Differential entropy, Gaussian channels

Differential entropy

- Differential entropy

The differential entropy $h(X)$ of a continuous random variable X with probability density function $f(x)$, where $\int_{-\infty}^{+\infty} f(x)\,dx = 1$ is:

$$h(X) = -\int_{-\infty}^{+\infty} f(x) \log f(x)\,dx$$

Examples of differential entropy

Example: Let X be uniform on $[a, b]$.

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{if } x < a, \text{ or } x > b \end{cases}$$

Then:

$$h(X) = -\int_a^b \frac{1}{b-a} \log \frac{1}{b-a} dx = -\log \frac{1}{b-a} = \log(b-a)$$

- Note that the differential entropy depends only on the size of the interval, not its location. In other words, differential entropy is translation invariant.
- The differential entropy of a random variable U that is uniformly distributed on $I = [0, 1]$ is $h(U) = \log 1 = 0$. If $(b-a) < 1$, then $h(U) < 0$!!!

What does it mean?

Entropy of Gaussian random variable

Let $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ be a Gaussian random variable with mean μ and variance σ^2. The density of X is:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

We will use the translation invariance property of differential entropy to simplify computation of $h(X)$ by setting $\mu = 0$. let $\phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$

Entropy of Gaussian random variable

$$
\begin{aligned}
&h(X) = -\int_{-\infty}^{+\infty} \phi(x) \ln \phi(x)\, dx = \\
&-\int_{-\infty}^{+\infty} \phi(x) \left(-\frac{x^2}{2\sigma^2} - \ln\sqrt{2\pi\sigma^2}\right) dx = \\
&\int_{-\infty}^{+\infty} \phi(x) \frac{x^2}{2\sigma^2} dx + \int_{-\infty}^{+\infty} \phi(x) \ln\sqrt{2\pi\sigma^2} dx = \\
&\frac{1}{2\sigma^2} \int_{-\infty}^{+\infty} x^2 \phi(x)\, dx + \ln\sqrt{2\pi\sigma^2} \int_{-\infty}^{+\infty} \phi(x)\, dx = \\
&\frac{1}{2\sigma^2} Var(X) + \ln\sqrt{2\pi\sigma^2} = \frac{1}{2} + \ln\sqrt{2\pi\sigma^2}\text{nats}
\end{aligned}
$$

hence:

$$
\begin{aligned}
&\frac{1}{2} + \ln\sqrt{2\pi\sigma^2} = \frac{1}{2}\ln e + \ln\sqrt{2\pi\sigma^2} = \frac{1}{2}\ln e + \frac{1}{2}\ln 2\pi\sigma^2 = \\
&\frac{1}{2}\ln 2\pi e\sigma^2\text{nats}, \text{ which is } \frac{1}{2}\log_2 2\pi e\sigma^2\text{bits}
\end{aligned}
$$

Conditional differential entropy and mutual information

The differential entropy of a continuous random vector $X_1, X_2, \ldots, X_n$ with probability density function $f(x_1, x_2, \ldots, x_n)$, where $\int \int f(x_1, x_2, \ldots, x_n)\, dx_1 \cdots dx_n = 1$ is:

$$h(X_1, X_2, \ldots, X_n) =$$
$$- \int \int f(x_1, x_2, \ldots, x_n) \log f(x_1, x_2, \ldots, x_n)\, dx_1 \cdots dx_n$$

Conditional differential entropy

The conditional differential entropy of Y given X is:

$$h(Y|X) = -\int\int f(x,y)\log f(y|x)\,dxdy$$

$$h(Y|X) = -E_{f(x,y)}[\log f(y|x)]$$

Kullback-Leibler distance

The relative entropy or Kullback-Leibler distance between probability densities is:

$$D\left(f\left(x\right)\|g\left(x\right)\right) = \int f\left(x\right)\log\frac{f\left(x\right)}{g\left(x\right)}dx$$

Using Jensen's inequality for continuous random variables, it can be shown that $D\left(f\left(x\right)\|g\left(x\right)\right) \geq 0$ and that $D\left(f\left(x\right)\|g\left(x\right)\right) = 0$ if and only if $f\left(x\right) \equiv g\left(x\right)$.

Maximum entropy probability density function

For any random variable with variance bounded by σ^2, the Gaussian density $\phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$ has the largest differential entropy. We will verify it now.

Suppose $f(x)$ is any density function with $Var(x) \leq \sigma^2$. By the translation invariance of differential entropy, we may assume that the mean of $f(x)$ is 0.

$$0 \leq D(f \| \phi) =$$

$$
\begin{aligned}
0 &\le D(f \| \phi) = \int_{-\infty}^{+\infty} f(x) \ln \frac{f(x)}{\phi(x)} dx \\
&= \int_{-\infty}^{+\infty} f(x) \ln \left[f(x) \sqrt{2\pi\sigma^2} \exp\left(+ \frac{x^2}{2\sigma^2} \right) \right] dx \\
&= \int_{-\infty}^{+\infty} f(x) \ln f(x)\, dx + \int_{-\infty}^{+\infty} f(x) \ln \sqrt{2\pi\sigma^2} dx + \int_{-\infty}^{+\infty} f(x) \frac{x^2}{2\sigma^2} dx \\
&= -h(f(x)) + \frac{1}{2} \ln 2\pi\sigma^2 + \frac{1}{2\sigma^2} Var(x) \\
&= -h(f(x)) + \frac{1}{2} \ln 2\pi\sigma^2 + \frac{1}{2}
\end{aligned}
$$

Hence: $0 \le -h(f(x)) + \frac{1}{2} \ln 2e\pi\sigma^2$, which implies:

$$
h(f(x)) \le h(\phi(x))
$$