Entropy and Quantum Information Theory

고려대학교 정보보호대학원

양 형 진

Quantum Information Theory

- Quantum theory
- Relativity
- Information theory

Entropy

- A basic & key concept of
  Classical & Quantum information theory

- Shannon & Von Neumann

- The 2nd law of thermodynamics

Shannon entropy

- Probability distribution $p_1, \ldots, p_N$

$$H(X) \equiv H(p_1, \ldots, p_N)$$
$$= -\sum_x p_x \lg p_x$$

- $H(X)$ quantifies the resources needed to store information

Entropy & Compression

Ex) Symbols : 1 2 3 4

Ex)	Symbols :	1	2	3	4
	Probabilities :	1/2	1/4	1/8	1/8
	Compression :	0	10	110	111

- Av length of the compression string

= 1/2 + 2/4 + 3/8 + 3/8

= 7/4

= $H(X)$

- $H(X)$ quantifies the optimal compression

  ➡ Noiseless coding theorem

Noiseless Coding Theorem 1

- Shannon's noiseless channel coding theorem

Suppose $\{X_i\}$ is an I.I.d. information source with entropy rate $H(X)$.

If $R > H(X)$, then there exists a reliable compression scheme of rate R for the source.

If $R < H(X)$, then any compression scheme will not be reliable.

* I.I.d. : independent &identically distributed

Noiseless Coding Theorem 2

- Schumacher's quantum noiseless channel coding theorem

Let $\{H, \rho\}$ be an I.I.d. quantum source.

If $R > S(\rho)$ then there exists a reliable compression scheme of rate R for the source.

If $R < S(\rho)$ then any compression scheme of rate R is not reliable.

Typical sequence

- Theorem of typical sequence
  1. Fix $\varepsilon > 0$. There for any $\delta > 0$, for sufficiently large n, the probability that a sequence is ε-typical is at least $1-\delta$
  2. For any fixed $\varepsilon > 0$ and $\delta > 0$, for sufficiently large n, the number $|T(n,\varepsilon)|$ of ε-typical sequence satisfies
  $$(1-\delta)2^{n(H(X)-\varepsilon)} \leq |T(n,\varepsilon)| \leq 2^{n(H(X)+\varepsilon)}$$
  3. Let $S(n)$ be a collection of size at most 2^{nR}, of length n sequences from the source, where $R < H(X)$ is fixed. Then for any $\delta > 0$ and for sufficiently large n,
  $$\sum_{x \in S(n)} p(x) \leq \delta$$

* Given $\varepsilon > 0$ a string of source symbols $x_1 x_2 \ldots x_n$ is ε-typical if
$$2^{-n(H(X)+\varepsilon)} \leq p(x_1, \ldots, x_n) \leq 2^{-n(H(X)-\varepsilon)}$$

Typical subspace

- Typical subspace theorem
  1. Fix $\varepsilon > 0$. Then for any $\delta > 0$, for sufficiently large n
     $$tr(P(n,\varepsilon)\rho^{\otimes n}) \geq 1-\delta$$
  2. For any fixed $\varepsilon > 0$ and $\delta > 0$, for sufficiently large n the dimension $|T(n,\varepsilon)| = tr(P(n,\varepsilon))$ of $T(n,\varepsilon)$ satisfies
     $$(1-\delta)2^{n(S(\rho)-\varepsilon)} \leq |T(n,\varepsilon)| \leq 2^{n(S(\rho)+\varepsilon)}$$
  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,
     $$tr(S(\rho)\rho^{\otimes n}) \leq \delta$$

Relative entropy

- $$H(p(x)\|q(x)) \equiv \sum_x p(x)\lg(p(x)/q(x))$$
  $$\equiv -H(X) - \sum_x p(x)\lg(q(x))$$
  $$S(\rho\|\sigma) \equiv tr(\rho\lg\rho) - tr(\rho\lg\sigma)$$
- Non-negativity of the relative entropy
  $$H(p(x)\|q(x)) \geq 0$$
  $S(\rho\|\sigma) \geq 0$: Klein's inequality

Joint entropy & Conditional entropy

- Joint entropy

$$H(X,Y) \equiv -\sum_{x,y} p(x,y) \lg p(x,y)$$

$$S(A,B) \equiv -tr(\rho^{AB} \lg \rho^{AB})$$

- Conditional entropy

$$H(X|Y) \equiv H(X,Y) - H(Y)$$

$$S(A|B) \equiv S(A,B) - S(B)$$

Mutual information

- $H(X:Y) \equiv H(X) + H(Y) - H(X,Y)$

$$= H(X) - H(X|Y)$$

- $S(A:B) \equiv S(A) + S(B) - S(A,B)$

$$= S(A) - S(A|B)$$

$$= S(B) - S(B|A)$$

Properties of Shannon entropy 1

1. $H(X,Y) = H(Y,X)$
2. $H(Y|X) \geq 0 \quad \rightarrow \quad H(X,Y) \geq H(X)$

   $H(Y) \geq H(X:Y)$
3. $H(X,Y) \leq H(X) + H(Y)$: subadditivity
4. $H(Y|X) \leq H(Y) \quad \rightarrow \quad H(X:Y) \geq 0$
5. $H(X,Y,Z) + H(Y) \leq H(X,Y) + H(Y,Z)$

   : Strong subadditivity

Properties of Shannon entropy 2

6. Conditioning reduces entropy

$$H(X\,|\,Y,Z) \le H(X\,|\,Y)$$

7. Chaining rule for conditional entropies

$$H(X_1,\ldots,X_N\,|\,Y) = \sum_{i=1}^{N} H(X_i\,|\,Y, X_1,\ldots,X_{i-1})$$

8. Data processing inequality

$$H(X) \ge H(X:Y) \ge H(X:Z)$$: Markov chain

9. Data pipelining inequality

$$H(Z:Y) \ge H(Z:X)$$: Markov chain

Von Neumann entropy 1

1. S is non-negative
2. $S \le \lg d$ for a d-dimensional H-space
3. $S(A) = S(B)$ when a composite system AB is in a pure state
4. $S(\sum_i p_i \rho_i) = H(p_i) + \sum_i p_i S(\rho_i)$
5. Joint entropy theorem

$$S(\sum_i p_i |i\rangle\langle i| \rho_i) = H(p_i) + \sum_i p_i S(\rho_i)$$

Von Neumann entropy 2

6. Concavity of entropy

$$S(\sum_i p_i \rho_i) \geq \sum_i p_i S(\rho_i)$$ (by 4)

7. Projective measurements increase entropy

$$S(\rho') \geq S(\rho)$$ with $$\rho' = \sum_i P_i \rho P_i$$

8. Subadditivity

$$S(A,B) \leq S(A) + S(B)$$

9. Triangle inequality

$$\left| S(A) - S(B) \right| \leq S(A,B)$$

Von Neumann entropy3

10. Strong subadditivity

$$S(A)+S(B)\le S(A,C)+S(B,C)$$

$$S(A,B,C)+S(B)\le S(A,B)+S(B,C)$$

11. Conditioning reduces entropy

$$S(A|B,C)\le S(A|B)$$

12. Discarding Q sys never increase m-inf

$$S(A:B)\le S(A:B,C)$$

13. Q operations never increase m-inf

$$S(A':B')\le S(A:B)$$

Holevo bound

- A prepares a state ρ_X where $X = 0, \ldots, N$

  with prob $p_0, \ldots, p_N$

  $$\rho = \sum_x p_x \rho_x$$

- B performs a POVM $\{E_y\} = \{E_1, \ldots, E_N\}$

  with measurement outcome Y

- $H(X:Y) < S(\rho) - \sum_x p_x S(\rho_x)$

  Holevo bound

- Upper bound on the accessible information