Information and Coding Theory

MA41024/ MA60020/ MA60262

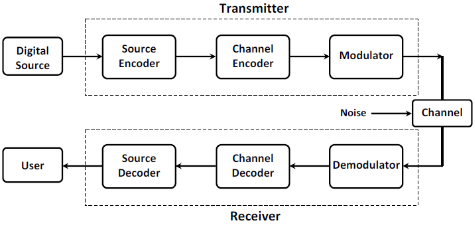
Bibhas Adhikari

Spring 2022-23, IIT Kharagpur

Lecture 2
January 9, 2023

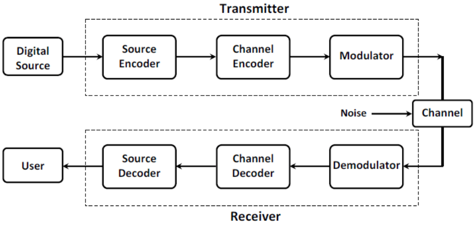
Content of the course

Model of a digital communication system



Content of the course

Model of a digital communication system



The encoding paradigm: Here

Entropy

Random variables Let Ω be a finite set, and $\mu : \Omega \to [0, 1]$ a function with

$$\sum_{\omega \in \Omega} \mu(\omega) = 1.$$

Then we refer Ω as the sample space and μ as a probability distribution (probability measure) on Ω.

Entropy

Random variables Let Ω be a finite set, and $\mu : \Omega \to [0, 1]$ a function with

$$\sum_{\omega \in \Omega} \mu(\omega) = 1.$$

Then we refer Ω as the sample space and μ as a probability distribution (probability measure) on Ω.

A random variable over Ω is any function $X : \Omega \to \mathbb{R}$. Then the expectation of X is

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} \mu(\omega) \cdot X(\omega)$$

Entropy

Random variables Let Ω be a finite set, and $\mu : \Omega \to [0, 1]$ a function with

$$\sum_{\omega \in \Omega} \mu(\omega) = 1.$$

Then we refer Ω as the sample space and μ as a probability distribution (probability measure) on Ω.
A random variable over Ω is any function $X : \Omega \to \mathbb{R}$. Then the expectation of X is

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} \mu(\omega) \cdot X(\omega)$$

If $\mathcal{X}$ is the range set of X then we can think of the probability distribution on $\mathcal{X}$ as

$$p(x) = \mathbb{P}[X = x] = \sum_{\omega : X(\omega) = x} \mu(\omega)$$

for any $x \in \mathcal{X}$

Entropy

Convex set and convex function

A set $S \subset \mathbb{R}^n$ is called convex if for any $\alpha \in [0, 1]$

$$\alpha \cdot x + (1 - \alpha) \cdot y \in S$$

for any $x, y \in S$.

Entropy

Convex set and convex function

A set $S \subset \mathbb{R}^n$ is called convex if for any $\alpha \in [0, 1]$

$$\alpha \cdot x + (1 - \alpha) \cdot y \in S$$

for any $x, y \in S$.
Let S be a convex subset of $\mathbb{R}^n$. Then a function $f : S \to \mathbb{R}$ is said to be a convex function on S if for all $\alpha \in [0, 1]$

$$f(\alpha \cdot x + (1 - \alpha) \cdot y) \leq \alpha \cdot f(x) + (1 - \alpha) \cdot f(y)$$

for all $x, y \in S$.

Entropy

Convex set and convex function

A set $S \subset \mathbb{R}^n$ is called convex if for any $\alpha \in [0, 1]$

$$\alpha \cdot x + (1 - \alpha) \cdot y \in S$$

for any $x, y \in S$.

Let S be a convex subset of $\mathbb{R}^n$. Then a function $f : S \to \mathbb{R}$ is said to be a convex function on S if for all $\alpha \in [0, 1]$

$$f(\alpha \cdot x + (1 - \alpha) \cdot y) \leq \alpha \cdot f(x) + (1 - \alpha) \cdot f(y)$$

for all $x, y \in S$. Equivalently, f is convex if the set

$$S_f = \{(x, z) : z \geq f(x)\}$$

is a convex subset of $\mathbb{R}^{n+1}$.

Entropy

Convex set and convex function

A set $S \subset \mathbb{R}^n$ is called convex if for any $\alpha \in [0, 1]$

$$\alpha \cdot x + (1 - \alpha) \cdot y \in S$$

for any $x, y \in S$.

Let S be a convex subset of $\mathbb{R}^n$. Then a function $f : S \to \mathbb{R}$ is said to be a convex function on S if for all $\alpha \in [0, 1]$

$$f(\alpha \cdot x + (1 - \alpha) \cdot y) \leq \alpha \cdot f(x) + (1 - \alpha) \cdot f(y)$$

for all $x, y \in S$. Equivalently, f is convex if the set

$$S_f = \{(x, z) : z \geq f(x)\}$$

is a convex subset of $\mathbb{R}^{n+1}$. If the inequality is strict then f is called strictly convex.

Entropy

If f satisfies the opposite inequality for all x, y, α then f is called concave function.

Entropy

If f satisfies the opposite inequality for all x, y, α then f is called concave function.
If f is a convex function then $-f$ is a concave function

Entropy

If f satisfies the opposite inequality for all x, y, α then f is called concave function.
If f is a convex function then $-f$ is a concave function

Note A function $f : \mathbb{R} \to \mathbb{R}$ is a twice differentiable function then f is convex if and only if $f"(x) \geq 0$ for all $x \in S$

Entropy

If f satisfies the opposite inequality for all x, y, α then f is called concave function.
If f is a convex function then $-f$ is a concave function

Note A function $f : \mathbb{R} \to \mathbb{R}$ is a twice differentiable function then f is convex if and only if $f''(x) \geq 0$ for all $x \in S$

Homework $f(x) = x^2$ is a convex function on $\mathbb{R}$.
$f(x) = \log(x), f(x) = x\log(x)$ are concave and convex functions respectively, on $(0, \infty)$

Entropy

Observation Consider a rv X which takes the values x with probability α and y with probability $1-\alpha$. Let $f(X)$ be a function. Then what happens to the convexity condition?

Jensen's inequality Let $S \subseteq \mathbb{R}^n$ be a convex set and let X be a random variable whose range set is a subset of S. Then for a convex function $f : S \to \mathbb{R}$,

$$\mathbb{E}[f(X)] \geq f(\mathbb{E}[X]).$$

Equivalently, for a concave function $f : S \to \mathbb{R}$,

$$\mathbb{E}[f(X)] \leq f(\mathbb{E}[X])$$

Entropy

Observation Consider a rv X which takes the values x with probability α and y with probability $1-\alpha$. Let $f(X)$ be a function. Then what happens to the convexity condition?

Jensen's inequality Let $S \subseteq \mathbb{R}^n$ be a convex set and let X be a random variable whose range set is a subset of S. Then for a convex function $f : S \to \mathbb{R}$,

$$\mathbb{E}[f(X)] \geq f(\mathbb{E}[X]).$$

Equivalently, for a concave function $f : S \to \mathbb{R}$,

$$\mathbb{E}[f(X)] \leq f(\mathbb{E}[X])$$

Homework Prove Cauchy-Schwarz inequality using Jensen's inequality.

Entropy

Shannon's assumptions - Let $\mathcal{X} = \{a_1, \ldots, a_n\}$ with pmf $p(a_i), 1 \leq i \leq n$. Denote the entropy by H

Entropy

Shannon's assumptions - Let $\mathcal{X} = \{a_1, \ldots, a_n\}$ with pmf $p(a_i), 1 \leq i \leq n$. Denote the entropy by H

- H must be a continuous function of p_is

Entropy

Shannon's assumptions - Let $\mathcal{X} = \{a_1, \ldots, a_n\}$ with pmf $p(a_i), 1 \leq i \leq n$. Denote the entropy by H

- H must be a continuous function of p_is
- H must be an increasing function of n when $p(a_i) = 1/n, 1 \leq i \leq n$

Entropy

Shannon's assumptions - Let $\mathcal{X} = \{a_1, \ldots, a_n\}$ with pmf $p(a_i), 1 \leq i \leq n$. Denote the entropy by H

- H must be a continuous function of p_is
- H must be an increasing function of n when $p(a_i) = 1/n, 1 \leq i \leq n$
- Bundling property

Entropy

Shannon's assumptions - Let $\mathcal{X} = \{a_1, \ldots, a_n\}$ with pmf $p(a_i), 1 \leq i \leq n$. Denote the entropy by H

- H must be a continuous function of p_is
- H must be an increasing function of n when $p(a_i) = 1/n, 1 \leq i \leq n$
- Bundling property

Question How much information is revealed when we know outcome of a random experiment?

Entropy

Shannon's assumptions - Let $\mathcal{X} = \{a_1, \ldots, a_n\}$ with pmf $p(a_i), 1 \leq i \leq n$. Denote the entropy by H

- H must be a continuous function of p_is
- H must be an increasing function of n when $p(a_i) = 1/n, 1 \leq i \leq n$
- Bundling property

Question How much information is revealed when we know outcome of a random experiment? How surprised are we?

Entropy

Entropy Suppose X is a rv distributed over $\mathcal{X} = \{a_1, \ldots, a_n\}$ such that each value $x \in \mathcal{X}$ occurs with probability $p(x)$. Then the entropy of X is

$$H(X) = \sum_{x \in \mathcal{X}} p(x) \cdot \underbrace{\log_2 \left(\frac{1}{p(x)} \right)}_{\text{surprise}}$$

Entropy

Entropy Suppose X is a rv distributed over $\mathcal{X} = \{a_1, \ldots, a_n\}$ such that each value $x \in \mathcal{X}$ occurs with probability $p(x)$. Then the entropy of X is

$$H(X) = \sum_{x \in \mathcal{X}} p(x) \cdot \underbrace{\log_2 \left(\frac{1}{p(x)} \right)}_{\text{surprise}} = - \sum_{x \in \mathcal{X}} p(x) \cdot \log_2(p(x))$$

Entropy

Entropy Suppose X is a rv distributed over $\mathcal{X} = \{a_1, \ldots, a_n\}$ such that each value $x \in \mathcal{X}$ occurs with probability $p(x)$. Then the entropy of X is

$$H(X) = \sum_{x \in \mathcal{X}} p(x) \cdot \underbrace{\log_2 \left(\frac{1}{p(x)} \right)}_{\text{surprise}} = -\sum_{x \in \mathcal{X}} p(x) \cdot \log_2(p(x))$$

Example Consider the entropy of coin toss with p as probability of head.

Entropy

Entropy Suppose X is a rv distributed over $\mathcal{X} = \{a_1, \ldots, a_n\}$ such that each value $x \in \mathcal{X}$ occurs with probability $p(x)$. Then the entropy of X is

$$H(X) = \sum_{x \in \mathcal{X}} p(x) \cdot \underbrace{\log_2 \left(\frac{1}{p(x)} \right)}_{\text{surprise}} = - \sum_{x \in \mathcal{X}} p(x) \cdot \log_2 (p(x))$$

Example Consider the entropy of coin toss with p as probability of head. What happens if $p_i = 1/n$?

Entropy

Entropy Suppose X is a rv distributed over $\mathcal{X} = \{a_1, \ldots, a_n\}$ such that each value $x \in \mathcal{X}$ occurs with probability $p(x)$. Then the entropy of X is

$$H(X) = \sum_{x \in \mathcal{X}} p(x) \cdot \underbrace{\log_2 \left(\frac{1}{p(x)} \right)}_{\text{surprise}} = -\sum_{x \in \mathcal{X}} p(x) \cdot \log_2 (p(x))$$

Example Consider the entropy of coin toss with p as probability of head. What happens if $p_i = 1/n$?

Proposition $0 \leq H(X) \leq \log(|\mathcal{X}|)$

Entropy

Entropy Suppose X is a rv distributed over $\mathcal{X} = \{a_1, \ldots, a_n\}$ such that each value $x \in \mathcal{X}$ occurs with probability $p(x)$. Then the entropy of X is

$$H(X) = \sum_{x \in \mathcal{X}} p(x) \cdot \underbrace{\log_2\left(\frac{1}{p(x)}\right)}_{\text{surprise}} = -\sum_{x \in \mathcal{X}} p(x) \cdot \log_2(p(x))$$

Example Consider the entropy of coin toss with p as probability of head. What happens if $p_i = 1/n$?

Proposition $0 \leq H(X) \leq \log(|\mathcal{X}|)$

Proof Let Y be a rv which takes the value $1/p(x)$ with probability $p(x)$. Then

$$\sum_{x \in \mathcal{X}} p(x) \cdot \log\left(\frac{1}{p(x)}\right) = \mathbb{E}[\log(Y)] \leq \log(\mathbb{E}[Y])$$

Entropy

Entropy Suppose X is a rv distributed over $\mathcal{X} = \{a_1, \ldots, a_n\}$ such that each value $x \in \mathcal{X}$ occurs with probability $p(x)$. Then the entropy of X is

$$H(X) = \sum_{x \in \mathcal{X}} p(x) \cdot \underbrace{\log_2 \left(\frac{1}{p(x)} \right)}_{\text{surprise}} = - \sum_{x \in \mathcal{X}} p(x) \cdot \log_2(p(x))$$

Example Consider the entropy of coin toss with p as probability of head. What happens if $p_i = 1/n$?

Proposition $0 \leq H(X) \leq \log(|\mathcal{X}|)$

Proof Let Y be a rv which takes the value $1/p(x)$ with probability $p(x)$. Then

$$\sum_{x \in \mathcal{X}} p(x) \cdot \log \left(\frac{1}{p(x)} \right) = \mathbb{E}[\log(Y)] \leq \log(\mathbb{E}[Y]) = \log \left(\sum_{x \in \mathcal{X}} p(x) \cdot \frac{1}{p(x)} \right)$$