Application of matrix in quantum mechanics

Koh Meng Hock

May 30, 2020

1 Introduction – measurement of quantal system

Let's begin by considering the act of measurement on a macroscopic (i.e. things that we can see e.g. car, table etc.) and microscopic (e.g. atoms, nucleus) objects. In order to make measurements of some physical quantities, one needs to send some kind of probes to the system. For example to measure the position of a macroscopic object such as a car, one needs to see where the object is. What is the probe here? It is the light that enables us see and thus complete the act of measurement.

Similarly, if one makes a measurement on a microscopic system one needs a probe. The difference is that the light that we use to see daily objects are visible light with wavelength of 400 - 700 nm. However, the wavelength needed to investigate microscopic objects must be shorter than those of the visible light. From the equation

$$E = \frac{hc}{\lambda} \tag{1}$$

we see that the energy for visible light is smaller than those needed to study microscopic objects.

This means that while making a measurement on microscopic object, we are bombarding the object with lots of energy. The consequence – we are disturbing the peace that the object once had! Say that before the measurement, the object is at a position x_1, then right after the measurement it will be shifted to a new position x_1.

Figure 1 Nice cartoon on the non-commutativity of position and momentum operators. Credit goes to the original cartoonist.

2 Eigenvalue equation

There are cases where the act of measurement does not change the state[1] of the quantal system. This means that the state of the system before and after the measurement is the same. In mathematical form, this is represented as

$$\hat{A}\psi = a\psi \tag{2}$$

where $\hat{A}$ is an operator[2] which represents a physical quantity to be measured, a is the value that one will get from the measurement process and Ψ is the wave function representing the state of the system. The equation (2) is known as the eigenvalue equation. It is important that the wave function remains the same on both sides of the equation.

3 Representating wave functions and operators in matrix form

The wave function Ψ is a vector and is represented by a column matrix. In the case of a 3D vector (ignoring other degrees of freedom), we have

$$\psi = \begin{pmatrix} \psi_x \\ \psi_y \\ \psi_z \end{pmatrix} \tag{3}$$

where ψ_x, ψ_y and ψ_z are essentially coefficients of the vector. For comparison we take a 3D vector for e.g. $\vec{r} = 2\hat{i} + 3\hat{j} + 4\hat{k}$ where in matrix form, the vector $\vec{r}$ is written as

$$\vec{r} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}. \tag{4}$$

Here we see that ψ_x is the value 2 in the vector $\vec{r}$ (the same for the others).

Operators, on the other hand, is represented by a square matrix

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \tag{5}$$

[1]Now we used a term that we see in quantum mechanics. Here the state of the system is represented by a wavefunction that contains all information about the system e.g. position, momentum, energy etc..

[2]The hat above the letter A signifies that this is an operator.

where Aij is the elements of the matrix A with i and j indices referring to the row and column of the matrix, respectively.

In equation (2) we say that a is the eigenvalue of $\hat{A}$. However, a might not necessary be a single value. It can have a spectrum of values. Assume that we are dealing with discrete values, we have for example $a = 1, 3, 5$ with certain probabilities of getting each of these values. Now we make lots of measurements and then to get an average value of these values, we calculate what is called the mean value of $\hat{A}$ by

$$\begin{pmatrix} \psi_x^* & \psi_y^* & \psi_z^* \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} \psi_x \\ \psi_y \\ \psi_z \end{pmatrix}. \tag{6}$$

4 Hermitian matrix – Hermitian operator

We have learnt about special matrices and one of them is **Hermitian** matrix which is defined as

$$A^\dagger = A^{*T} = A. \tag{7}$$

Hermitian matrix is of huge importance in quantum mechanics. Recall that any physical (observable) quantities to be measured are represented by operators. The square matrix representing these operators must be Hermitian. That is to say a non-Hermitian matrix cannot represent an observable quantity.

5 Commutator of two matrices – Simultaneous measurements of two observables

In matrix we have also learnt of **commutator** between two matrices say A and B i.e.

$$[A, B] = AB - BA. \tag{8}$$

If $[A, B] = 0$ then the two matrices are said to commute with one another.

This is important in quantum mechanics because it tells us whether we can measure two (or more) quantities at the same time. One of the three sets of Heisenberg's uncertainty principle given as

$$\Delta x \, \Delta p \geq \frac{\hbar}{2} \tag{9}$$

comes about from the equation[3]

$$\Delta A \, \Delta B \geq \frac{1}{2}|\langle [A, B] \rangle|. \tag{10}$$

The symbol ΔA represents the uncertainty in the measurement of a physical quantity represented by A. From the relation, we see that the uncertainty in measuring two physical quantities boils down to the commutation between the two matrices (or operators) representing those two quantities.

In the case there A is the position operator and B is the momentum operator, we have

$$[\hat{X}, \hat{P}_x] = i\hbar \tag{11}$$

i.e. the two matrices does not commute. Taking the absolute value of the commutator, one then obtains equation (9).

In Section 1 we speak about how measurements changes the state of the system. Here we can relate it to equation (9). If we send a light (with short wave length i.e. high in energy) to observe the position of a particle, then the high amount of energy supplied to the particle will alter the momentum of the particle. So now that we know the exact position of the particle[4], we loose track of its momentum.

Therefore, knowing whether two or more matrices commute or not allows us to determine the possibility for simultaneous measurements of physical quantities.

6 Eigenvalues of a matrix – Eigenvalue of an operator

In linear algebra class, we have learnt how to find eigenvalues of a square matrix. This is very relevant in quantum mechanics since an operator is always represented by a square matrix. Therefore, eigenvalues of the square matrix are in fact the eigenvalues of the operator. This means that when one performs repetitive measurements for the quantity represented by this operator, one will always obtain one of these eigenvalues.

The probability to get one eigenvalue over the other is of course not the same. But we will wait until a few sections later before coming back to this issue of finding the probability of finding an eigenvalue. For completeness, let us mention here that the eigenvalues of a matrix e.g. A above are obtained by equating the determinant

$$\begin{vmatrix} A_{11} - \lambda & A_{12} & A_{13} \\ A_{21} & A_{22} - \lambda & A_{23} \\ A_{31} & A_{32} & A_{33} - \lambda \end{vmatrix} = 0. \tag{12}$$

[3]For those interested to find out how we get this equation, do check out on **Uncertainty relation between two operators** in *Quantum Mechanics* by Nouredine Zettili or any other related textbook.

[4]This is the position before the particle flies off due to the extra energy supplied to it.

7 Diagonalization of matrix & searching for eigenvectors

Recall that for each eigenvalue we can find an eigenvector by inserting the eigenvalue into

$$\begin{pmatrix} A_{11}-\lambda & A_{12} & A_{13} \\ A_{21} & A_{22}-\lambda & A_{23} \\ A_{31} & A_{32} & A_{33}-\lambda \end{pmatrix} \begin{pmatrix} \psi_x \\ \psi_y \\ \psi_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{13}$$

Here $\begin{pmatrix} \psi_x \\ \psi_y \\ \psi_z \end{pmatrix}$ is the eigenvector in which coefficients ψ_x, ψ_y and ψ_z are to be determined[5]. Assuming we have a 3×3 matrix, we will then have 3 eigenvectors that can be combined into a new matrix denoted as C where

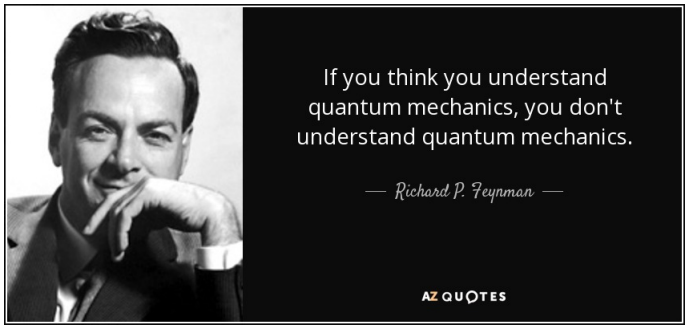
$$C = \begin{pmatrix} \psi_x^{(1)} & \psi_x^{(2)} & \psi_x^{(3)} \\ \psi_y^{(1)} & \psi_y^{(2)} & \psi_y^{(3)} \\ \psi_z^{(1)} & \psi_z^{(2)} & \psi_z^{(3)} \end{pmatrix}. \tag{14}$$

The matrix C is used to diagonalize the square matrix A representing the operator $\hat{A}$

$$C^{-1}\, A\, C = D \tag{15}$$

where C^{-1} is the inverse of matrix C and D is a diagonal matrix where the values along the diagonal are the 3 eigenvalues that we have found earlier.

Supposed that the eigenvalues of the matrix A is **degenerate**[6] then you will see that the eigenvectors of the degenerate eigenvalue are not orthogonal. We can employ the **Gram-Schmidt** method[7] in order to find orthogonal eigenvectors. At the end of this procedure, we will have 3 orthogonal eigenvectors that will diagonalize the square matrix A.



[5] Relate this to eq. (3).
[6] There are more than one eigenvector sharing the same eigenvalue.
[7] Find a short note for simple 2 degenerate case here.

8 State of the system, eigenvectors and probability of measurements

In this last section, I want to gel up all the things that we have discussed here. Especially on how to relate a general wave function for a system with eigenvectors and probability in measurements.

We first discussed about the state of the system and that it is described by a wave-function. But the state of the system does not have to exist in its eigenstate only, but rather in a **superposition of states**. Let us consider only three eigenstates so that the state of the system is

$$\Psi = a_1\psi_1 + a_2\psi_2 + a_3\psi_3 \tag{16}$$

where $\psi_i = \begin{pmatrix} \psi_x^{(i)} \\ \psi_y^{(i)} \\ \psi_z^{(i)} \end{pmatrix}$ (compare this to eq. (3) and (14)). In order for the wave function Ψ to be normalized, we must have the inner product of Ψ with itself to be equal to 1 so that

$$a_1^2 + a_2^2 + a_3^2 = 1. \tag{17}$$

The **probability** of obtaining an eigenvalue λ_i during measurement is then the square of the **probability amplitude** a_i i.e.

$$\mathcal{P}(\lambda_i) = |a_i|^2 \tag{18}$$

You will see from equation (17) that the sum of the probability is exactly 1.

One last note is that the probability amplitude a_i can be obtained by projecting Ψ onto ψ_i. Take for example $i = 1$, the probability amplitude a_1 can be obtained in matrix form by

$$a_1 = \underbrace{\begin{pmatrix} \psi_x^{(1)} & \psi_y^{(1)} & \psi_x^{(1)} \end{pmatrix}}_{\psi_1^\dagger} \cdot \underbrace{\left[a_1 \begin{pmatrix} \psi_x^{(1)} \\ \psi_y^{(1)} \\ \psi_z^{(1)} \end{pmatrix} + a_2 \begin{pmatrix} \psi_x^{(2)} \\ \psi_y^{(2)} \\ \psi_z^{(2)} \end{pmatrix} + a_3 \begin{pmatrix} \psi_x^{(3)} \\ \psi_y^{(3)} \\ \psi_z^{(3)} \end{pmatrix} \right]}_{\Psi} \tag{19}$$

Since the eigenvectors are orthogonal, the dot product between ψ_1 with ψ_2 and ψ_3 is exactly 0. The only surviving term is the dot product of ψ_1 with itself ψ_1.