Introduction to Quantum Information Theory

Based on *Wilde, MM. (Cambridge University Press, Cambridge) (2013).*

Ali Almasi

Institut Polytechnique de Paris

Entropy and Information

Von Neumann Entropy

Definition

The entropy $H(A)_\rho$ of a state $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ is defined as follows:

$$H(A)_\rho \equiv -\operatorname{Tr}\{\rho_A \log \rho_A\}.$$

Von Neumann Entropy

Definition

The entropy $H(A)_\rho$ of a state $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ is defined as follows:

$$H(A)_\rho \equiv -\operatorname{Tr}\{\rho_A \log \rho_A\}.$$

Some remarks:

- Suppose that the spectral decomposition of a state ρ_A is

$$\rho_A = \sum_x p_X(x) |x\rangle \langle x|_A .$$

  Then $H(A)_\rho = H(X)$.
- $H(\rho) \geq 0$.
- $\min_{\rho_A} H(\rho) = 0$ and $\max_{\rho_A} H(\rho) = \log d$.
- $H(\rho)$ is concave and it is invariant under isometries.

Joint Quantum Entropy

Definition

For a density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ living in a bipartite system AB, we define:

$$H(AB)_\rho \equiv -\operatorname{Tr}\{\rho_{AB} \log \rho_{AB}\}$$

Joint Quantum Entropy

Definition

For a density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ living in a bipartite system AB, we define:

$$H(AB)_\rho \equiv -\operatorname{Tr}\{\rho_{AB} \log \rho_{AB}\}$$

Some remarks:

- For a pure state $|\phi\rangle_{AB}$, $H(A)_\phi = H(B)_\phi$, while $H(AB)_\phi = 0$.
- $H(\rho_A \otimes \sigma_B) = H(\rho_A) + H(\sigma_B)$.
- For a state of the form $\rho_{XB} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_B^x$, the joint entropy $H(XB)_\rho$ is as follows:

$$H(XB)_\rho = H(X) + \sum_x p_X(x) H(\rho_B^x).$$

Conditional Quantum Entropy and Mutual Information

Definition

For $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the conditional quantum entropy is defined as:

$$H(A \mid B)_\rho \equiv H(AB)_\rho - H(B)_\rho.$$

Conditional Quantum Entropy and Mutual Information

Definition

For $\rho_{AB} \in \mathcal{D}\left(\mathcal{H}_A \otimes \mathcal{H}_B\right)$, the conditional quantum entropy is defined as:

$$H(A \mid B)_\rho \equiv H(AB)_\rho - H(B)_\rho .$$

Some remarks:

- For a state of the form $\rho_{XB} \equiv \sum_x p_X(x) |x\rangle \langle x|_X \otimes \rho_B^x$, $H(B \mid X)_\rho = \sum_x p_X(x) H\left(\rho_B^x\right)$.
- The conditional entropy can be negative!

Conditional Quantum Entropy and Mutual Information

Definition

For $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the conditional quantum entropy is defined as:

$$H(A \mid B)_\rho \equiv H(AB)_\rho - H(B)_\rho.$$

Some remarks:

- For a state of the form $\rho_{XB} \equiv \sum_x p_X(x) |x\rangle \langle x|_X \otimes \rho_B^x$, $H(B \mid X)_\rho = \sum_x p_X(x) H(\rho_B^x)$.
- The conditional entropy can be negative!

Definition

The quantum mutual information of a state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is defined as:

$$I(A;B)_\rho \equiv H(A)_\rho + H(B)_\rho - H(AB)_\rho.$$

Quantum Relative Entropy

Definition

The quantum relative entropy $D(\rho\|\sigma)$ between a density operator $\rho \in \mathcal{D}(\mathcal{H})$ and a positive semi-definite operator $\sigma \in \mathcal{L}(\mathcal{H})$ is defined as follows:

$$D(\rho\|\sigma) \equiv \text{Tr}\{\rho[\log\rho - \log\sigma]\}$$

if the following support condition is satisfied

$$\text{supp}(\rho) \subseteq \text{supp}(\sigma),$$

and it is defined to be equal to $+\infty$ otherwise.

Theorem

For any two density operators $\rho, \sigma \in \mathcal{D}(\mathcal{H})$*,*

$$D(\rho\|\sigma) \geq 0,$$

and $D(\rho\|\sigma) = 0$ *if and only if* $\rho = \sigma$*.*

Quantum Evolution

Quantum Evolutions

What is the most general formulation of quantum evolution?

- A quantum evolution Φ should be a linear map acting on the space of (density) matrices.
- A quantum evolution should take quantum states to quantum states.

Quantum Channels

Definition

Let $\mathcal{H}_A, \mathcal{H}_B$ be Hilbert spaces. Then a linear map $\Phi : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$, is a quantum channel from system A to system B if it satisfies

1. $\text{Tr}[X] = \text{Tr}[\Phi(X)] \quad$ for all $X \in \mathcal{L}(\mathcal{H}_A)$.
2. For any additional Hilbert space $\mathcal{H}_C$ and bipartite operator $X_{AC} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C)$ with $X_{AC} \geq 0$ we have

$$(\Phi \otimes \mathcal{I}_C)(X_{AC}) \geq 0.$$

Kraus-Choi Representation of Quantum Channels

Theorem

Let $\Phi : \mathcal{L}(A) \to \mathcal{L}(B)$ be a linear map. Then the following are equivalent.

1. *Φ is a quantum channel.*
2. *(Kraus representation) There exist matrices $K_i \in \mathcal{L}(A, B)$ with $\sum_i K_i^\dagger K_i = I_A$ such that $\Phi(X) = \sum_i K_i X K_i^\dagger$.*
3. *Define the Choi matrix $C \in \mathcal{L}(AB)$ of the map Φ by*

$$C_\Phi := \begin{pmatrix} \Phi(|0\rangle\langle 0|) & \Phi(|0\rangle\langle 1|) & \dots & \Phi(|0\rangle\langle d-1|) \\ \Phi(|1\rangle\langle 0|) & \Phi(|1\rangle\langle 1|) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \Phi(|d-1\rangle\langle 0|) & \dots & \dots & \Phi(|d-1\rangle\langle d-1|) \end{pmatrix}.$$

Then $C_{AB} \geq 0$ and $\mathrm{Tr}_B\left[C_{AB}\right] = I_A$.

Examples of Quantum Channels

- **Quantum bit-flip channel**: For $p \in [0, 1]$:

$$\rho \to (1 - p)\rho + pX\rho X$$

- **Quantum erasure channel**: For $p \in [0, 1]$:

$$\rho \to (1 - p)\rho + p|e\rangle\langle e|,$$

  where $\langle e|\rho|e\rangle = 0$ for all inputs ρ.

- **Unitary mixture channel**: For unitary U_i's and a probability distribution $\{p_i\}$:

$$\Phi(X) = \sum_{i=1}^{n} p_i U_i X U_i^\dagger$$

- **Partial trace**.

Quantum Channels as Generalizations of Classical Channels

Any discrete channel with the set of conditional distributions $p_{Y|X}(y \mid x)$ can be implemented by a quantum channel with the following Kraus operators

$$\left\{ \sqrt{p_{Y|X}(y \mid x)} |y\rangle\langle x| \right\}_{x,y}.$$

Monotonicity of Quantum Relative Entropy

Theorem

Let $\rho \in \mathcal{D}(\mathcal{H}), \sigma \in \mathcal{L}(\mathcal{H})$ *be positive semi-definite, and* $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ *be a quantum channel.*

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)).$$

Schumacher's Quantum Data Compression

Shannon's Source Coding Theorem

For an i.i.d source X, if we denote a sequence of N outputs of this source by $X^{\otimes N}$, we have:

$$\text{For all } \delta \in (0,1),$$

$$\frac{1}{N} H_\delta\left(X^{\otimes N}\right) \underset{N \to +\infty}{\longrightarrow} H(X).$$

What if we have a quantum source?

Quantum Setting

A source: A Hilbert space $\mathcal{H}$ and a density matrix ρ over $\mathcal{H}$.

$$\rho = \sum_x p(x)|x\rangle\langle x|$$

A compression scheme of rate R:

- The compression operation: a quantum operation $\mathcal{C}^n$, taking states in $H^{\otimes n}$ to states in a 2^{nR}-dimensional state space, the compressed space.
- The decompression operation: takes states in the compressed space to states in the original state space.

Our criteria for reliability: In the limit of large n the entanglement fidelity $F(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n)$ should tend towards one.

Fidelity $F(\rho, \sigma)$ between density matrices ρ and σ is defined as

$$F(\rho, \sigma) \equiv \|\sqrt{\rho}\sqrt{\sigma}\|_1^2.$$

Typical Subspaces

An ϵ-typical sequence:

$$\left|\frac{1}{n}\log\left(\frac{1}{p\left(x_1\right)p\left(x_2\right)\ldots p\left(x_n\right)}\right) - S(\rho)\right| \leq \epsilon.$$

An ϵ-typical state: A state $|x_1\rangle|x_2\rangle\ldots|x_n\rangle$ for which the sequence $x_1, x_2, \ldots, x_n$ is ϵ-typical.

The ϵ-typical subspace: The subspace $T(n,\epsilon)$, spanned by all ϵ-typical states, $|x_1\rangle\ldots|x_n\rangle$.

The projector onto the ϵ-typical subspace: The projector $P(n,\epsilon)$ which is defined as:

$$P(n,\epsilon) = \sum_{x \text{ is } \epsilon\text{-typical}} |x_1\rangle\langle x_1| \otimes |x_2\rangle\langle x_2| \otimes \ldots |x_n\rangle\langle x_n|$$

Typical Subspace Theorem

Theorem

1. *Fix* $\epsilon > 0$*. Then for any* $\delta > 0$*, for sufficiently large* n*,*

$$\mathrm{tr}\left(P(n,\epsilon)\rho^{\otimes n}\right) \geq 1 - \delta.$$

2. *For any fixed* $\epsilon > 0$ *and* $\delta > 0$*, for sufficiently large* n*, the dimension* $|T(n,\epsilon)| = \mathrm{tr}(P(n,\epsilon))$ *of* $T(n,\epsilon)$ *satisfies:*

$$(1-\delta)2^{n(S(\rho)-\epsilon)} \leq |T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}.$$

3. *Let* $S(n)$ *be a projector onto any subspace of* $H^{\otimes n}$ *of dimension at most* 2^{nR}*, where* $R < S(\rho)$ *is fixed. Then for any* $\delta > 0$*, and for sufficiently large* n*,*

$$\mathrm{tr}\left(S(n)\rho^{\otimes n}\right) \leq \delta.$$

Quantum Data Compression

Theorem

Suppose that ρ_A is the density operator corresponding to a quantum information source. Then the quantum entropy $H(A)_\rho$ is equal to the quantum data compression limit of ρ.

- Main idea for quantum data compression: measure typical subspace. Successful with probability $1 - \varepsilon$.
- If successful, perform a unitary that rotates typical subspace to space of dimension $\leq 2^{n[H(\rho)+\delta]}$ ($n[H(\rho) + \delta]$ qubits).
- Send qubits to Bob, who then undoes the compression unitary.
- Scheme is guaranteed to meet the fidelity criterion.

Holevo Bound

Accessible Information

- Alice prepares an ensemble $\mathcal{E} \equiv \{p_X(x), \rho_x\}$.
- Bob performs a POVM $\{\Lambda_y\}$.
- Bob wants to retrieve as much information as possible about the random variable X.
- Bob can choose which measurement he would like to perform.
- It would be good for Bob to perform the measurement that maximizes his information about X.

Accessible Information

- Alice prepares an ensemble $\mathcal{E} \equiv \{p_X(x), \rho_x\}$.
- Bob performs a POVM $\{\Lambda_y\}$.
- Bob wants to retrieve as much information as possible about the random variable X.
- Bob can choose which measurement he would like to perform.
- It would be good for Bob to perform the measurement that maximizes his information about X.

Definition

The accessible information $I_{\text{acc}}(\mathcal{E})$ of the ensemble $\mathcal{E}$ is defined as:

$$I_{\text{acc}}(\mathcal{E}) \equiv \max_{\{\Lambda_y\}} I(X;Y),$$

where the marginal density $p_X(x)$ is that from the ensemble and the conditional density $p_{Y|X}(y \mid x) = \text{Tr}\{\Lambda_y \rho_x\}$.

Holevo Theorem

Definition

The Holevo information of an ensemble $\mathcal{E}$ is defined as:

$$\chi(\mathcal{E}) \equiv H(\rho) - \sum_{x} p_X(x) H(\rho^x) \, .$$

Theorem

$$I_{acc}(\mathcal{E}) \leq \chi(\mathcal{E}).$$

Conclusion

Conclusion

- We can define information theoretic quantities in the quantum realm analogous to the classical definitions, but some radical departures from the classical notions may arise.
- We need to generalize the notion of quantum evolution to be able to capture the notion of noisiness of an evolution.
- In quantum setting, entropy still governs the ultimate limit of the compression rate.

Thank you!