Introduction to quantum information processing

Measurements and quantum probability

Brad Lackey

25 October 2016
OUTLINE

1. Probability
2. Density Operators
3. Review: the Spectral Theorem
Probability and Measurement

As taught in Quantum Mechanics

In traditional quantum mechanics courses:

1. A system is associated with a Hilbert space $\mathcal{H}$.
   - In quantum information we typically work with finite dimensional spaces.

2. Probability is determined by the “state” of the system $|\psi\rangle \in \mathcal{H}$.
   - This is a unit vector, but only determined up to “phase.”

3. Observables are given by Hermitian operators $A$, and $\langle A \rangle = \langle \psi | A | \psi \rangle$.
   - In particular, if $|\phi\rangle$ is an eigenvector of $A$ with eigenvalue $\lambda$ then
     \[
     \Pr\{\text{Getting outcome } \lambda \text{ upon measuring } A\} = \Pr\{\text{Seeing the state } |\phi\rangle \text{ upon observing } A\} = |\langle \psi | \phi \rangle|^2.
     \]
   - This produces the idea of a “measurement” basis: the eigenbasis of $A$. 

Measurements and Quantum Probability

3 of 22
OUTLINE

1. Probability

2. Density Operators

3. Review: the Spectral Theorem
**Ein Gedankenexperiment**

Alice flips a (fair) coin, and without telling us the result does the following:

- on heads, she prepares the state $|0\rangle$; and
- on tails, she prepares the state $|1\rangle$.

Then sends the resulting state. We make a measurement. What do we get?

Say our measurement basis is

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

$$|\phi^\perp\rangle = \beta^*|0\rangle - \alpha^*|1\rangle.$$

Then

$$\Pr\{\text{Seeing } |\phi\rangle\} = \frac{1}{2} |\langle 0 | \phi \rangle|^2 + \frac{1}{2} |\langle 1 | \phi \rangle|^2 = \frac{\alpha^2 + |\beta|^2}{2} = \frac{1}{2}.$$

Heads Tails

$$\Pr\{\text{Seeing } |\phi^\perp\rangle\} = \frac{1}{2} |\langle 0 | \phi^\perp \rangle|^2 + \frac{1}{2} |\langle 1 | \phi^\perp \rangle|^2 = \frac{|\beta|^2 + |\alpha|^2}{2} = \frac{1}{2}.$$
Quantum Probability

Hmmm... Alice managed to prepare a state where *any* measurement basis \(\{|\phi\rangle, |\phi\perp\rangle\}\) we use gives us the statistics:

\[
\Pr\{\text{Seeing } |\phi\rangle\} = \Pr\{\text{Seeing } |\phi\perp\rangle\} = \frac{1}{2}.
\]

So what is this state?

Suppose this state is \(|\psi\rangle\). But then measuring in the basis \(\{|\psi\rangle, |\psi\perp\rangle\}\) gives

\[
\Pr\{\text{Seeing } |\psi\rangle\} = 1 \text{ and } \Pr\{\text{Seeing } |\psi\perp\rangle\} = 0,
\]

not 50-50 as we’ve already shown it must be.

So this state is never of the form \(|\psi\rangle\)!

*Answer*: states of the form \(|\psi\rangle\) are special, called “pure” states. There’s many more quantum states than just these. Alice prepared something called a “mixed” state.
**Mixed states versus pure states**

Let figure it out what it is by pushing symbols around:

\[
\text{Pr}\{\text{Seeing } |\phi\rangle\} = \frac{1}{2} |\langle 0 |\phi \rangle|^2 + \frac{1}{2} |\langle 1 |\phi \rangle|^2
\]

\[
= \frac{1}{2} (\langle 0 |\phi \rangle)^* \langle 0 |\phi \rangle + \frac{1}{2} (\langle 1 |\phi \rangle)^* \langle 1 |\phi \rangle
\]

\[
= \frac{1}{2} \langle \phi |0\rangle \langle 0 |\phi \rangle + \frac{1}{2} \langle \phi |1\rangle \langle 1 |\phi \rangle
\]

\[
= \langle \phi | \left( \frac{1}{2} |0\rangle \langle 0 | \right) |\phi \rangle + \langle \phi | \left( \frac{1}{2} |1\rangle \langle 1 | \right) |\phi \rangle
\]

\[
= \langle \phi | \left( \frac{1}{2} (|0\rangle \langle 0 | + |1\rangle \langle 1 |) \right) |\phi \rangle
\]

\[
= \langle \phi | \rho |\phi \rangle
\]

where

\[
\rho = \frac{1}{2} (|0\rangle \langle 0 | + |1\rangle \langle 1 |).
\]
MIXED STATES VERSUS PURE STATES

What’s with this notation? Let’s really compute this mixed state:

\[
\frac{1}{2} \langle \phi | 0 \rangle \langle 0 | \phi \rangle = \frac{1}{2} \left[ \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] \left[ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right] \\
= \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} \left[ \frac{1}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right].
\]

So \(\frac{1}{2} |0\rangle \langle 0|\) is the matrix (or operator)

\[
\frac{1}{2} |0\rangle \langle 0| = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}
\]

And similarly for \(\frac{1}{2} |1\rangle \langle 1|\), and so

\[
\rho = \frac{1}{2} (|0\rangle \langle 0| + |1\rangle \langle 1|) = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \frac{1}{2} \cdot \mathbf{1}.
\]
Mixed states versus pure states

Now things are falling into place.

- The state Alice prepared is an operator $\rho = \frac{1}{2} \cdot 1$.
- We’ve shown the general probability rule

$$\text{Pr}\{\text{Seeing } |\phi\rangle\} = \langle \phi | \rho | \phi \rangle.$$

- So we simply compute: for any $|\phi\rangle$,

$$\frac{1}{2} \langle \phi | 1 | \phi \rangle = \frac{1}{2} \langle \phi | \phi \rangle = \frac{1}{2}.$$
OUTLINE

1 Probability

2 Density Operators

3 Review: the Spectral Theorem
Mixed states and ensembles

Instead, imagine Alice prepared a huge number of quantum states:

- half of them are prepared in $|0\rangle$, and the other half in $|1\rangle$;
- we take one at random and measure it.

In general: an ensemble is a collection $\{(|\psi_j\rangle, p_j)\}_{j=1}^N$

- $|\psi_j\rangle \in \mathcal{H}$ is any pure state, and
- $p_j$ is the proportion of the ensemble prepared in $|\psi_j\rangle$.

Then our statistics are given by

$$\Pr\{\text{Seeing } |\phi\rangle\} = \sum_{j=1}^N p_j |\langle \psi_j | \phi \rangle|^2$$

$$= \sum_{j=1}^N p_j \langle \phi | \psi_j \rangle \langle \psi_j | \phi \rangle = \langle \phi | \left( \sum_{j=1}^N p_j |\psi_j\rangle \langle \psi_j| \right) |\phi\rangle.$$ 

So ensembles are (generally) types of mixed states.
The state of an ensemble has the form \( \rho = \sum_{j=1}^{N} p_j |\psi_j\rangle \langle \psi_j| \).

- This operator is Hermitian since \( p_j \) is real.

**Definition**

A Hermitian operator \( \rho \) is *positive* (denoted \( \rho \geq 0 \)) if all its eigenvalues are nonnegative.

Why isn’t this called nonnegative? I don’t know.

**Proposition**

*An operator \( \rho \geq 0 \) if and only if for every \( |\phi\rangle \) we have \( \langle \phi|\rho|\phi\rangle \geq 0 \).*

We’ll prove this soon, but we can apply it now.

- The state of any ensemble is positive. In fact,

\[
\langle \phi|\rho|\phi \rangle = \sum_{j=1}^{N} p_j |\langle \psi_j|\phi \rangle|^2 = \Pr\{\text{Seeing } |\phi\rangle\} \geq 0.
\]


**TRACE**

Recall from the linear algebra:

- the trace of a matrix is the sum of its diagonal entries.
- Important fact 1: \( \text{tr}(AB) = \text{tr}(BA) \).
- Important fact 2: \( \text{tr}(c_1A + c_2B) = c_1\text{tr}(A) + c_2\text{tr}(B) \).

Fact 1 implies that the trace can be computed in any basis:

- if \( P \) is the change-of-basis matrix, then

\[
\text{tr}(P^{-1}AP) = \text{tr}((P^{-1}A)P) = \text{tr}(P(P^{-1}A)) = \text{tr}(PP^{-1}A) = \text{tr}(A).
\]

In Dirac notation the trace is expressed as follows.

- Let \( \{|\phi_j\rangle\} \) be any basis of \( \mathcal{H} \) (typically we use an orthonormal one).
- The trace of an operator is given by

\[
\text{tr}(\rho) = \sum_j \langle \phi_j | \rho | \phi_j \rangle.
\]

*(Warning: in general \( \langle \phi_j | \) refers to the “dual” basis; this can be a tricky to compute if the basis is not orthonormal!)*
**Density Operators**

The state of an ensemble is \( \rho = \sum_{j=1}^{N} p_j |\psi_j\rangle \langle \psi_j| \).

**Proposition**

*The state of any ensemble has trace one.*

Now we use an *orthonormal* basis, \( \{|\phi_k\rangle\} \). We first compute

\[
\text{tr}(|\psi_j\rangle \langle \psi_j|) = \sum_{k} \langle \phi_k | \psi_j \rangle \langle \psi_j | \phi_k \rangle = \sum_{k} |\langle \psi_j | \phi_k \rangle|^2
\]

\[
= \|\phi_j\|^2 = 1 \text{ (Pythagoras’ Theorem)}.
\]

But then

\[
\text{tr}(\rho) = \text{tr} \left( \sum_{j=1}^{N} p_j |\psi_j\rangle \langle \psi_j| \right) = \sum_{j=1}^{N} p_j \text{tr}(|\psi_j\rangle \langle \psi_j|) = \sum_{j=1}^{N} p_j = 1.
\]

**Definition**

A *density operator* is an operator \( \rho \) on \( \mathcal{H} \) with (i) \( \rho \geq 0 \) and (ii) \( \text{tr}(\rho) = 1 \).
Example: densities on the Bloch sphere

Recall the Bloch sphere:

- A qubit can be written as $|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\varphi} \sin(\theta/2)|1\rangle$,
- On the Bloch sphere it is $(r_x, r_y, r_z) = (\cos \varphi \sin \theta, \sin \varphi \sin \theta, \cos \theta)$.

Consider $|\psi\rangle$ as an ensemble by itself $(|\psi\rangle, 1)$, so $\rho = |\psi\rangle\langle\psi|$. We compute:

$$
\rho = \begin{pmatrix}
\cos(\theta/2) \\
e^{i\varphi} \sin(\theta/2)
\end{pmatrix}
\begin{pmatrix}
\cos(\theta/2) & e^{-i\varphi} \sin(\theta/2) \\
e^{i\varphi} \sin(\theta/2) \cos(\theta/2) & \sin^2(\theta/2)
\end{pmatrix}
= \frac{1}{2}
\begin{pmatrix}
1 + \cos \theta & \cos \varphi \sin \theta - i \sin \varphi \sin \theta \\
\cos \varphi \sin \theta + i \sin \varphi \sin \theta & 1 - \cos \theta
\end{pmatrix}
= \frac{1}{2}(1 + \cos \theta Z + \cos \varphi \sin \theta X + \sin \varphi \sin \theta Y)
= \frac{1}{2}(1 + r_x X + r_y Y + r_z Z).
$$

So the Bloch sphere coordinates of $|\psi\rangle$ are just the coefficients of $\rho = |\psi\rangle\langle\psi|$ with respect to the Pauli matrices.
OUTLINE

1 Probability

2 Density Operators

3 Review: the Spectral Theorem
EIGENVECTORS AND EIGENVALUES

We’re familiar with eigenvectors and eigenvalues of matrix $M \vec{v} = \lambda \vec{v}$.

- Dirac notation: $M |\psi \rangle = \lambda |\psi \rangle$.

Important fact: let $H$ be Hermitian on $\mathcal{H}$,

- then $\mathcal{H}$ has a orthonormal basis of eigenvectors for $H$, $\{ |\psi_j \rangle \}_{j=1}^{\dim \mathcal{H}}$,

and all the eigenvalues of $H$ are real.

Proposition

The operator $\Pi = |\psi \rangle \langle \psi |$ is the orthogonal projection onto $\text{span}\{ |\psi \rangle \}$.

In other words: (i) $\Pi |\psi \rangle = |\psi \rangle$ and (ii) if $\langle \psi | \phi \rangle = 0$ then $\Pi |\phi \rangle = 0$.

Using the notation above $\Pi_j = |\psi_j \rangle \langle \psi_j |$ is an eigenprojector of $H$.

- The eigenvalue equation reads $H \Pi_j = \lambda_j \Pi_j$.

- Orthogonality implies $\Pi_j \Pi_k = 0$ whenever $j \neq k$. 
SPECTRAL MEASURES

What if $H$ has degenerate eigenvalues? E.g. $H = Z \otimes 1$:
- eigenvalue $+1$ has eigenvectors $|00\rangle$ and $|01\rangle$,
- eigenvalue $-1$ has eigenvectors $|10\rangle$ and $|11\rangle$.

This would come up when we “measure the first qubit of a register.”
- The eigenprojection $\Pi_{+1}$ maps onto $\text{span}\{|00\rangle, |01\rangle\}$.
- The eigenprojection $\Pi_{-1}$ maps onto $\text{span}\{|10\rangle, |11\rangle\}$.

If $\{|\psi_k\rangle\}_{k=1}^d$ is an orthonormal basis of eigenvectors for eigenvalue $\lambda$:
- the eigenprojection is $\Pi_\lambda = |\psi_1\rangle\langle \psi_1| + \cdots + |\psi_d\rangle\langle \psi_d|$.
- Still, if $\lambda \neq \mu$ are eigenvalues then $\Pi_\lambda \Pi_\mu = 0$.

Definition

A spectral (or projection-valued) measure is a collection of orthogonal projections $\{\Pi_j\}$ such that (i) $\Pi_j \Pi_k = 0$ whenever $j \neq k$, and (ii) $\sum_j \Pi_j = 1$. 
The Spectral Theorem

**Theorem (The Spectral Theorem)**

Let $H$ be Hermitian, $\sigma(H) = \{\lambda_1, \ldots, \lambda_r\}$ be its eigenvalues, and 
$\{\Pi_1, \ldots, \Pi_r\}$ be the associated eigenprojections. Then for any function $f$
defined on $\sigma(H)$, we have $f(H) = \sum_{k=1}^{r} f(\lambda_k) \Pi_k$.

This is not just notation: consider $H^2 = \sum_{k=1}^{r} \lambda_k^2 \Pi_k$.
- I know how to calculate each side independently.
- The spectral theorem shows they always give the same answer.

More important examples:
- (spectral resolution) for $f(x) = x$ we have $H = \sum_{k=1}^{r} \lambda_k \Pi_k$;
- (square-root) if $\rho \geq 0$ and $f(x) = \sqrt{x}$ we have $\rho^{1/2} = \sum_{k=1}^{r} \sqrt{\lambda_k} \Pi_k$;
- (modulus) for $f(x) = |x|$ we have $|A| = \sum_{k=1}^{r} |\lambda_k| \Pi_k$;
- (propagation) for $f(x) = e^{ixt}$ we have $e^{iHt} = \sum_{k=1}^{r} e^{i\lambda_k t} \Pi_k$;
- (resolution-of-the-identity) for $f(x) = 1$ we have $1 = \sum_{k=1}^{r} \Pi_k$. 

**Example: Densities versus Ensembles**

We claim every density operator is just an ensemble.

- Use the spectral resolution to write \( \rho = \sum_{j=1}^{\dim \mathcal{H}} \lambda_j |\psi_j\rangle \langle \psi_j| \).
- Here, if an eigenvalue is degenerate (say \( \lambda_k \)) we split up its projection \( \Pi_k = |\psi_{k_1}\rangle \langle \psi_{k_1}| + \cdots + |\psi_{k_d}\rangle \langle \psi_{k_d}| \).
- Always \( \lambda_j \geq 0 \) since \( \rho \geq 0 \).

Trace can be computed in any basis, so use the basis \( \{ |\psi_j\rangle \} \):

\[
\text{tr}(\rho) = \sum_{jk} \lambda_j |\langle \psi_j | \psi_k \rangle|^2 = \sum_j \lambda_j = 1.
\]

Therefore \( \{(|\psi_j\rangle, \lambda_j)\} \) is an ensemble.

So we see (i) mixed states, (ii) ensembles, and (iii) density operators are all the same.
Some unfinished business:

**Proposition**

An operator $A \geq 0$ if and only if for every $|\psi\rangle$, we have $\langle \psi | A | \psi \rangle \geq 0$.

**Proof.** Suppose $\langle \psi | A | \psi \rangle \geq 0$ for all $|\psi\rangle$. Let $\lambda$ be an eigenvalue of $A$, so that $A |\phi\rangle = \lambda |\phi\rangle$ for some $|\phi\rangle$. Then

$$0 \leq \langle \phi | A | \phi \rangle = \langle \phi | \lambda | \phi \rangle = \lambda.$$  

Since $\lambda$ was arbitrary, all eigenvalues of $A$ are nonnegative.

Now suppose $A \geq 0$, and write $A = \sum_{j=1}^{n} \lambda_{j} |\phi_{j}\rangle \langle \phi_{j}|$ using the spectral theorem. Then since each $\lambda_{j} \geq 0$, every $|\psi\rangle$ has

$$\langle \psi | A | \psi \rangle = \sum_{j=1}^{n} \lambda_{j} |\langle \psi | \phi_{j}\rangle|^{2} \geq 0.$$  

□
NEXT TIME...

- Schmidt decomposition.
- Partial trace.
- State purification.
- Superoperators.