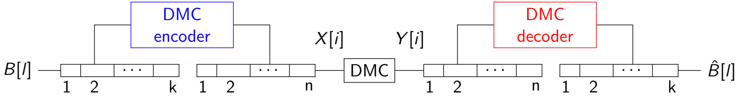
Noisy-channel coding theorem and differential entropy

Communication Theory

Manuel A. Vázquez

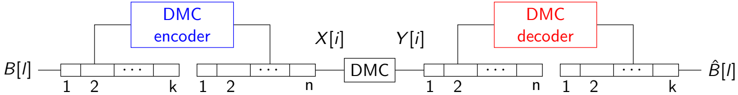
April 18, 2024

Noisy-channel coding theorem



Rate: $R = \frac{k}{n}$ | Capacity: $C = \max_{p(x_i), i=1,\cdots,M} I(X, Y)$

Noisy-channel coding theorem



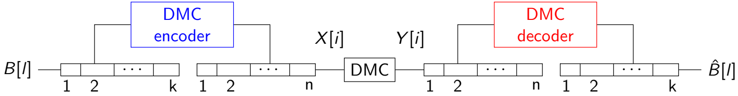
Rate: $R = \frac{k}{n}$ | Capacity: $C = \max_{p(x_i),i=1,\cdots,M} I(X, Y)$

Theorem: Noisy-channel coding (Shannon, 1948)

1. $mR < C \Rightarrow \forall\delta > 0, \exists$ code yielding $P_e < \delta$
2. $mR > C \Rightarrow P_e > \epsilon$, where $\epsilon > 0$ is a constant.

$m = \log_2 M \equiv$ number of bits per symbol

Noisy-channel coding theorem



Rate: $R = \frac{k}{n}$ | Capacity: $C = \max_{p(x_i), i=1,\cdots,M} I(X,Y)$

Theorem: Noisy-channel coding (Shannon, 1948)

1. $mR < C \Rightarrow \forall \delta > 0, \exists$ code yielding $P_e < \delta$
2. $mR > C \Rightarrow P_e > \epsilon$, where $\epsilon > 0$ is a constant.

$m = \log_2 M \equiv$ number of bits per symbol

There *exist* codes attaining the channel capacity, and

- low R: easy to find one
- high R: hard to find one

Differential entropy

Definition: Differential entropy

$$h(X) = \int_{-\infty}^{\infty} f_X(x) \log_2 \frac{1}{f_X(x)} dx \text{ bits}$$

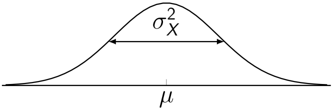
Differential entropy

Definition: Differential entropy

$$h(X) = \int_{-\infty}^{\infty} f_X(x) \log_2 \frac{1}{f_X(x)} dx \text{ bits}$$

Gaussian random variable

$$X \sim \mathcal{N}\left(\mu, \sigma_X^2\right)$$



$$h(X) = \frac{1}{2} \log_2 \left(2\pi e \sigma_X^2\right) \text{ bits}$$

(**irregardless of the mean**!!)

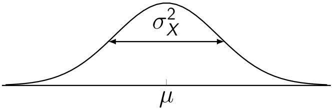
Differential entropy

Definition: Differential entropy

$$h(X) = \int_{-\infty}^{\infty} f_X(x) \log_2 \frac{1}{f_X(x)} dx \text{ bits}$$

<u>Gaussian random variable</u>

$$X \sim \mathcal{N}\left(\mu, \sigma_X^2\right)$$



$$h(X) = \frac{1}{2}\log_2\left(2\pi e \sigma_X^2\right) \text{ bits}$$

(irregardless of the mean!!)

<u>Uniform random variable</u>

$$X \sim \mathcal{U}[a, b]$$

$$h(X) = \log_2(b - a) \text{ bits}$$

Bounds on the differential entropy

- for X unbounded, i.e. $X \in (-\infty, \infty)$, with variance σ_X^2,

$$h(X) \text{ maximum} \Leftrightarrow X \sim \mathcal{N}\left(\cdot, \sigma_X^2\right),$$

Bounds on the differential entropy

- for X unbounded, i.e. $X \in (-\infty, \infty)$, with variance σ_X^2,

$$h(X) \text{ maximum} \Leftrightarrow X \sim \mathcal{N}\left(\cdot, \sigma_X^2\right),$$

and hence, for X unbounded,

$$h(X) \leq \frac{1}{2} \log_2 \left(2\pi e \sigma_X^2 \text{bits}\right) \text{ bits}$$

Bounds on the differential entropy

- for X unbounded, i.e. $X \in (-\infty, \infty)$, with variance σ_X^2,

$$h(X) \text{ maximum} \Leftrightarrow X \sim \mathcal{N}\left(\cdot, \sigma_X^2\right),$$

and hence, for X unbounded,

$$h(X) \leq \frac{1}{2} \log_2 \left(2\pi e \sigma_X^2 \text{bits}\right) \text{ bits}$$

- for X bounded between a and b

$$h(X) \text{ maximum} \Leftrightarrow X \sim \mathcal{U}\left[a, b\right],$$

Bounds on the differential entropy

- for X unbounded, i.e. $X \in (-\infty, \infty)$, with variance σ_X^2,

$$h(X) \text{ maximum} \Leftrightarrow X \sim \mathcal{N}\left(\cdot, \sigma_X^2\right),$$

and hence, for X unbounded,

$$h(X) \leq \frac{1}{2}\log_2\left(2\pi e\sigma_X^2 \text{bits}\right) \text{ bits}$$

- for X bounded between a and b

$$h(X) \text{ maximum} \Leftrightarrow X \sim \mathcal{U}\left[a, b\right],$$

and hence for X bounded,

$$h(X) \leq \log_2(b-a) \text{ bits}$$

Definition: Joint differential entropy

$$h(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) \log_2 \frac{1}{f_{X,Y}(x, y)} dxdy$$

Definition: Joint differential entropy

$$h(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) \log_2 \frac{1}{f_{X,Y}(x, y)} dxdy$$

Definition: Conditional differential entropy

$$h(X|Y) = \int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{\infty} f_{X|Y}(x|y) \log_2 \frac{1}{f_{X|Y}(x|y)} dxdy$$

or, equivalently,

$$h(X|Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) \log_2 \frac{1}{f_{X|Y}(x|y)} dxdy$$

Definition: Joint differential entropy

$$h(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\log_2 \frac{1}{f_{X,Y}(x,y)}dxdy$$

Definition: Conditional differential entropy

$$h(X|Y) = \int_{-\infty}^{\infty} f_Y(y)\int_{-\infty}^{\infty} f_{X|Y}(x|y)\log_2 \frac{1}{f_{X|Y}(x|y)}dxdy$$

or, equivalently,

$$h(X|Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\log_2 \frac{1}{f_{X|Y}(x|y)}dxdy$$

Definition: Mutual information

$$I(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\log_2 \frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)}dxdy$$

Mutual information and conditional entropy

Properties

- $I(X, Y) \geq 0$ (non negative function)
- $I(X, Y) = 0 \Leftrightarrow X$ and Y independent
- $I(X, Y) = I(Y, X)$

Mutual information and conditional entropy

Properties

- $I(X, Y) \geq 0$ (non negative function)
- $I(X, Y) = 0 \Leftrightarrow X$ and Y independent
- $I(X, Y) = I(Y, X)$

Identities

(counterparts of those for the discrete case)

- mutual information

$$\begin{aligned} I(X, Y) &= h(Y) - h(Y|X) \\ &= h(X) - h(X|Y) \end{aligned}$$

Mutual information and conditional entropy

Properties

- $I(X, Y) \geq 0$ (non negative function)
- $I(X, Y) = 0 \Leftrightarrow X$ and Y independent
- $I(X, Y) = I(Y, X)$

Identities

(counterparts of those for the discrete case)

- mutual information

$$\begin{aligned} I(X, Y) &= h(Y) - h(Y|X) \\ &= h(X) - h(X|Y) \end{aligned}$$

- joint entropy

$$\begin{aligned} h(X, Y) &= h(X|Y) + h(Y) \\ &= h(Y|X) + h(X) \end{aligned}$$