Quantum Algorithms and Learning Theory
Notes and Exercises

Faris Sbahi

March 16, 2018

1 Nielsen & Chuang: Chapter 2

1.1 Postulates of Quantum Mechanics

First, we cover the fundamental postulates of quantum mechanics.

1.1.1 State Space and State Vector

Associated with an isolated physical system is a Hilbert space, $\mathcal{H}$. A Hilbert space is a complete inner-product vector space. Note that completeness holds trivially in a finite-dimensional vector space because we have closure with respect to all sequences (and hence any Cauchy sequence in the vector space must converge to a vector in the same space). Nevertheless, the state space of a physical system may be infinite-dimensional.

A system is completely described by a unit vector $u \in \mathcal{H}$ called the state vector.

For example, consider a system given by a single qubit, which has a two-dimensional state space. Let $|0\rangle$ and $|1\rangle$ be an orthonormal basis for this space. Hence, a state vector in this space is given by

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle$$

where $a, b \in \mathbb{C}$ and $|a|^2 + |b|^2 = 1$.

1.1.2 Evolution

The evolution of a closed quantum system is described by a unitary transformation. Recall that an operator U is unitary iff $U^\dagger U = I = UU^\dagger$ (and hence preserves inner products[1]).

So, let the state of a system at time t_1 be given by $|\psi\rangle$ and $|\psi'\rangle$ at t_2. Hence,

$$|\psi'\rangle = U\,|\psi\rangle$$

[1] and furthermore has a spectral decomposition because it is normal

1.1.3 Evolution in Continuous Time

Schrodinger's equation provides the time evolution of the state of a quantum system

$$i\hbar \frac{d\left|\psi\right\rangle}{dt} = H\left|\psi\right\rangle \tag{1}$$

where H is the (Hermitian) Hamiltonian of the closed system. Because the Hamiltonian is Hermitian it has spectral decomposition

$$H = \sum_E E\left|E\right\rangle\left\langle E\right|$$

where E is the energy eigenvalue corresponding to energy eigenstate $\left|E\right\rangle$.

For example, consider the Hamiltonian $H = \hbar\omega X$ (recall that $X = \sigma_X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$). Hence, we solve for its eigenvalues and eigenvectors

$$\begin{aligned}
\det\left\{\hbar\omega\begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix}\right\} &= 0 \\
\lambda^2 - 1^2 &= 0 \\
\lambda &= \pm 1 \\
\Rightarrow E_\pm &= \pm\hbar\omega \\
\hbar\omega\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\left|E_+\right\rangle &= 0 \\
\left|E_+\right\rangle &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} := \left|+\right\rangle \\
\left|E_-\right\rangle &= \left|-\right\rangle
\end{aligned}$$

Now, notice that we can solve Schrodinger's equation (1) and have

$$\left|\psi(t_2)\right\rangle = \exp\left[\frac{-iH(t_2 - t_1)}{\hbar}\right]\left|\psi(t_1)\right\rangle$$

and equivalently from 1.1.2 we can represent this transformation with unitary operator $U = \exp\left[\frac{-iH(t_2-t_1)}{\hbar}\right]$. This holds in general and so we can consider the two descriptions from 1.1.2 and 1.1.3 interchangeably (the authors prefer the latter).

1.1.4 Quantum Measurement

Quantum measurements are described by a collection of measurements operators $\{M_m\}$ (where m refers to the potential measurement outcomes of the experiment) which act on the state space of the system being observed.

Hence, if the pre-measurement state is $|\psi\rangle$, then

$$p(m) = \langle\psi| M_m^\dagger M_m |\psi\rangle$$

and the post-measurement state is

$$\frac{M_m |\psi\rangle}{\sqrt{p(m)}}$$

Furthermore, $\{M_m\}$ satisfy the completeness equation

$$\sum_m M_m^\dagger M_m = I$$

Now, we see an interesting implication. If we seek to distinguish our physical system from a set of orthogonal states, then we can reliably do so by simply defining each measurement operator to be the outer product of our states of interest. We add a final operator defined to be the remaining complement of the identity in order to satisfy the completeness equation.

On the flipside, two non-orthogonal states $|\psi_1\rangle$ and $|\psi_2\rangle$ necessarily share a parallel component in their orthogonal decomposition. Hence, a measurement outcome that corresponds to the pre-measurement state being $|\psi_1\rangle$ with probability $p > 0$ has a probability $p' > 0$ of having been in state $|\psi_2\rangle$.

1.1.5 Projective Measurements

There exists a special class of quantum measurements known as projective measurements. These measurements can be described by an observable M, a hermitian operator on the state space being observed. M has spectral decomposition

$$M = \sum_m m P_m$$

where P_m is the projector onto the eigenspace of M with eigenvalues m.

Furthermore, if the pre-measurement state is $|\psi\rangle$, then

$$p(m) = \langle\psi| P_m |\psi\rangle$$

and the post-measurement state is

$$\frac{P_m \left|\psi\right\rangle}{\sqrt{p(m)}}$$

This simplifies the formula for the expected value of a measurement

$$\begin{aligned}\langle M\rangle &= \sum_m mp(m)\\ &= \left\langle\psi\right| (\sum_m mP_m) \left|\psi\right\rangle\\ &= \left\langle\psi\right| M \left|\psi\right\rangle\end{aligned}$$

For example, consider the system given by single qubits with observable Pauli matrix Z. Hence, Z has eigenvalues $+1$ and -1 and eigenstates $|0\rangle$ and $|1\rangle$, respectively. So, consider state $|\psi\rangle = |+\rangle \Rightarrow p(+1) = \langle +|0\rangle \langle 0|+\rangle = \frac{1}{2}$. Similarly, $p(-1) = \frac{1}{2}$.

1.1.6 POVM measurements

POVMs are best viewed as a special case of the general measurement formalism, providing the simplest means to study post-measurement statistics without knowledge of the post measurement state.

From above, $p(m) = \left\langle\psi\right| M_m^\dagger M_m \left|\psi\right\rangle$ so if we define $E_m := M_m^\dagger M_m$ then these E_m's are sufficient for the purpose of computing probabilities. We denote $\{E_m\}$ as a POVM.

Note that projective operators are the special case of being equivalent to their respective POVM element because $E_m = P_m^\dagger P_m = P_m$.

Nevertheless, the POVM formalism is a useful guide in for our intuition in quantum information. Consider if Alice prepares some state for Bob that is either $|\psi_1\rangle = |0\rangle$ or $|\psi_2\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$. Recall from 1.1.4, Bob can't determine which state was prepared with full certainty (because of the shared orthogonal component $|0\rangle$). Still, we can define a POVM[2]

$$\begin{aligned}E_1 &= \frac{\sqrt{2}}{1+\sqrt{2}} \left|1\right\rangle \left\langle 1\right|\\ E_2 &= \frac{\sqrt{2}}{1+\sqrt{2}} \frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2}\\ E_3 &= I - E_1 - E_2\end{aligned}$$

Now, notice what happens.

[2] verify that completeness and these being positive operators holds

$$
\begin{aligned}
\langle\psi_1| E_1 |\psi_1\rangle &= \langle 0| \frac{\sqrt{2}}{1+\sqrt{2}} |1\rangle \langle 1|0\rangle \\
&= 0 \\
\langle\psi_2| E_1 |\psi_2\rangle &= \frac{\langle 0| + \langle 1|}{\sqrt{2}} \frac{\sqrt{2}}{1+\sqrt{2}} |1\rangle \langle 1| \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\
&= \frac{\sqrt{2}}{2\sqrt{2}+2} > 0
\end{aligned}
$$

Hence, if we observe E_1 after the measurement described by $\{E_1, E_2, E_3\}$, then Alice must've prepared $|\psi_2\rangle$. Similarly,

$$
\begin{aligned}
\langle\psi_1| E_2 |\psi_1\rangle &= \langle 0| \frac{\sqrt{2}}{1+\sqrt{2}} \frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2} |0\rangle \\
&= \frac{\sqrt{2}}{2\sqrt{2}+2} > 0 \\
\langle\psi_2| E_2 |\psi_2\rangle &= \frac{\langle 0| + \langle 1|}{\sqrt{2}} \frac{\sqrt{2}}{1+\sqrt{2}} \frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2} \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\
&= 0
\end{aligned}
$$

so if we observe E_2, then Bob concludes that Alice prepared $|\psi_1\rangle$. Our routine is imperfect because we may observe E_3 and hence would infer nothing of the original state. Still, we would never *incorrectly* guess given that we allow ourselves to abstain when we see E_3.

Exercise 1.1. *(2.64) Suppose Bob is given a quantum state chosen from a set $S = |\psi_1\rangle, \cdots, |\psi_m\rangle$ of linearly independent states. Construct a POVM $\{E_1, \cdots, E_{m+1}\}$ such that if outcome E_i occurs, $1 \leq i \leq m$, then Bob knows with certainty that he was given state $|\psi_i\rangle$.*

To distinguish the states we require $\langle\psi_i| E_j |\psi_i\rangle = p_i\delta_{ij}$ where $p_i > 0$ and $1 \leq i, j \leq m$.

So, we can use the Gram-Schmidt process using S as our linearly independent set. This will give us an orthonormal set $U = |\varphi_1\rangle, \cdots, |\varphi_m\rangle$ that spans the same subspace as S. Next, we can represent each $|\psi_i\rangle$ in this orthonormal basis, U. Finally, for each i we can find a vector $|\psi_i'\rangle$ in the span of U that is orthogonal to all $|\psi_j\rangle, j \neq i$. Hence, we can define $E_i = |\psi_i'\rangle\langle\psi_i'|, 1 \leq i \leq m$. Finally, take $E_{m+1} = I - \sum_m E_i$.

Creating an optimal POVM is much trickier (in the sense of minimizing the probability p_{m+1}).

From this exercise, we see that POVMs present a reliable way to distinguish non-orthogonal (but linearly independent) states given that we allow for the slack of an "inconclusive" measurement (E_{m+1}).

1.1.7 Composite Systems

The state space of a composite physical system is the tensor product of the state spaces of the component physical systems.

Exercise 1.2. *(2.66) Show that the average value of the observable $X_1 Z_2$ (X acting on the first qubit and Z on the second) for a two qubit system measured in the state $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$ is zero.*

Proof. Let observable $M = X_1 Z_2$. Hence,

$$\begin{aligned}
\langle M \rangle &= \frac{\langle 00| + \langle 11|}{\sqrt{2}} M \frac{|00\rangle + |11\rangle}{\sqrt{2}} \\
&= \frac{\langle 00| + \langle 11|}{\sqrt{2}} \frac{X_1 |0\rangle Z_2 |0\rangle + X_1 |1\rangle Z_2 |1\rangle}{\sqrt{2}} \\
&= \frac{\langle 00| + \langle 11|}{\sqrt{2}} \frac{|1\rangle |0\rangle - |0\rangle |1\rangle}{\sqrt{2}} \\
&= 0
\end{aligned}$$

□

Interestingly, we can show that a general quantum measurement (as described in 1.1.4) can be implemented as a projective measurement coupled with unitary dynamics.

Consider a quantum system with state space Q and measurements M_m on this system. We can introduce an *ancilla* system M with orthonormal basis $|m\rangle$ which is in one-to-one correspondence with the possible outcomes of the measurement we wish to implement.

So, let $|0\rangle$ be a fixed state of M and define an operator U on $|\psi\rangle |0\rangle$ (with $|\psi\rangle$ as a state of Q) by

$$U |\psi\rangle |0\rangle := \sum_m M_m |\psi\rangle |m\rangle$$

Hence,

$$\langle \varphi| \langle 0| U^\dagger U |\psi\rangle |0\rangle = \sum_m \sum_{m'} \langle \varphi| M_m^\dagger M_{m'} |\psi\rangle \langle m|m'\rangle$$

So, because the states $|m\rangle$ are orthonormal

$$= \sum_m \langle \varphi| M_m^\dagger M_m |\psi\rangle$$

and finally by the completeness of M_m

$$= \langle \varphi | \psi \rangle$$

This tells us that U preserves inner products between states of the form $|\psi\rangle |0\rangle$. Furthermore, we can show that U can be extended to a unitary operator on $Q \otimes M$ (exercise).

Hence, consider a projective measurement on the two systems $(U |\psi\rangle |0\rangle)$ given by projectors $P_m := I_Q \otimes |m\rangle \langle m|$. So,

$$\begin{aligned} p(m) &= \langle \psi| \langle 0| U^\dagger P_m U |\psi\rangle |0\rangle \\ &= \sum_{m'} \sum_{m''} \langle \psi| M_{m'}^\dagger \langle m'| (I_Q \otimes |m\rangle \langle m|) M_{m''} |\psi\rangle |m''\rangle \\ &= \langle \psi| M_m^\dagger M_m |\psi\rangle \end{aligned}$$

which agrees with the general result from 1.1.4. Similarly, the post-measurement state is as expected. Hence, we've shown that unitary dynamics, projective measurements, and ancillary systems can be used together to describe any general measurement.

Exercise 1.3. *(2.68) Prove that $|\psi\rangle \neq |a\rangle |b\rangle$ for all single qubit state $|a\rangle$ and $|b\rangle$ where $|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$.*

Proof. First, decompose the qubit state in their basis, $|a\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle$ and $|b\rangle = \beta_0 |0\rangle + \beta_1 |1\rangle$. Now, we prove by contradiction

$$\begin{aligned} |a\rangle |b\rangle &= \alpha_0\beta_0 |00\rangle + \alpha_0\beta_1 |01\rangle + \alpha_1\beta_0 |10\rangle + \alpha_1\beta_1 |11\rangle \\ &= \frac{|00\rangle + |11\rangle}{\sqrt{2}} \end{aligned}$$

which would imply that either α_0 or β_1 are zero in order to remove the $|01\rangle$ term. However, this would also remove either the $|00\rangle$ or $|11\rangle$ term, so we have a contradiction. □

A state of a composite system having this property is said to be entangled.

1.2 Superdense Coding

Suppose Alice is in possession of two classical bits of information she wishes to transmit to Bob, but is only allowed to send a single qubit to Bob.

Now, suppose that Alice and Bob initially share a pair of qubits in the entangled state from above

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

where Alice is initially holding the first qubit and Bob the second. She can then apply a particular gate to send a bit string. Below shows the corresponding gate and resulting state

Bit String	Applied gate	Resulting state
00	–	$\frac{\lvert 00\rangle+\lvert 11\rangle}{\sqrt{2}}$
01	Z	$\frac{\lvert 00\rangle-\lvert 11\rangle}{\sqrt{2}}$
10	X	$\frac{\lvert 10\rangle+\lvert 01\rangle}{\sqrt{2}}$
11	iY	$\frac{-\lvert 10\rangle+\lvert 01\rangle}{\sqrt{2}}$

Observe that these are the Bell states. Furthermore, Bell states form an orthonormal basis and hence can be distinguished (as we've discussed in 1.1.6). Hence, Alice needs only to interact with the single qubit to transmit two classical bits of information to Bob.

Exercise 1.4. *(2.70) Suppose E is any positive operator acting on Alices qubit. Show that $\langle\psi| E \otimes I |\psi\rangle$ takes the same value when $|\psi\rangle$ is any of the four Bell states. Suppose some malevolent third party ('Eve') intercepts Alices qubit on the way to Bob in the superdense coding protocol. Can Eve infer anything about which of the four possible bit strings 00, 01, 10, 11 Alice is trying to send? If so, how, or if not, why not?*

Proof.

$$\begin{aligned}
\langle 00| + \langle 11| \,(E \otimes I)\, |00\rangle + |11\rangle &= \langle 0| E |0\rangle + \langle 1| E |1\rangle \\
\langle 00| - \langle 11| \,(E \otimes I)\, |00\rangle - |11\rangle &= \langle 0| E |0\rangle + \langle 1| E |1\rangle \\
\langle 10| + \langle 01| \,(E \otimes I)\, |10\rangle + |01\rangle &= \langle 0| E |0\rangle + \langle 1| E |1\rangle \\
-\langle 10| + \langle 01| \,(E \otimes I) - |10\rangle + |01\rangle &= \langle 0| E |0\rangle + \langle 1| E |1\rangle
\end{aligned}$$

Hence, Eve can't infer anything. The states are only distinguishable if one can perform a measurement that acts on both qubits. □

1.3 The Density Operator

An alternative formulation of quantum mechanics is possible using a tool known as the density operator.

Suppose a quantum system is one of a number of states $|\psi\rangle$ with probability p_i. We call $\{p_i, |\psi_i\rangle\}$ an ensemble of pure states. The density operator is defined

$$\rho := \sum_i p_i |\psi_i\rangle \langle\psi_i|$$

Evolution of the density operator (under a unitary transformation) can be derived readily

$$\sum_i p_i U |\psi_i\rangle \langle\psi_i| U^\dagger = U\rho U^\dagger$$

If we perform a measurement with operator M_m with initial state $|\psi_i\rangle$ then

$$\begin{aligned} p(m \mid i) &= \langle\psi_i| M_m^\dagger M_m |\psi_i\rangle \\ &= tr(M_m^\dagger M_m |\psi_i\rangle \langle\psi_i|) \end{aligned}$$

using 9.2.
Hence, summing this conditional probability across all initial states we have

$$\begin{aligned} p(m) &= \sum_i p_i tr(M_m^\dagger M_m |\psi_i\rangle \langle\psi_i|) \\ &= tr(M_m^\dagger M_m \rho) \end{aligned}$$

The state after obtaining measurement result m on initial state $|\psi_i\rangle$ is

$$|\psi_i^m\rangle = \frac{M_m |\psi_i\rangle}{\sqrt{\langle\psi_i| M_m^\dagger M_m |\psi_i\rangle}}$$

and so the density operator after result m is given by

$$\begin{aligned} \rho_m &= \sum_i p(i \mid m) |\psi_i^m\rangle \langle\psi_i^m| \\ &= \sum_i p(i \mid m) \frac{M_m |\psi_i\rangle \langle\psi_i| M_m^\dagger}{\langle\psi_i| M_m^\dagger M_m |\psi_i\rangle} \end{aligned}$$

Furthermore, from Baye's rule we have that $p(i \mid m) = \frac{p(m|i)p(i)}{p(m)}$ so we can simplify

$$\begin{aligned} \rho_m &= \sum_i \frac{p(m \mid i)p(i)}{p(m)} \frac{M_m |\psi_i\rangle \langle\psi_i| M_m^\dagger}{\langle\psi_i| M_m^\dagger M_m |\psi_i\rangle} \\ &= \sum_i \frac{p(i) \langle\psi_i| M_m^\dagger M_m |\psi_i\rangle}{tr(M_m^\dagger M_m \rho)} \frac{M_m |\psi_i\rangle \langle\psi_i| M_m^\dagger}{\langle\psi_i| M_m^\dagger M_m |\psi_i\rangle} \\ &= \sum_i \frac{p(i) M_m |\psi_i\rangle \langle\psi_i| M_m^\dagger}{tr(M_m^\dagger M_m \rho)} \\ &= \frac{M_m \rho M_m^\dagger}{tr(M_m^\dagger M_m \rho)} \end{aligned}$$

1.4 EPR and the Bell Inequality

2 Nielsen & Chuang: Chapter 4

3 Nielsen & Chuang: Chapter 5

1. Quantum Fourier Transform
2. Quantum Phase Estimation Algorithm
3. Shor's algorithm for Factoring

4 Nielsen & Chuang: Chapter 6

Quantum Search algorithms

5 Algorithms for solving linear systems of equations

`https://arxiv.org/abs/0811.3171`

6 Density Matrix Exponentiation Algorithms

`https://www.nature.com/articles/nphys3029`

7 Review of Quantum Machine Learning

`https://www.nature.com/articles/nature23474`

8 Learnability of Quantum States

1. `https://arxiv.org/abs/quant-ph/0608142`
2. `https://arxiv.org/abs/1711.01053`
3. `https://arxiv.org/abs/1801.05721`

9 Appendix

9.1 Pauli Matrices

$$\begin{aligned} X &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ Y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ Z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

9.2 Trace of an Operator

Let $|i\rangle$ be an orthonormal basis for A and so

$$\begin{aligned} tr(A) &= \sum_i A_{ii} \\ &= \sum_i \langle i| A |i\rangle \end{aligned}$$

Hence, if we extend $|\psi\rangle$ to the orthonormal basis $|i\rangle$ which includes $|\psi\rangle$ as the first element (for example via the Gram-Schmidt procedure) then

$$\begin{aligned} tr(A |\psi\rangle \langle\psi|) &= \sum_i \langle i| A |\psi\rangle \langle\psi|i\rangle \\ &= \langle\psi| A |\psi\rangle \end{aligned}$$