ECE 612, Information Theory

Ganyu (Bruce) Xu (g66xu)

Winter, 2024

Preliminares

Definition 0.1. *The normal distribution $N(\mu, \sigma^2)$ has the probability density function*

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

Definition 0.2. *The joint normal distribution $N(\mu, K)$ is defined by probability density function:*

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(K)}} \exp(-\frac{1}{2}(\mathbf{x}-\mu)^\intercal K^{-1}(\mathbf{x}-\mu))$$

1 Entropy, mutual information, divergence

2 Entropy rate

3 Asymptotic equipartition property

4 Data compressions

5 Channel capacity

6 Differential entropy

Theorem 6.1 (Differential entropy of Gaussian distribution)**.** *Let X be Gaussian $N(0, \sigma^2)$, then*

$$h(X) = \frac{1}{2}\log(2\pi e \sigma^2)$$

Theorem 6.2. *Let $\mathbf{X}$ follow joint Gaussian distribution $N(\mathbf{0}, K)$, then:*

$$h(\mathbf{X}) = \frac{1}{2}\log((2\pi e)^n \det K)$$

7 Gaussian channel

Definition 7.1 (Gaussian channel with power constraint)**.**

Definition 7.2 (Information channel capacity)**.**

Theorem 7.1. *The information channel capacity of a Gaussian channel is*

$$\max_{f_X : E[X^2] \leq P} I(X;Y) = \frac{1}{2}\log(1 + \frac{P}{\sigma^2}) \tag{1}$$

Where P is the power constraint, and σ^2 is the variance of the Gaussian noise. The maximum is achieved when X follows Gaussian distribution $X \xleftarrow{\$} N(0, P)$

8 Rate distortion theory