2.11.3 微分熵的极大化问题

离散熵的极大化：？？？？条件

求连续变量的最大微分熵需要附加一些约束条件

1、幅值受限：随机变量的取值受限于某个区间之间。

定理 **2.2** 设X的取值受限于有限区间$[a,b]$，则X服从均匀分布时，其熵达到最大。

证明：因为X的取值受限于有限区间$[a,b]$，则有

$$\int_a^b f_X(x)dx = 1$$

因为要在以上约束条件下求微分熵的最大值，利用拉各朗日乘数法，令

$$F[f_x(x)] = -\int_a^b f_X(x)\log f_X(x)dx + \lambda[\int_a^b f_X(x)dx - 1]$$

$$F[f_x(x)] = \log e\int_a^b f_X(x)\ln\frac{2^\lambda}{f_X(x)}dx - \lambda$$

$$\le \log e\int_a^b f_X(x)\left[\frac{2^\lambda}{f_X(x)} - 1\right]dx - \lambda$$ （信息论不等式）

等号成立的充要条件是：$\dfrac{2^{\lambda}}{f_X(x)}=1$，即$f_X(x)=2^{\lambda}$

由约束条件决定常数 λ

$$\int_a^b f_X(x)dx=\int_a^b 2^{\lambda}dx=2^{\lambda}(b-a)=1$$

得 $f_X(x)=2^{\lambda}=\dfrac{1}{b-a}$,

即X服从均匀分布$f_X(x)=\begin{cases}\dfrac{1}{b-a}, & x\in[a,b]\\ 0, & x\notin[a,b]\end{cases}$

此时，最大微分熵为

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

2、方差受限：

设X的方差受限为σ^2,即

$$\int_{-\infty}^{\infty}(x-\mu)^2 f_X(x)dx=\sigma^2,\quad \mu\text{为X的均值}$$

$$\because \int_{-\infty}^{\infty}(x-\mu)^2 f_X(x)dx=\int_{-\infty}^{\infty}(x^2-2x\mu+\mu^2)f_X(x)dx$$

$$=\int_{-\infty}^{\infty}x^2 f_X(x)dx-\mu^2$$

$$\therefore \int_{-\infty}^{\infty}x^2 f_X(x)dx=\mu^2+\sigma^2=P$$

即，均值一定时，方差受限为σ^2等价平均功率受限于$P=\sigma^2+\mu^2$

定理2.3 设X的均值为 μ ，方差受限为 σ^2，则**X**服从高斯分布时，其熵达到最大。

证明：考虑方差受限条件，概率完备性条件，利用拉各朗日乘数法如下

构造函数

$$F[f_x(x)]$$

$$= -\int_{-\infty}^{\infty} f_X(x)\log f_X(x)dx$$

$$+\lambda_1[\int_{-\infty}^{\infty} f_X(x)dx-1]+\lambda_2[\int_{-\infty}^{\infty}(x-\mu)^2 f_X(x)dx-\sigma^2]$$

$$\lambda_1 \int_{-\infty}^{\infty} f_X(x)\,dx$$

$$= \int_{-\infty}^{\infty} f_X(x) \log 2^{\lambda_1}\,dx$$

$$= \log e \int_{-\infty}^{\infty} f_X(x) \ln 2^{\lambda_1}\,dx$$

同理$\lambda_2 \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\,dx$

$$= \log e \int_{-\infty}^{\infty} f_X(x) \ln 2^{\lambda_1 (x-\mu)^2}\,dx$$

$$\therefore F[f_x(x)]$$

$$= -\log e\int_{-\infty}^{\infty} f_X(x)\ln f_X(x)dx + \log e\int_{-\infty}^{\infty} f_X(x)\ln 2^{\lambda_1}dx$$

$$+\log e\int_{-\infty}^{\infty} f_X(x)\ln 2^{\lambda_2(x-\mu)^2}dx - \lambda_1 - \lambda_2\sigma^2$$

$$= \log e\int_{-\infty}^{\infty} f_X(x)\ln \frac{2^{\lambda_1}2^{\lambda_2(x-\mu)^2}}{f_X(x)}dx - \lambda_1 - \lambda_2\sigma^2$$

应用信息论不等式 $\ln x \le x-1$

$$F[f_x(x)] \le$$

$$\log e\int_{-\infty}^{\infty} f_X(x)\left[\frac{2^{\lambda_1}2^{\lambda_2(x-\mu)^2}}{f_X(x)} - 1\right]dx - \lambda_1 - \lambda_2\sigma^2$$

等号成立的充要条件为

$$\frac{2^{\lambda_1}2^{\lambda_2(x-\mu)^2}}{f_X(x)}=1,\ 即 f_X(x)=2^{\lambda_1}2^{\lambda_2(x-\mu)^2}$$

确定常数2^{λ_1}和常数2^{λ_2}：

$$\because \int_{-\infty}^{\infty} 2^{\lambda_2(x-\mu)^2}\,dx=\int_{-\infty}^{\infty} e^{\ln 2^{\lambda_2(x-\mu)^2}}\,dx=\int_{-\infty}^{\infty} e^{(x-\mu)^2\ln 2^{\lambda_2}}\,dx$$

$$=\int_{-\infty}^{\infty} e^{\frac{(x-\mu)^2}{1/(\ln 2^{\lambda_2})}}\,dx=\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{1/(\ln 2^{-\lambda_2})}}\,dx=\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\frac{1}{2\ln 2^{-\lambda_2}}}}\,dx$$

$$=\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\left(\sqrt{\frac{1}{2\ln 2^{-\lambda_2}}}\right)^2}}\,dx$$

由方差受限条件

$$\int_{-\infty}^{\infty}(x-\mu)^2 f_X(x)dx=\sigma^2，\text{即}$$

$$\int_{-\infty}^{\infty}(x-\mu)^2 2^{\lambda_1}2^{\lambda_2(x-\mu)^2}dx$$

$$=2^{\lambda_1}\int_{-\infty}^{\infty}(x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\left(\sqrt{\frac{1}{2\ln 2^{-\lambda_2}}}\right)^2}}dx$$

$$=2^{\lambda_1}\sqrt{2\pi}\sqrt{\frac{1}{2\ln 2^{-\lambda_2}}}\left(\sqrt{\frac{1}{2\ln 2^{-\lambda_2}}}\right)^2=\sigma^2 \qquad (1)$$

由概率完备性条件

$$\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x=1\text{，即}$$

$$\int_{-\infty}^{\infty} 2^{\lambda_1} 2^{\lambda_2 (x-\mu)^2} \mathrm{d}x = 2^{\lambda_1} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\left(\sqrt{\frac{1}{2\ln 2^{-\lambda_2}}}\right)^2}} dx$$

$$= 2^{\lambda_1} \sqrt{2\pi} \sqrt{\frac{1}{2\ln 2^{-\lambda_2}}} = 1 \qquad (2)$$

由式(1)和(2)可以求出

$$2^{\lambda_1}=\frac{1}{\sqrt{2\pi\sigma}};\quad 2^{\lambda_2}=e^{-\frac{1}{2\sigma^2}}$$

$$\therefore\quad f_X(x)=2^{\lambda_1}2^{\lambda_2(x-\mu)^2}=\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

即**X**服从均值为 μ 、方差为的高斯分布时，其熵最大为

$$\boldsymbol{h(X)=\frac{1}{2}\log_2 2\pi e\sigma^2}$$

见前面例题

特例：均值为零，平均功率与 σ^2相等，此时，

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

$$h(X) = \log\sqrt{2\pi eP}$$

意义：

实际信号平均功率总是受限的，若信号服从高斯分布时，其熵最大。传输信息时，使用高斯分布的输入信号较为有利；而噪声服从高斯分布则不利