Lecture 22 - Quantum Probability

Recap of last lecture:

Mixed (qudit) state: "prob. p_i of $|\psi_i\rangle \in \mathbb{C}^d$", $i = 1 \ldots m$.

↳ encoded by density matrix $\rho \in \mathbb{C}^{d \times d}$,

$$\rho = p_1 |\psi_1\rangle\langle\psi_1| + \cdots + p_m |\psi_m\rangle\langle\psi_m|$$

Hermitian matrix ρ (meaning $\rho^\dagger = \rho$) is a density matrix iff

① "$\rho \geq 0$": "ρ is positive semidefinite" ($\langle u|\rho|u\rangle \geq 0 \;\; \forall |u\rangle$)

② "$tr(\rho) = 1$" $\sum_{i=1}^{d} \rho_{ii}$

Every Hermitian matrix M acts on $\mathbb{C}^d$ by

"stretch by factor λ_i in direction $|v_i\rangle$ $\forall 1 \leq i \leq d$"

(real) eigenvalues of M (orthonormal) eigenvectors of M

① $\Leftrightarrow$ ρ's eigvals λ_i all ≥ 0 ② $\Leftrightarrow$ $\lambda_1 + \cdots + \lambda_d = 1$

$\therefore$ any density mtx ρ equiv to mixed state "λ_i prob. of $|v_i\rangle$", $i = 1 \ldots m$.

Actions on (mixed) states:

- Applying unitary transformation $U \in \mathbb{C}^{d \times d}$:

$$\rho \longmapsto U \rho U^{\dagger}$$

- Measuring in basis $|u_1\rangle, \ldots, |u_d\rangle$:

  get outcome "i" with prob. $\langle u_i | \rho | u_i \rangle$

  [[And state collapses, changing ρ. You can work out formulas, but we won't bother about this now. Today, we'll never use a state again after measuring it.]]

- Adjoining two states:

  If ρ is density mtx of a d-dim. mixed state,
  σ ~~~ e-dim. ~~~,
  & we decide to call the whole system one de-dim. state, its density mtx is .... $\rho \otimes \sigma$. (exercise/homework)

  [[Kronecker prod]]

Linear Algebra Interlude

prop: Let $A \in \mathbb{C}^{d \times d'}$, $B \in \mathbb{C}^{d' \times d}$. Then

$$tr(AB) = tr(BA).$$

proof:

$$tr(AB) = \sum_{i=1}^{d} (AB)_{ii} = \sum_{i=1}^{d} \sum_{j=1}^{d'} A_{ij} B_{ji}$$

$$= \sum_{j=1}^{d'} \sum_{i=1}^{d} B_{ji} A_{ij}$$

$$tr(BA) = \sum_{j=1}^{d'} (BA)_{jj}$$

Classic trick: $\langle u_i | \rho | u_i \rangle = tr(\underbrace{\langle u_i |}_{A} \underbrace{\rho | u_i \rangle}_{B})$ 1x1 matrix!

[Recall: this is prob["i"] when measuring in $|u_1\rangle, \ldots, |u_d\rangle$ basis]

$$= tr(\rho \underbrace{|u_i \rangle\langle u_i|}_{d \times d \text{ "proj. onto } |u_i\rangle\text{" mtx}})$$

Also: $tr(A^\dagger B) = \sum_{ij} (A^\dagger)_{ij} B_{ji} = \underbrace{\sum_{ij} A_{ji}^* B_{ji}}$ [Could reverse i, j here]

"dot product of A, B when they're viewed as lists of #'s"

def: $\langle A, B \rangle = \cdot$. "Inner product for matrices".

[Don't worry much about the dagger! Often A, B will be Hermitian, so $A = A^\dagger$.]

Today: "Quantum Probability 101"

(vs. "Classical Probability 101")

Classical Def #1: A d-outcome prob. distrib. is a vector $p \in \mathbb{R}^d$ with $p \geq 0$, $\sum_{i=1}^{d} p_i = 1$.

[Outcomes called $1, 2, \ldots, d$; $p_i = \Pr[i]$]

Quantum Def #1: A d-dim. state/density mtx is a Hermitian $\rho \in \mathbb{C}^{d \times d}$ with $\rho \geq 0$, $\sum_{i=1}^{d} \rho_{ii} = 1$.

[Think of ρ as a "source of quantum randomness"]

[classical → quantum : replace "vector" by "Hermitian matrix"

quantum → classical : take every Hermitian matrix to be diagonal]

Example:

Recall: if state ρ measured in $\{|u_1\rangle, \dots, |u_d\rangle\}$ basis, $\Pr[\text{"}i\text{"}] = \langle u_i|\rho|u_i\rangle$

$$= tr(\rho\, |u_i\rangle\langle u_i|)$$

⟦Needn't worry about dagger†, ∵ $\rho^\dagger = \rho$⟧ $\longrightarrow$ $= \langle \rho, E_i \rangle$, $\underbrace{E_i := |u_i\rangle\langle u_i|}$

("proj. onto $|u_i\rangle$" mtx)

Remark: $E_i \geq 0$ & $E_1 + \dots + E_d = I_{d\times d} = \mathbb{1}_{d\times d}$

⟦why?⟧

⟦It's a pure-state dens. mtx⟧

⟦Matrix that stretches by 1 in every direction: all eigvals are 1.⟧

⟦Two notations for the $d\times d$ identity matrix⟧

Compare: Given prob. dist $p \in \mathbb{R}^d$, let

$$e_i = (0, \dots, 0, 1, 0, \dots, 0)$$

↑ ith position.

⟦"Indicator of ith outcome"⟧

$p_i = \Pr_p[\text{"}i\text{"}] = \langle p, e_i \rangle$.

And $e_i \geq 0$, $e_1 + \dots + e_d = (1, 1, \dots, 1) = \text{"}\vec{\mathbb{1}}\text{"}$.

Probability 101 def. #2: Events

[[My descriptions here will be a bit weird, to accommodate the subsequent quantum analogy.]]

Given prob. dist $p \in \mathbb{R}^d$, let's define some "mutually exclusive, collectively exhaustive" events
[[meaning always exactly one happens]]

$\text{Event}_1, \ldots, \text{Event}_m$

Naivest: $m = d$, Event_i = "i is drawn from p"
[[like in preceding example]]

Lumping: Some outcomes grouped together.

E.g. Event_1 = "the draw from p is odd"
Event_2 = " ~~~~~~~~ even"

Identify with "indicator vectors":

$$e_1 = (1, 0, 1, 0, \ldots\ldots)$$
$$e_2 = (0, 1, 0, 1, \ldots\ldots)$$

$\Pr_p[\text{Event}_1] = p_1 + p_3 + p_5 + \cdots = \langle p, e_1 \rangle$; $\Pr_p[\text{Event}_2] = \langle p, e_2 \rangle$

Note: e_i's nonneg. & sum to $\vec{\mathbb{1}} = (1, 1, \ldots, 1)$.

[[This ensures $\langle p, e_i \rangle \geq 0$ & $\sum_i \langle p, e_i \rangle = 1$; exactly one Event occurs.]]

[In "lumping", you have $\le d$ events. Now we'll describe a way to have $> d$ events. Involves injecting additional randomness.]

Using additional randomness:

Suppose $e_1, \dots, e_m$ are any "indicator" vectors in $\mathbb{R}^d$ with $e_i \geq 0$ [entrywise] & $e_1 + \dots + e_m = \vec{\mathbb{1}}$.

[in quotes b/c need not be 0/1]

E.g.: $e_1 = (0, .2, .3, 0, \dots)$

$e_2 = (1, .7, .2, 0, \dots)$

$e_3 = (0, .1, .5, 1, \dots)$

[$m < d$ here, but could easily have $m > d$]

Event$_1$, Event$_2$, Event$_3$??

Think: Draw from p.

If outcome 1 : Event$_2$ occurs.

If outcome 2 : $\begin{cases} \text{Event}_1 \text{ occurs w.p. } .2 \\ \text{Event}_2 \text{ ——— } .7 \\ \text{Event}_3 \text{ ——— } .1 \end{cases}$

If outcome 3 : $\{$ Event$_1$ occurs w.p. .3
etc.

[Yes, you need additional randomness to implement, But still makes sense to say "exactly one event always occurs", and...]

$$\Pr[\text{Event}_j] = \sum_{i=1}^{d} p_i (e_j)_i = \langle p, e_j \rangle.$$

Quantum Generalization

How can you measure [[and thereby get classical outcomes]] a qudit state ρ?

- in orthonormal basis $\rightarrow$ d outcomes
- "lumping": Say $d = 2^q$, ρ has q qubits.
  Can do a "partial measurement" of r qubits: yields $2^r < d$ outcomes
- "additional randomness": Can add additional [["ancilla"]] s qubits, do a 2^{r+s}-dim unitary, partially/fully measure: up to $2^{r+s} > d$ outcomes

[[Can also do non-powers-of-2 things]]

Most general: Think of whole procedure as a "BODM": Big Ol' Device for Measuring.

Call readouts/outcomes "1", "2", ..., "m"

[[m may be $<d$, $=d$, $>d$]]

Turns out: [[if you do the math, which we won't]]

BODM yields m Hermitian $E_1, \dots, E_m \in \mathbb{C}^{d \times d}$

satisfying: (i) "$E_i \geq 0$" $\forall i$ (positive semidefinite)

$\langle w | E_i | w \rangle \geq 0 \;\; \forall |w\rangle$

(ii) $E_1 + \dots + E_m = \mathbb{1}_{d \times d}$ (identity matrix)

<u>And</u>: $\Pr[\text{BODM applied to } \rho \text{ reads out "}i\text{"}] = \langle \rho, E_i \rangle$.

<u>ex</u>: check (i) $\Rightarrow \langle \rho, E_i \rangle \geq 0 \;\; \forall i$ (uses $\rho \geq 0$)

(ii) $\Rightarrow \langle \rho, E_1 \rangle + \dots + \langle \rho, E_m \rangle = 1$ (uses $\sum_i \rho_{ii} = 1$)

Conversely: given Hermitian $E_1, \dots, E_m$ satisfying (i), (ii), can in principle build a physical BODM with associated behavior.

Real name of BODM is "POVM".

(You'll see "POVM" a lot. Stands for "Positive Operator-Valued Measure" for some abstruse math reason)

[[Anyway, POVM $\{E_1, \dots, E_m\}$ is quantum generalization of M.E.C.E. events.]]

<u>Question:</u> If POVM registers outcome "i", how does state ρ collapse?

<u>Answer</u>: Depends on how it's implemented! Can't tell just from $E_1, \dots, E_m$.

[If you want to know, need to know the circuit implementing the measurement. Then, once you know this, can work out the formula. For posterity, if (nonuniquely) $E_i = M_i^\dagger M_i$ for $d \times d$ M_i, then you can implement it in such a way that upon reading out "i", ρ collapses to

$$\frac{M_i \rho M_i^\dagger}{\langle \rho, M_i M_i^\dagger \rangle}.$$

↑ see homework

But generally, when people discuss POVMs, they usually only care about measurement outcome probabilities, and don't plan to continue "using" the state.]

Probability 101 def #3: Random variables

Given prob. distribution $p \in \mathbb{R}^d$, a random variable is just a real number x_i for each outcome $1 \le i \le d$.

So it's any old vector $x \in \mathbb{R}^d$.

Expected value is $E_p[x] = \sum_{i=1}^{d} p_i x_i = \langle p, x \rangle$.

Quantum generalization: "observable"

for d-dim. states ρ:

Any old Hermitian matrix $X \in \mathbb{C}^{d \times d}$.

recall: can express $X = \sum_{i=1}^{d} x_i \, |u_i \rangle\langle u_i|$

real eigenvalues

projection onto eigenvec. $|u_i\rangle$

$$X = \sum_{i=1}^{d} x_i \, |u_i\rangle\langle u_i|$$

real eigenvalues

projection onto eigenvec. $|u_i\rangle$

Physically, could build an instrument that:
- measured in $|u_1\rangle, \ldots, |u_d\rangle$ basis
- on outcome "i", read out real # x_i

If you applied this instrument to ρ, what's expected value of readout?

$$\sum_{i=1}^{d} \underbrace{\langle \rho, |u_i\rangle\langle u_i| \rangle}_{\Pr[\text{measure "}i\text{"}]} x_i = \langle \rho, \sum_{i=1}^{d} x_i |u_i\rangle\langle u_i| \rangle$$

$$= \langle \rho, X \rangle \quad ☺$$

[[Matches classical situation.]]

<u>Notation:</u> $E_\rho[X]$ ← [[Good notation: linearity of expectation holds.]]

<u>Rem:</u> X^2 operator stretches by x_i^2 factor in direction $|u_i\rangle$. I.e., it's $\sum_{i=1}^{d} x_i^2 |u_i\rangle\langle u_i|$

$\therefore$ expected $(\text{readout})^2 = E_\rho[X^2]$.

[[This is not tautological! We had to check it. But further reassures that notation is good.]]

Quantum probability theory:

sources of randomness: ρ ✓

events: POVMs ✓

random: observables ✓

Done ☺ [Really, what else is there in Probability 101? ☺]

Warning: If X, Y are observables (Hermitian matrices)

$XY \neq YX$ necessarily.

Indeed, "quantum probability" often called "noncommutative probability".

[Holds iff XY is Hermitian; i.e., iff XY is itself an "observable".]

[Actually, of course, there's much much more. From quantum probability, can develop quantum.... • statistics
• information theory
• learning theory
• .....

We'll see a little statistics on homework, including... the "Uncertainty Principle",

$$\mathrm{stddev}_\rho[X] \cdot \mathrm{stddev}_\rho[Y] \geq \left| E_\rho\left[\tfrac{i}{2}(XY - YX) \right] \right|.$$

Some info theory in next lecture.

My research is about learning & statistics of quantum states; e.g., the "tomography = state learning" problem: How many copies, n, of state ρ are needed to approximately learn it, by applying POVMs to $\underbrace{\rho \otimes \rho \otimes \cdots \otimes \rho}_{(n \text{ times})}$?"]