8.S372/18.S996 Quantum Information Science III **Fall 2020**

Lecture 3: Sep 8, 2020

Lecturer: Aram Harrow *Scribe: Zane Rossi, Andrey Boris Khesin*

3.0.3 Norms

Given a tuple $x \in \mathbb{C}^d$ we define

$$\|x\|_{\ell_p} \equiv \left(\sum_{i=1}^{d} |x_i|^p \right)^{1/p} = \|x\|_p,$$

is known as the ℓ_p norm. For ℓ_1 this is useful for probability distributions (where the norm is required to be 1) while ℓ_2 is useful for pure quantum states (where the norm of a valid state vector is again required to be 1). The ℓ_∞ norm finds use when describing classical observables.

We can also define norms over matrices

$$\|M\|_{S_p} \equiv \|\text{sing. val. of } M\|_{\ell_p},$$

is known as the *Schatten p-norm.* Note that the $\|M\|_{S_\infty}$ is simply the largest of the absolute values of the singular values of M.

This is relevant for measurements, e.g., $\{M, I - M\}$ is a legal set of measurements iff $0 \leq M \leq I$ in the *PSD ordering* iff M is PSD and its Schatten-∞ norm is less than or equal to 1. Note that $A \geq B$ iff $A - B$ is PSD (possitive-semi-definite).

We can consider the Schatten-2 norm, which is nice operationally,

$$\|M\|_{S_2} \equiv \|\text{vec}(M)\|_{\ell_2} = \sqrt{\text{tr}(M^\dagger M)},$$

We can also consider the Schatten-1 norm,

$$\|M\|_{S_1} \equiv \sum_i |\lambda_i| = \text{tr}(M),$$

where the last equality applies only if the singular values are positive (this means that M is PSD). This is equivalently $\text{tr}\sqrt{(M^\dagger M)}$ if M is a Hermitian operator. All of these results can be seen by taking singular value decompositions of the operator M.

The reductions given above are part of a more general scheme of defining inner products over matrices, e.g.,

$$\mathrm{tr} A^\dagger B = \sum_{ij} A^*_{ij} B_{ij} = \langle A, B\rangle.$$

This apparently direct analogy between ℓ_p and S_p norms is often useful, and can be a good technique for visualization.

We define the *dual norm* as the result, given a norm $\|\cdot\|_*$, the norm

$$\|x\|_{\mathrm{dual},*} \equiv \max_{\|y\|_* \le 1} |\langle x, y\rangle|.$$

The ℓ_2 norm is dual to itself. We can establish

$$|\langle x, y\rangle| \le \|x\|_{\ell_2}\|y\|_{\ell_2} \le \|x\|_{\ell_2},$$

while the maximum is achieved by $y = x/\|x\|_{\ell_2}$. This establishes duality of the norm with itself. Note that the direct analogy to this choice does not always hold for other norms.

Note that, under our level of rigour, the dual operation is its own inverse.

We can summarize dualities between ℓ_p norms by the following relation,

$$\|\cdot|_{\ell_p,\mathrm{d}} = \|\cdot|_{\ell_q},$$

iff $1/p + 1/q = 1$, which is closely related to the Hölder inequality.

For matrices we use the Hilbert-Schmidt inner product

$$\langle A, B\rangle = \mathrm{tr}(A^\dagger B) = \langle \mathrm{vec}(A), \mathrm{vec}(B)\rangle.$$

We also present the following fact without proof.

$$\|M\|_{S_1} = \max_{U \mathrm{unitary}} |MU|$$

3.0.4 Comparing probability distributions

Total variation distance

$$T(p, q) = \frac{1}{2}\|p - q\|_1 = \max_S \left[p(S) - q(S)\right],$$

where this maximum is taken over subsets of elements,

$$\sum_x |p(x) - q(x)|.$$

Sometimes we want something that looks like the ℓ_2 norm, though, because of its nice, seemingly geometric properties, but the standard inner product is not great: summing $p(x)q(x)$ over events is generally less than 1.

We can instead consider the fidelity

$$\langle \sqrt{p}, \sqrt{q} \rangle = \sum_x \sqrt{p(x)q(x)} = F.$$

Properties of this include $1 - F \leq T \leq \sqrt{2(1-F)}$, as well as $F(p_1 \otimes p_2, q_1 \otimes q_2) = F(p_1, q_1)F(p_2, q_2)$, which is a much nicer behavior than how T acts over tensor products. We also note $F(p, q) \leq 1$ and attains equality iff $p = q$.

These facts help us to prove asymptotic statements like the one below

$$1 - T(p^{\otimes n}, q^{\otimes n}) \sim e^{-cn},$$

where $e^{c_1 n} \leq 1 - T \leq e^{c_2 n}$

3.0.5 Comparing quantum states

The naive natural distance here is the ℓ_2 norm

$$\| |\alpha\rangle - |\beta\rangle \|_{\ell_2} = \sqrt{2(1 - \mathrm{Re}(\langle \alpha | \beta \rangle))},$$

but this ignores phase and can allow identical states to appear far from each other. Ignoring this we recover the familiar $|\langle \alpha | \beta \rangle|$, which resolves this ambiguity.

For density matrices the trace distance seems natural

$$T(\rho, \sigma) = \|\rho - \sigma\|_{S_1},$$

which is the maximum over 2-outcome measurements M of $\mathrm{tr}(M(\rho - \sigma))$.

We can ask questions now like what happens if we apply a noisy quantum operation?

$$T(N(\rho), N(\sigma)) \quad \text{vs} \quad T(\rho, \sigma)?$$

The easiest way to think about this, says the lecturer, is an isometry followed by a partial trace. Here $N(\rho) = \mathrm{tr}_E(V \rho V^\dagger)$, where V is an isometry iff $V^\dagger V = I$. We know that the isometry cannot change the trace distance, as it implies a corresponding change in the optimal measurement $M \mapsto V^\dagger M V$. The trace over the environment can certainly not increase the chance of distinguishing the two states, and a proof might stem from considering the natural ideal measurement $I \otimes M$ on the entire system with respect to the ideal measurement M on the relevant (i.e., non-traced over) system.

The lecturer makes a comment on another proof of this fact relating operator norms, e.g., that we might alternatively measure $\|N\|_{S_1 \to S_1} \leq 1$.

We can also consider the *mixed state fidelity* or Schatten-1 norm

$$\|\sqrt{\rho}\sqrt{\sigma}\|_{S_1} = \mathrm{tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\sigma}},$$

which is used incredibly rarely. This can make some more natural sense if σ is a pure state. Note also that this definition reduces to $|\langle\psi|\phi\rangle|$ in the case of two pure states, and thus might reasonably generalize fidelity to mixed states.

On the problem set we will show $T^2 + F^2 = 1$ for pure states. We might again ask what happens to this measure if we act on it with a quantum channel:

$$F(\rho, \sigma) \leq F(N(\rho), N(\sigma)).$$

Theorem 3 *Uhlmann's theorem: $F(\rho, \sigma) = \max(|\langle\alpha|\beta\rangle|)$ over $|\alpha\rangle_{AB}$ and $|\beta\rangle_{AB}$ which agree with ρ and σ by partial trace over subsystem A (there are many such purifications possible).*

Note that if $|\Gamma\rangle = |i\rangle_A \otimes |i\rangle_B$ (summation implied) is the maximally entangled state, then $\mathrm{tr}_B(\Gamma) = I_A$ the maximally mixed state (unnormalized).

We also introduce the canonical purification $|\phi^\rho\rangle = (\sqrt{\rho} \otimes I_B)\,|\Gamma\rangle$. We can check that this is both normalized and a valid purification. This gives us an equivalent formulation of Uhlmann's theorem.

Theorem 4 *The alternative form of Uhlmann's theorem is given with proof below:*

$$\begin{aligned}
F(\rho, \sigma) &= \max_U |\langle\phi^\rho| I \otimes U |\phi^\rho\rangle| \\
&= \max_U |\langle\Gamma|(\sqrt{\rho} \otimes I)(I \otimes U)(\sqrt{\sigma} \otimes I)|\Gamma\rangle| \\
&= \max_U |\langle\Gamma|(\sqrt{\rho}\sqrt{\sigma} \otimes U)|\Gamma\rangle| \\
&= \max_U |\mathrm{tr}\sqrt{\rho}\sqrt{\sigma}U^T| \\
&= \|\sqrt{\rho}\sqrt{\sigma}\|_1 \\
&= F
\end{aligned}$$

There are many corollaries to this result, including the Fuch's von-Graf inequalities

$$1 - T \leq F \leq \sqrt{1 - T^2}$$

as well as stronger results against the no-go theorem for quantum bit commitment.