Quantum info in TCS
Lecture 5: Data processing inequality

Data processing inequality Recall that the quantum relative entropy between two two density operators ρ and σ is defined as

$$D(\rho\|\sigma) := \operatorname{tr}\rho(\log\rho - \log\sigma) \tag{1}$$

We can extend this definition to when ρ is any positive operator. The data processing inequality states that for any quantum channel $\mathcal{N}$, we have

$$D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \leq D(\rho\|\sigma) \tag{2}$$

The intuition is that when you process your data, you loose information. Therefore, different states get closer to each other under any information processing. The same inequality holds for other distance measures such as trace norm and

Applications

1. **(Strong) Subadditivity of entropy** If X and Y are two random variables, one can check that $H(X|Y) = H(X,Y) - H(Y)$. So it makes sense to define the *quantum* conditional entropy for a bipartate density operator ρ_{AB} as

   $$H(A|B)_\rho = H(\rho_{AB}) - H(\rho_B) \tag{3}$$

   where $H(\rho) := -\operatorname{tr}(\rho\log\rho)$. This definition coincides with classical definition if ρ_{AB} is classical. For classical definition, we always have $0 \leq H(X|Y) \leq H(X)$. The first inequality saturates when X is a function of Y and the second inequality saturates when X is independent of Y. However, for quantum definition, $H(A|B)_\rho$ can be negative. For example, take ρ_{AB} to be the maximally entangled state. Intuitively, when A and B are entangled, the correlation between A and B is stronger than when A is just a classical function of B.

   The upper bound $H(A|B)_\rho \leq H(A)_\rho$ still holds (known as subadditivity of entropy). To see this, we can write

   $$H(A)_\rho - H(A|B)_\rho = D(\rho_{AB}\|\rho_A \otimes \rho_B) \geq 0, \tag{4}$$

   where the last inequality is due to the positivity of quantum relative entropy.

   Now consider a tripartite quantum density operator ρ_{ABC}. Strong subadditivity of entropy states that $H(A|BC)_\rho \leq H(A|B)_\rho$. We cannot prove this by just using positivity of quantum relative entropy. However, note that

   $$H(A|BC)_\rho = -D(\rho_{ABC}\|\mathbb{1}_A \otimes \rho_{BC}) \tag{5}$$
   $$H(A|B)_\rho = -D(\rho_{AB}\|\mathbb{1}_A \otimes \rho_B). \tag{6}$$

   Therefore, applying data processing inequality for quantum channel partial trace wrt to C, we obtain strong subadditivtvity of quantum entropy.

2. **Holevo bound** Consider a classical random variable X and some quantum side information A about X. In general, this can be described by a classical-quantum state $rho_{XA} = \sum_x P_X(x) |x\rangle\langle x| \otimes \rho_A^x$. Now you perform a POVM $\{M_y\}_y$ on register A and obtain a classical random variable Y. If ρ_{XY} is (classical) density operator of random variables X, Y, we can write

$$\rho_{XY} = (\mathbb{1}_X \otimes \mathcal{M})(\rho_{XA}) \tag{7}$$

for quantum channel

$$\mathcal{M}(\rho) = \sum_y \mathrm{tr}(M_y \rho) |y\rangle\langle y| , \tag{8}$$

describing POVM $\{M_y\}_y$. Data processing inequality implies that

$$I(X : A) \geq I(X : Y) \tag{9}$$

Intuitively, this says that by performing any measurement on A, you cannot obtain more than $I(X : A)$ bits of information about X.

3. **Joint convexity of quantum relative entropy** Let ρ_0, ρ_1 and σ_0, σ_2 be four density operator and $\lambda \in [0, 1]$. Then, we have

$$D(\lambda\rho_0 + (1-\lambda)\rho_1 \| \lambda\sigma_0 + (1-\lambda)\sigma_1) \leq \lambda D(\rho_0\|\sigma_0) + (1-\lambda)D(\rho_1\|\sigma_1) \tag{10}$$

We can prove this by applying data processing inequality to states

$$\rho = \lambda |0\rangle\langle 0| \otimes \rho_0 + (1-\lambda) |1\rangle\langle 1| \otimes \rho_1 \tag{11}$$

$$\sigma = \lambda |0\rangle\langle 0| \otimes \sigma_0 + (1-\lambda) |1\rangle\langle 1| \otimes \sigma_1 \tag{12}$$

and quantum channel being partial trace wrt to the first system.