Information Theory: Lecture Notes 8

zqy1018

2020. 4. 15

Contents

1 Differential Entropy

We first give some basic definitions.

Definition. Let $F(x)$ be the c.d.f. (Cumulative Distribution Function) and $f(x)$ be the p.d.f. (Probability Density Function) of X. If $F(x)$ is continuous, then X is **continuous**. Let $S = \text{supp}(F) = \{x : f(x) > 0\}$ be the support of X.

We now try to define entropy for continuous random variables.

Definition. Then the **differential entropy** $h(X)$ of a continuous random variable X with density $f(x)$ is defined as

$$h(X) = -\int_S f(x) \log f(x) \mathrm{dx}$$

Note.

(1) Like the discrete case, $h(X)$ only depends on $f(x)$. So sometimes we use $h(f)$ instead of $h(X)$.

(2) $\forall c, h(X + c) = h(X)$. So transition does not change the entropy.

Remark. As in every example involving an integral, or even a density, we should include the statement *if it exists*. But sometimes for convenience we will assume that it exists.

2 Some Important Examples

2.1 Example: Uniform Distribution

For $X \sim U[0, a](a > 0)$, $h(X) = \log a$. So we can see that *differential entropy may be negative!*

2.2 Example: Normal Distribution

For $X \sim N(\mu, \sigma^2)$, we have

$$\begin{aligned} h(f) &= -\int f(x) \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) \mathrm{dx} + \int f(x) \frac{(x-\mu)^2}{2\sigma^2} \mathrm{dx} \\ &= \frac{1}{2} \log 2\pi\sigma^2 + \frac{DX}{2\sigma^2} \\ &= \frac{1}{2} \log 2\pi e\sigma^2 \end{aligned}$$

Note. Here we assume that the log function takes e as base.

2.3 Example: Multi-variable Normal Distribution

For $X \sim N(\mu, K)$, where K is the covariance matrix, we have

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}|K|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mu)^T K^{-1}(\mathbf{x}-\mu)\right]$$

Thus

$$\begin{aligned} h(f) &= -\int f(\mathbf{x})\left[-\frac{1}{2}(\mathbf{x}-\mu)^T K^{-1}(\mathbf{x}-\mu)\right] \mathrm{d}\mathbf{x} + \frac{1}{2}\log(2\pi)^n|K| \\ &= \frac{1}{2}E \end{aligned}$$

3 Other Definitions

Differential entropy is much similar as discrete entropy.

Acknowledgment

The contents are mainly based on the course materials of CS258, Shanghai Jiao Tong University and *Elements of Information Theory* by Thomas M. Cover.