion in an article: How is a vector rotated?. Express his result as a 3×3 matrix. Prove it using linear algebra.