Linear Algebra for Quantum Mechanics

Anisha Yeddanapudi

October 15, 2023

§1 Introduction

A couple notes:

1. I will be using mathematical notation for most of the linear algebra discussed here, however I will always have an aside describing the same system in Dirac notation.
2. My goal here is to provide a very general overview of the necessary QM and Linear Algebra that will be useful for understanding the experiment we're planning to conduct. So, I'll be skipping over some important content that would normally be further discussed in 137A.
3. I'm going to be assuming the simplest cases of all the following proofs so that means completeness, well-defined-ness, discreteness, pure quantum states, etc.
4. Feel free to ignore my asides...

§2 Hilbert Spaces

The state of a particle is represented as a vector which lives in a Hilbert space. A Hibert space is a infinite-dimension vector space. To understand this better, let us go over some important defintions:

Defintion: Field ($\mathbb{F}$)

For our purposes you can consider a **field** to be either the real or complex numbers ($\mathbb{R}$ or $\mathbb{C}$)

Elements in $\mathbb{F}$ are called scalars, which is a fancy word for 'number', as opposed to a vector.

Defintion: Vector Space

A **vector space** is a set V with operations addition and scalar multiplication such that the following properties hold:
(1) Commutativity: $u + v = v + u$ for all $u, v \in V$.
(2) Associativity: $(u + v) + w = u + (v + w)$ and $(ab)v = a(bv)$ for all $u, v, w \in V$ and $a, b \in \mathbb{F}$.
(3) Additive Identity: There exists an element $0 \in V$ such that $v + 0 = v$ for all $v \in V$.
(4) Additive Inverse: For every $v \in V$ there exists a $w \in V$ such that $v + w = 0$.
(5) Multiplicative Identity: $1v = v$ for all $v \in V$.
(6) Distributive Property: $a(u + v) = au + av$ and $(a + b)v = av + bv$ for all $a, b \in \mathbb{F}$ and $u, v \in V$.

Vectors in Dirac Notation

In quantum we commonly use Dirac Notation, where vectors such as v and u above are represented as kets $v \to |\psi\rangle$.

In general, the Hilbert spaces that we work with in Quantum Mechanics are $\mathbb{C}$-vector spaces. This means that the vector space itself 'originates' from the field $\mathbb{C}$. Some important properties that come form this defintion are:
(1) Vector spaces are closed under addition. So if $u, v \in V$ then $u + v \in V$.
(2) Vector spaces are closed under scalar multiplication. So if $u \in V$ then for all $\lambda \in \mathbb{F}$, $\lambda u \in V$.

Sets and Subsets

A **set** is a collection of objects. A **subset** is a smaller set whose objects all belong to a bigger separate set. For example $\mathbb{R}$ is a subset of $\mathbb{C}$.

Defintion: Subspace

A subset U of V is called a **subspace** of V if U is also a vector space (using the same addition and scalar multiplication as on V).

§2.1 Dimensionality

Now, we mentioned previously that Hilbert spaces are infinite dimensional, what does this mean?

Defintion: Linear Combination

A **linear combination** of a list of vectors $v_1, \cdots, v_m \in V$ is a vector of the form:

$$a_1 v_1 + \cdots + a_m v_m \quad \text{where } a_i \in \mathbb{F}$$

Linear Combinations in Dirac Notation

$$a_1|\psi_1\rangle + a_2|\psi_2\rangle + \cdots + a_m|\psi_m\rangle \quad \text{where } a_i \in \mathbb{F}$$

Now, in general, a list of vectors can span a space V, this means that every vector in V can be represented as a linear combination of that list.

Defintion: Span

A list of vectors $\{v_1, \cdots, v_m\}$ **spans** a vector space V if every $u \in V$ can be represented as:

$$u = a_1 v_1 + \cdots + a_m v_m \quad \text{where } a_i \in \mathbb{F}$$

Defintion: Linearly Independent

A list of vectors $v_1, \cdots, v_m \in V$ is called **linearly independent** if:

$$a_1 v_1 + \cdots + a_m v_m = 0 \text{ iff all } a_i = 0$$

Mathemathically, a specific vector space V is called finite-dimensional if there exists a list of vectors within V that spans V. And so a infinite-dimensional vector space, is one that is not

finite-dimensional. However, I would like to introduce this concept in a more scientific manner.

§2.1.1 Bases

This brings us to an important concept of a basis of a vector space.

Defintion: Basis

A **basis** of V is a list of vectors in V that is linearly independent and spans V.

Now, we can use the basis to define the dimensionality of V.

Defintion: Dimension

The **dimension** of V is the length of the list of basis vectors.

Therefore, an infinite-dimensional set has a infinite list of basis vectors. Now you may be wondering... since the choice of basis is non-unique, is it possible to have varying lengths of basis vector lists which would imply that a vector space could have multiple different dimensions. However, thre is a important theorem (which I can prove if asked) which states:

Basis length does not depend on basis

Any two bases of a vector space have the same length.

Hence, even though the list of a basis vectors is non-unique, the length of such lists is unique.

Basis in Dirac Notation

In a n-dimensional Hilbert space, we can write any ket $|\psi\rangle$ in terms of $|n\rangle$ other kets (the basis vectors) multiplied by specific scalars.

$$|\psi\rangle = a_1|1\rangle + a_2|2\rangle + a_3|3\rangle + \cdots + a_n|n\rangle$$

ASIDE: We can also add different kets. Assuming $|\psi_1\rangle = a_1|1\rangle + \cdots + a_n|n\rangle$ and $|\psi_2\rangle = b_1|1\rangle + \cdots + b_n|n\rangle$ then:

$$|\psi_1\rangle + |\psi_2\rangle = (a_1 + b_1)|1\rangle + \cdots + (a_n + b_n)|n\rangle$$

The concept of a basis will be referenced frequently as we move forward, so please be sure to keep this in mind!

§2.2 Inner Product Space

Finally, Hilbert spaces are also inner product spaces.

Defintion: Inner Product

An **inner product** on V is a function that takes an ordered pair (u, v) of $u, v \in V$ to a number $\langle u, v \rangle \in \mathbb{F}$ and has the following properties:
(1) Positivity: $\langle v, v \rangle \geq 0$ for all $v \in V$.
(2) Definiteness: $\langle v, v \rangle = 0$ iff $v = 0$.
(3) Additivity in First slot: $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ for all $u, v, w \in V$.
(4) Homogeneity in First slot: $\langle \lambda u, v \rangle = \lambda \langle u, v \rangle$ for all $\lambda \in \mathbb{F}$ and all $u, v \in V$.
(5) Conjugate Symmetry: $\langle u, v \rangle = \overline{\langle v, u \rangle}$ for all $u, v \in V$.

Complex Conjugate

As I stated previously, the Hilbart spaces we work in for quantum mechanics are $\mathbb{C}$-vector spaces. Within the complex numbers you can define something known as a complex conjugate. All complex numbers can be represented as:

$$z = a + bi \quad \text{where } a, b \in \mathbb{R}$$

So the complex conjugate $\bar{z}$ is:

$$\bar{z} = a - bi \quad \text{where } a, b \in \mathbb{R}$$

Basically where you flip the sign on the complex portion.

Defintion: Inner Product Space

An **inner product space** is a vector space V along with the inner product on V.

There are a couple of relevant examples of inner products which become important in qunatum mechanics. However, an inner product can be defined in any of these ways within a specific vector space, but you should never think that these are the only ways to define inner products. Anything that satisfies the defintion above can be characterized as a inner product below are just a few important examples.

- The Euclidean Inner Product (Dot Product).

$$\langle (v_1, \cdots, v_n), (u_1, \cdots, u_n) \rangle = v_1 \bar{u}_1 + \cdots + v_n \bar{u}_n$$

- Continuous Real Valued Funtions.

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx$$

- On $\mathcal{P}(\mathbb{R})$ (the set of all polynomails with real coefficients).

$$\langle p, q \rangle = \int_0^\infty p(x) q(x) e^{-x} dx$$

We commonly come across forms of these throughout quantum mechanics. However one important note before moving forward.

L^2 Spaces

L^2 defines the set of all square integrable functions on a specific interval $[a, b]$.

$$L^2(a,b) := \left\{ f(x) : \int_a^b |f(x)|^2 dx < \infty \right\}$$

For physicists Hilbert spaces are L^2 spaces, by defintion. But mathemathicians can refer to them as separate things.

Inner Products in Dirac Notation

As stated previously, we represent states in QM as 'kets' $|\psi\rangle$ where they are simply vectors living in the Hilbert space. We can also define something known as a 'bra' $\langle\phi|$ such that:

$$\langle\phi|\psi\rangle \text{ for states } |\psi\rangle \text{ and } |\phi\rangle \text{ is known as a 'bra-ket'}$$

The above statement defines the inner product between $|\psi\rangle$ and $|\phi\rangle$. We consider two main ways to define the inner product in QM:
(1) $\langle\psi(x)|\phi(x)\rangle = \int_a^b \overline{\psi}(x)\phi(x)dx$ where $|\psi\rangle$ and $|\phi\rangle$ are two (time-independent) wave functions.
(2) $\langle\psi|\phi\rangle = \sum_k^n \overline{\psi}_k\phi_k$ where $|\psi\rangle$ and $|\phi\rangle$ have some matrix representations (refer to section 3.1.1 below for clarification) and $\langle\psi|$ represents the conjugate transpose of $|\psi\rangle$.
My goal here is to prove that these two representations are equivalent.

Proof. If we have two functions $f(x)$ and $g(x)$ we can sample these functions at a series of discrete locations (with regular intervals $\Delta x = \frac{b-a}{n-1}$). Such that $f_i = f(x_i)$ and $g_i = g(x_i)$. Refer to the figure below for clarification. Now we rewrite $\vec{f}$ and $\vec{g}$ as a matrix containing these n points.

$$\vec{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} \qquad \text{and} \qquad \vec{g} = \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix}$$

Let us proceed to define the inner product of these vectors as the Euclidean Inner Product (Dot Product). Hence:

$$\langle\vec{f},\vec{g}\rangle = \vec{g}^*\vec{f} = \sum_{k=1}^{n} f_k\bar{g}_k$$

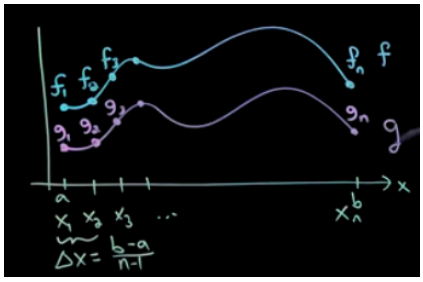
Now we proceed to normalize this to ensure that our inner product doesn't get infinitely big as we add more data points.

$$\langle\vec{f},\vec{g}\rangle\Delta x = \sum_{k=1}^{n} f(x_k)\bar{g}(x_k)\Delta x$$

Notice that this expression above is simply the Reimann Approximation of:

$$\langle f(x), g(x)\rangle = \int_a^b f(x)\bar{g}(x)dx$$

Therefore, if we take the limit as $\Delta x \to 0$ then summation can be converted to an integral and we obtain the final result. □



There are reasons why the inner product is defined in this way, and not differently. I won't get into the nuance of that here because it's outside of what's needed for our work together. But let me know if you're interested and we can always chat about it.

§2.2.1 Orthogonal and Orthonormal

Defintion: Orthogonal

Two vectors $u, v \in V$ are called **orthogonal** if $\langle u, v \rangle = 0$.

Defintion: Norm

For $v \in V$ the **norm** of v, denoted as $||v||$, is defined by:

$$||v|| = \sqrt{\langle v, v \rangle}$$

Defintion: Orthonormal

A list of vectors is called **orthonormal** if each vector in the list has norm 1 and is orthogonal to all other vectors in the list.

Defintion: Orthonormal Basis

An **orthonormal basis** of V is an orthonormal list of vectors in V that is also a basis of V.

There are a variety of reasons why orthonormal basis are particularly useful in QM. We shall be covering a few in Sec. 4. Additionally, a orthonormal basis for a vector space always exists due to the Gramm-Schmidt Procedure. We don't need to worry about the procedure itself just the fact that it implies that any basis can be converted to an orthonormal one.

§3 Observables

In general, observables represent things that you can measure within a qunatum system (momentum, position, spin, angular momentum, etc.). Observables are a more specific concept based on some general principles which we will first discuss.

§3.1 Linear Maps

Defintion: Linear Maps

A **linear map** from V to W is a function $T : V \to W$ which the following properties:
(1) Additivity: $T(u + v) = T(u) + T(v)$ for all $u, v \in V$.
(2) Homodeneity: $T(\lambda v) = \lambda(Tv)$ for all $\lambda \in \mathbb{F}$ and all $v \in V$.

§3.1.1 Matrices

Now, you may have heard that vectors (or states) can be represented as matrices where:

$$|\psi\rangle = \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_n \end{bmatrix}$$

The reason for this, is because of the general defintion of matrices and their inherent connection with vectors and vector spaces. We won't discuss that too deeply here, however, if you're curious read Ch 3C in Axler.

In addition to states, linear maps can also be represented in matrix form, which inherently depends on the chosen basis.

Defintion: Matrix of Linear Map

Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \cdots, v_n$ is a basis of V and $w_1, \cdots, w_m$ is a basis of W. The **matrix of** T with respect to these basis is a $m \times n$ matrix $\mathcal{M}(T)$ whose entries A_{jk} are defined by:

$$Tv_k = A_{1,k}w_1 + \cdots + A_{m,k}w_m$$

It's for this definitional reason that operators have a matrix form. If this seems confusing, don't worry! It's honestly not too important for what we're about to discuss, but a good concept to have rattling around in the back of your head.

§3.2 Hermitian Operators

Defintion: Operator

A linear map from a vector space to itself is called an **operator**.

Defintion: Adjoint

Suppose $T \in \mathcal{L}(V, W)$. The **adjoint** of T is a function $T^\dagger : W \to V$ such that:

$$\langle Tv, w\rangle = \langle v, T^\dagger w\rangle$$

for every $v \in V$ and $w \in W$.

Defintion: Hermitian

An operator is called **hermitian** if $T = T^\dagger$.

$\sim$ ASIDE: A BIT OF A TANGENT $\sim$

The bra is defined to be the Hermitian conjugate of it's corresponding ket. The matrix representation of $T^\dagger$ is the conjugate transpose of $\mathcal{M}(T)$, as a result of the following theorem:

The matrix of $T^\dagger$

Let $T \in \mathcal{L}(V, W)$. Suppose $e_1, \cdots, e_n$ is an orthonormal basis of V and $f_1, \cdots, f_m$ is an orthonormal basis of W then:

$$\mathcal{M}(T^\dagger, (f_1, \cdots, f_m), (e_1, \cdots, e_n))$$

is the conjugate transpose of $\mathcal{M}(T, (e_1, \cdots, e_n), (f_1, \cdots, f_m))$

I'm not going to prove this theorem here, but check out 7A in Axler (pg 208) if you're curious.
$\sim$ END OF ASIDE $\sim$

Observables are represented by hermitian operators. There is a more nuanced reason for why this is, but the main one I'm going to focus on is that hermtian operators have real eigenvalues (refer to Sec. 4).

§4 Statistical Interpretation

Defintion: Eigenvalues/Eigenvectors

Suppose $T \in \mathcal{L}(V)$. A scalar $\lambda \in \mathbb{F}$ is called an **eigenvalue** of T if there exists a $v \in V$ such that $v \neq 0$ and $Tv = \lambda v$. And v is the corresponding **eigenvector**.

> If you measure an observable $\hat{Q}$ on a particle in some state $|\Psi\rangle$ you will get one of the eigenvalues of $\hat{Q}$.

This is an assumption that's made in QM, and there isn't a concrete proof for why this happens. Many different interpretations of QM try to build this up... but we're not discussing those here.

Interestingly, the eigenvalues of observables are real because they are hermitian.

Eigenvalues of Hermitian Operators are Real

Let $\mathcal{H}$ be a hilbert space and $T \in \mathcal{L}(\mathcal{H})$ be an Hermitian operator. Then all the eigenvalues of T are real.

Proof. Let λ be an eigenvalue of T corresponding with eigenvector $v \in \mathcal{H}$. Then $Tv = \lambda v$. Consider the following:

$$\begin{aligned}
\lambda\langle v, v\rangle &= \langle \lambda v, v\rangle \\
&= \langle Tv, v\rangle \\
&= \langle v, Tv\rangle \\
&= \overline{\langle Tv, v\rangle} \\
&= \overline{\langle \lambda v, v\rangle} \\
&= \bar{\lambda}\langle v, v\rangle
\end{aligned}$$

The only way $\lambda = \bar{\lambda}$ is if $\lambda \in \mathbb{R}$. □

This is one of the reasons why we represent observables as hermitian operators, because when we measure the system we wish to get real values.

$\sim$ ASIDE: A BIT OF PERSPECTIVE $\sim$
The reason why linear algebra is so good at representing QM is non-trivial. There is no reason for why the quantum world should obey the rules of linear algebra, the fact that it does it honestly really cool! So when I say thing like:

> Observables are represented as hermitian operators because they have real eigenvalues.

You may be tempted to ask whether observables implies hermitian operator OR if it goes the other way around hermitian operator implies observables. This becomes a bit like which came first the chicken or the egg. And to be honest, I don't have a good answer because it bcomes a more philosphical question about the inherent nature of QM.
$\sim$ END OF ASIDE. $\sim$

Repeating what was said previously mathematically...
The eigenvectors of a observable form a complete orthonormal basis (omitting proof). Let us define some operator $\hat{Q}$ where the eigenvectors(functions) of this operator are the set $\{|q_1\rangle, \cdots, |q_n\rangle\}$. This is a basis for the wave function, so $|\Psi\rangle$ can be written as a linear combination of q_i's:

$$|\Psi\rangle = c_1|q_1\rangle + \cdots + c_n|q_n\rangle \quad \text{where } c_i \in \mathbb{C}$$

When conducting a measurement of $\hat{Q}$ we would obtain one of the q_i eigenvalues (denoted as λ_i). Now, an obvious question one can ask is what is the probability of getting a specific λ_i when you conduct a measurement?

§4.1 Probabilities

The probabilities referenced previously are given by the Born Rule.

Born Rule

If a system is in a state $|\Psi\rangle$ (assuming pure state) then the probability $\mathbb{P}$ that an eigenvalue λ_i of q_i is found when $\hat{Q}$ is measured is:

$$\mathbb{P}(\lambda_i) = |(q_i, \Psi)|^2$$

Let's proceed to break this down. Consider again our previous representation of $|\Psi\rangle$ in the $\hat{Q}$ basis:

$$|\Psi\rangle = c_1|q_1\rangle + \cdots + c_n|q_n\rangle$$

Assume that this state is properly normalized such that $\sum_i |c_i|^2 = 1$. Then the Born rule is basically saying that:

$$\begin{aligned}
\mathbb{P}(\lambda_i) &= |(q_i, \Psi)|^2 \\
&= \left| \left\langle q_i \middle| c_1|q_1\rangle + \cdots + c_n|q_n\rangle \right\rangle \right|^2 \\
&= |c_1\langle q_i|q_1\rangle + \cdots + c_i\langle q_i|q_i\rangle + \cdots + c_n\langle q_i|q_n\rangle|^2 \\
&= |c_i\langle q_i|q_i\rangle|^2 \quad \text{all } \langle q_i|q_j\rangle = 0 \text{ if } i \neq j \text{ by orthogonality} \\
&= |c_i|^2 \quad \text{by orthonormality } \langle q_1|q_i\rangle = 1
\end{aligned}$$

So the probability is the magnitude squared of the constant multiple associated with linear expansion of $|\Psi\rangle$ in the $\hat{Q}$ eigenbasis. This is one of the reason why it's inherently useful to have a orthonormal basis as calculating the inner products becomes much easier.

$\sim$ ASIDE: DEGENERACY $\sim$
We can also have the same eigenvalue associated with multiple different eigenvectors (qunatum states). If this is the case then that quantum state is called degenerate with dengeneracy equivalent to the multiplicty of the eigenvalue.
In this case calculating the probability associated with a specific eigenstate can get slightly more complicated.
$\sim$ END OF ASIDE $\sim$

§5 Time Evolution

Given a quantum system in some state $|\Psi\rangle$, we can also ask how the system will look like in the future. The time-evolution of $|\Psi\rangle$ is given by the Schrodinger Eqn:

$$i\hbar\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}|\Psi\rangle$$

Here $\hat{H}$ representes the hermitian operator called the Hamiltonian which, broadly, representes the total energy of a quantum system.

§5.1 Measurement Collapse

Let say we conduct a measurement of some observable $\hat{Q}$ and obtain the eigenvalue λ_j associated with eigenvector $|q_j\rangle$. Immediately after the measurement, regardless of the state of the system beforehand, the system will now solely be in the state $|q_j\rangle$.

$$|\Psi_{\text{before}}\rangle = c_1|q_1\rangle + \cdots + c_n|q_n\rangle \Longrightarrow |\Psi_{\text{after}}\rangle = |q_j\rangle$$

So if you were to measure $\hat{Q}$ again, you would get λ_j with 100% probability. This is known as 'the collapse of the wavefunction'. After this collapse the system continues to evolve via the Schrodinger Eqn.