# Quantum Computing (Fall 2013) Instructor: Shengyu Zhang.

# Lecture 10 Quantum information 2: entropy and mutual information

## Review of classical information theory

Let’s first review some basic concepts in information theory.

Suppose that is a discrete random variable on the sample space X, and the probability distribution function is . Following the standard notation in information theory, we also use the capital letter to denote the distribution, and by writing , we mean to draw a sample from the distribution .

**Entropy**. The entropy of a discrete random variable is defined by

Here we use the convention that .

One intuitive explanation about this definition is that entropy measures the amount of the uncertainty of a random variable. The more entropy has, the more uncertain it is. In particular, iff is fixed/deterministic. At the other extreme, the maximum entropy for any random variable on X is , achieved by the uniform distribution. The range of the entropy is

Another interpretation of entropy is the compression rate, which will be the subject of the next lecture.

A special case is binary entropy, when takes only 2 possible values. Suppose , then the binary entropy is . How does this quantity vary with ? See the following figure.



In particular, it reaches its maximum when .

**Joint distribution**. Sometimes we have more than one random variable under the concern. If is a pair of random variables on , distributed according to a joint distribution , then the total entropy is simply the following, treating as one random variable in a bigger sample space :

**Marginal distribution and conditional distribution**. The joint variable has a *marginal distribution* over defined by . For any fixed with positive probability for some , there is a *conditional distribution* defined by . The *conditional entropy* is defined as the average entropy of the conditional distribution:

From this definition, it is easily seen that .

In general, conditioning reduces entropy:

Knowing something else () always helps you to knowing the target (*X*).

**Chain rule**. Another basic fact is the *chain rule*:

Think of this as the following: The uncertainty of is that of one variable , plus that of the other variable after knowing .

The chain rule also works with conditional entropy,

and with more variables:

where the inequality is usually referred to as the subadditivity of entropy.

**Relative entropy**. Another important concept is that of the relative entropy of two distributions and .

Relative entropy can sometimes serve as a good measure as distance of two measures. A basic property is that it is nonnegative.

**Fact**. with equality iff .

Proof. Use the fact that in the definition of .

**Mutual information**. For a joint distribution , the *mutual information* between and is

It is a good exercise to verify the above equalities. But the second one has a clear explanation: We mentioned earlier that conditioning always reduces the entropy. How much does it reduce? Exactly the mutual information. So the mutual information is the amount of uncertainty of minus that when is known. In this way, it measures how much ***contains*** the information of . It is not hard to verify that it’s symmetric: , and we’ve implicitly said that it’s nonnegative:

, i.e.

One can also add conditioning on this. The *conditional mutual information* is defined by

It is not hard to see that

and

**Data processing inequality.** A sequence of random variables form a *Markov Chain*, denoted by if

Namely, depends only on , but not the “earlier history”.

**Fact**. If , then .

Intuitively, it says that processing data always reduces the contained information.

## Quantum entropy

We can extend the classical Shannon entropy to the quantum case. The *von Newmann entropy* of a quantum state is defined by

Equivalently, if the eigenvalues of are , then . (Note that and , so is a distribution.)

Similar to the classical case, the quantum entropy is a measure of uncertainty of the quantum states.

**Fact**. , with equality iff is a pure state.

**Relative entropy**. The quantum relative entropy is defined by

Let’s first verify that this extends the classical relative entropy. When and commute and they have eigenvalues and , then

consistent with the classical case.

The relative entropy can be viewed as a distance measure between states.

**Theorem** (Klein’s inequality). with equality holding iff .

Another “distance” property is that taking partial trace reduces relative entropy.

**Fact**. .

The theorem can be used to prove the following result.

**Fact**. Projective measurements can only increase entropy. More precisely, suppose is a

complete set of orthogonal projectors and is a density operator. Then with equality iff .

**Conditional entropy**. Suppose that a joint quantum system is in state . For convenience, we sometimes use to denote . We define the conditional entropy as

Note that classically conditional entropy can also be defined as the average of , but we cannot do it for quantum entropy. Actually, different than the classical case, the quantum conditional entropy *can be negative*. For example, consider an EPR pair, , but the average definition won't give us this: A and B are in either 00 or 11.

On the other hand, the quantum conditional entropy cannot be *too* negative.

**Convexity**. Suppose ’s are quantum states and ’s form a distribution.

**Mutual information**.

Classically the mutual information is upper bounded by individual entropy, i.e. , but quantum mutual information isn’t… Well it still holds---up to a factor of 2:

**Strong subadditivity**. .

## Notes

The definitions and theorems in this lecture about classical information theory can be found in the standard textbook such as [CT06] (Chapter 1). The quantum part can be found in [NC00] (Chapter 11).

## References

[CT06] Thomas Cover, Joy Thomas. **Elements of Information Theory**, Second Edition, *Wiley InterScience*, 2006.

[NC00] Michael Nielsen and Isaac Chuang, **Quantum Computation and Quantum Information**, *Cambridge University Press*, 2000.

## Exercise

1. What’s the von Newmann entropy of the following quantum states?
* .
1. What’s the mutual information between the two parts of an EPR state?