Ph219C/CS219C

Exercises
Due: Thursday 14 April 2022

1.1 Positivity of quantum relative entropy

a) Show that $\ln x \leq x - 1$ for all positive real x, with equality iff $x = 1$.

b) The (classical) relative entropy of a probability distribution $\{p(x)\}$ relative to $\{q(x)\}$ is defined as

$$D(p \parallel q) \equiv \sum_x p(x) \left(\log p(x) - \log q(x)\right) \ . \tag{1}$$

Show that

$$D(p \parallel q) \geq 0 \ , \tag{2}$$

with equality iff the probability distributions are identical. **Hint**: Apply the inequality from (a) to $\ln\left(q(x)/p(x)\right)$.

c) The quantum relative entropy of the density operator $\boldsymbol{\rho}$ with respect to $\boldsymbol{\sigma}$ is defined as

$$D(\boldsymbol{\rho} \parallel \boldsymbol{\sigma}) = \mathrm{tr}\ \boldsymbol{\rho}\left(\log \boldsymbol{\rho} - \log \boldsymbol{\sigma}\right) \ . \tag{3}$$

Let $\{p_i\}$ denote the eigenvalues of $\boldsymbol{\rho}$ and $\{q_a\}$ denote the eigenvalues of $\boldsymbol{\sigma}$. Show that

$$D(\boldsymbol{\rho} \parallel \boldsymbol{\sigma}) = \sum_i p_i \left(\log p_i - \sum_a D_{ia} \log q_a\right) \ , \tag{4}$$

where D_{ia} is a doubly stochastic matrix. Express D_{ia} in terms of the eigenstates of $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$. (A matrix is doubly stochastic if its entries are nonnegative real numbers, where each row and each column sums to one.)

d) Show that if D_{ia} is doubly stochastic, then (for each i)

$$\log\left(\sum_a D_{ia} q_a\right) \geq \sum_a D_{ia} \log q_a \ , \tag{5}$$

with equality only if $D_{ia} = 1$ for some a.

e) Show that

$$D(\boldsymbol{\rho} \parallel \boldsymbol{\sigma}) \geq D(p \parallel r) \ , \tag{6}$$

where $r_i = \sum_a D_{ia} q_a$.

f) Show that $D(\boldsymbol{\rho} \parallel \boldsymbol{\sigma}) \geq 0$, with equality iff $\boldsymbol{\rho} = \boldsymbol{\sigma}$.

1.2 Properties of Von Neumann entropy

a) Use nonnegativity of quantum relative entropy to prove the *subadditivity* of Von Neumann entropy

$$H(\boldsymbol{\rho}_{AB}) \leq H(\boldsymbol{\rho}_A) + H(\boldsymbol{\rho}_B), \tag{7}$$

with equality iff $\boldsymbol{\rho}_{AB} = \boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B$. **Hint**: Consider the relative entropy of $\boldsymbol{\rho}_{AB}$ and $\boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B$.

b) Use subadditivity to prove the concavity of the Von Neumann entropy:

$$H(\sum_x p_x \boldsymbol{\rho}_x) \geq \sum_x p_x H(\boldsymbol{\rho}_x) \ . \tag{8}$$

Hint: Consider

$$\boldsymbol{\rho}_{AB} = \sum_x p_x \left(\boldsymbol{\rho}_x\right)_A \otimes \left(|x\rangle\langle x|\right)_B \ , \tag{9}$$

where the states $\{|x\rangle_B\}$ are mutually orthogonal.

c) Use the condition

$$H(\boldsymbol{\rho}_{AB}) = H(\boldsymbol{\rho}_A) + H(\boldsymbol{\rho}_B) \quad \text{iff} \quad \boldsymbol{\rho}_{AB} = \boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B \tag{10}$$

to show that, if all p_x's are nonzero,

$$H\left(\sum_x p_x \boldsymbol{\rho}_x\right) = \sum_x p_x H(\boldsymbol{\rho}_x) \tag{11}$$

iff all the $\boldsymbol{\rho}_x$'s are identical.

1.3 Monotonicity of quantum relative entropy

Quantum relative entropy has a property called *monotonicity*:

$$D(\boldsymbol{\rho}_A \| \boldsymbol{\sigma}_A) \leq D(\boldsymbol{\rho}_{AB} \| \boldsymbol{\sigma}_{AB}); \tag{12}$$

The relative entropy of two density operators on a system AB cannot be less than the induced relative entropy on the subsystem A.

a) Use monotonicity of quantum relative entropy to prove the strong subadditivity property of Von Neumann entropy. **Hint**: On a tripartite system ABC, consider the relative entropy of $\boldsymbol{\rho}_{ABC}$ and $\boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_{BC}$.

b) Use monotonicity of quantum relative entropy to show that the action of a quantum channel $\mathcal{N}$ cannot increase relative entropy:

$$D(\mathcal{N}(\boldsymbol{\rho})\|\mathcal{N}(\boldsymbol{\sigma}) \leq D(\boldsymbol{\rho}\|\boldsymbol{\sigma}), \tag{13}$$

Hint: Recall that any quantum channel has an isometric dilation.

1.4 Separability and majorization

The hallmark of entanglement is that in an entangled state the whole is less random than its parts. But in a separable state the correlations are essentially classical and so are expected to adhere to the classical principle that the parts are less disordered than the whole. The objective of this problem is to make this expectation precise by showing that if the bipartite (mixed) state $\boldsymbol{\rho}_{AB}$ is separable, then

$$\lambda(\boldsymbol{\rho}_{AB}) \prec \lambda(\boldsymbol{\rho}_A) \ , \quad \lambda(\boldsymbol{\rho}_{AB}) \prec \lambda(\boldsymbol{\rho}_B) \ . \tag{14}$$

Here $\lambda(\boldsymbol{\rho})$ denotes the vector of eigenvalues of $\boldsymbol{\rho}$, and $\prec$ denotes majorization.

A separable state can be realized as an ensemble of pure product states, so that if $\boldsymbol{\rho}_{AB}$ is separable, it may be expressed as

$$\boldsymbol{\rho}_{AB} = \sum_a p_a \ |\psi_a\rangle\langle\psi_a| \otimes |\varphi_a\rangle\langle\varphi_a| \ . \tag{15}$$

We can also diagonalize $\boldsymbol{\rho}_{AB}$, expressing it as

$$\boldsymbol{\rho}_{AB} = \sum_j r_j |e_j\rangle\langle e_j| \ , \tag{16}$$

where $\{|e_j\rangle\}$ denotes an orthonormal basis for AB; then by the HJW theorem, there is a unitary matrix V such that

$$\sqrt{r_j}|e_j\rangle = \sum_a V_{ja}\sqrt{p_a}|\psi_a\rangle \otimes |\varphi_a\rangle \ . \tag{17}$$

Also note that $\boldsymbol{\rho}_A$ can be diagonalized, so that

$$\boldsymbol{\rho}_A = \sum_a p_a|\psi_a\rangle\langle\psi_a| = \sum_\mu s_\mu|f_\mu\rangle\langle f_\mu| \ ; \tag{18}$$

here $\{|f_\mu\rangle\}$ denotes an orthonormal basis for A, and by the HJW theorem, there is a unitary matrix U such that

$$\sqrt{p_a}|\psi_a\rangle = \sum_\mu U_{a\mu}\sqrt{s_\mu}|f_\mu\rangle \ . \tag{19}$$

Now show that there is a doubly stochastic matrix D such that

$$r_j = \sum_\mu D_{j\mu} s_\mu \ . \tag{20}$$

That is, you must check that the entries of $D_{j\mu}$ are real and nonnegative, and that $\sum_j D_{j\mu} = 1 = \sum_\mu D_{j\mu}$. Thus we conclude that $\lambda(\boldsymbol{\rho}_{AB}) \prec \lambda(\boldsymbol{\rho}_A)$. Just by interchanging A and B, the same argument also shows that $\lambda(\boldsymbol{\rho}_{AB}) \prec \lambda(\boldsymbol{\rho}_B)$.

Remark: Note that it follows from the Schur concavity of Shannon entropy that, if $\boldsymbol{\rho}_{AB}$ is separable, then the von Neumann entropy has the properties $H(AB) \geq H(A)$ and $H(AB) \geq H(B)$. Thus, for separable states, conditional entropy is nonnegative: $H(A|B) = H(AB) - H(B) \geq 0$ and $H(B|A) = H(AB) - H(A) \geq 0$. In contrast, if the state of AB is an entangled pure state, then $H(AB) = 0$ and $H(B|A) = H(A|B) < 0$.

1.5 The first law of Von Neumann entropy

We'll use $S(\boldsymbol{\rho}) = -\mathrm{tr}\,(\boldsymbol{\rho} \ln \boldsymbol{\rho})$ to denote the entropy of a density operator when using natural logarithms instead of logarithms with base 2. As in §10.2.6, a $d \times d$ density matrix can be expressed as

$$\boldsymbol{\rho} = \frac{e^{-\boldsymbol{K}}}{\mathrm{tr}\left(e^{-\boldsymbol{K}}\right)}, \tag{21}$$

where $\boldsymbol{K}$ is a $d \times d$ Hermitian matrix called the *modular Hamiltonian* associated with $\boldsymbol{\rho}$. (Under this definition of $\boldsymbol{K}$, we have the freedom to shift $\boldsymbol{K}$ by a multiple of the identity operator without changing $\boldsymbol{\rho}$.) We assume that $\boldsymbol{\rho}$ has full rank; that is, it has d positive eigenvalues. We will see that when $\boldsymbol{\rho}$ changes slightly, the first-order change in $S(\boldsymbol{\rho})$ can be related to the change in the expectation value of $\boldsymbol{K}$.

a) Suppose $\boldsymbol{A}(\lambda)$ is a bounded Hermitian operator smoothly parametrized by the real number λ. Show that

$$\frac{d}{d\lambda}\left(\mathrm{tr}\boldsymbol{A}^n\right) = n\ \mathrm{tr}\left(\frac{d\boldsymbol{A}}{d\lambda}\boldsymbol{A}^{n-1}\right) . \tag{22}$$

Do not assume that $d\boldsymbol{A}/d\lambda$ commutes with $\boldsymbol{A}$.

b) Suppose the density operator is perturbed slightly:

$$\boldsymbol{\rho} \to \boldsymbol{\rho}' = \boldsymbol{\rho} + \delta\boldsymbol{\rho}. \tag{23}$$

Since $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$ are both normalized density operators, we have $\operatorname{tr}(\delta\boldsymbol{\rho}) = 0$. Show that

$$S(\boldsymbol{\rho}') - S(\boldsymbol{\rho}) = \operatorname{tr}\left(\boldsymbol{\rho}'\boldsymbol{K}\right) - \operatorname{tr}\left(\boldsymbol{\rho}\boldsymbol{K}\right) + O\left((\delta\boldsymbol{\rho})^2\right); \tag{24}$$

that is,

$$\delta S = \delta\langle\boldsymbol{K}\rangle \tag{25}$$

to first order in the small change in $\boldsymbol{\rho}$. This statement generalizes the first law of thermodynamics; for the case of a thermal density operator with $\boldsymbol{K} = \boldsymbol{H}/T$ (where $\boldsymbol{H}$ is the Hamiltonian and T is the temperature), it becomes the more familiar statement

$$\delta E = \delta\langle\boldsymbol{H}\rangle = T\delta S. \tag{26}$$