信息论第十一讲作业解答

中国科学技术大学《信息论 A》006125.01 班助教组

2024 年 6 月 13 日

第 1题

Consider a channel with additive exponential noise, $Y = X + Z$, where the noise Z obeys an exponential distribution with mean λ, independent of X, and the input X has support $[0, \infty)$ and a mean constraint $\mathbf{E}X \le \mu$. Calculate the capacity-cost function $C(\mu)$ of this channel.

解: 对每个 x 有

$$h(Y|X=x) = h(Y-x|X=x) = h(Y-X|X=x) = h(Z|X=x) = h(Z) = \log_2(e\lambda),$$

所以 $h(Y|X) = \log_2(e\lambda)$. 因为 Y 取值于 $[0,\infty)$, $\mathbf{E}[Y] = \mathbf{E}[X] + \mathbf{E}[Z] \le \mu + \lambda$, 所以 $h(Y) \le \log_2(e(\mu+\lambda))$ 这样

$$I(X;Y) = h(Y) - h(Y|X) \le \log_2\left(1 + \frac{\mu}{\lambda}\right).$$

设随机变量 X 以概率 $\lambda/(\mu+\lambda)$ 取 0, 以概率 $\mu/(\mu+\lambda)$ 服从均值为 $\mu+\lambda$ 的指数分布. 这样 0 有邻域 A 使对每个 $t \in A$ 有

$$\mathbf{E}[e^{tX}] = \frac{\lambda}{\mu+\lambda} + \frac{\mu}{\mu+\lambda}\frac{1}{1-(\mu+\lambda)t}, \tag{1}$$

$$\mathbf{E}[e^{tZ}] = \frac{1}{1-\lambda t}. \tag{2}$$

用矩母函数的性质可得对每个 $t \in A$ 有 $\mathbf{E}[e^{tY}] = \mathbf{E}[e^{tX}]\mathbf{E}[e^{tZ}]$ 即

$$\mathbf{E}[e^{tY}] = \frac{1}{1-(\mu+\lambda)t}. \tag{3}$$

所以 Y 服从均值为 $\mu+\lambda$ 的指数分布, $h(Y) = \log_2(e(\mu+\lambda))$, $I(X;Y) = \log_2(1+\mu/\lambda)$. 同时 $P[X \ge 0] = 1$, $\mathbf{E}[X] = (\mu/(\mu+\lambda))(\mu+\lambda) = \mu$.

作为 $P[X \ge 0] = 1$ 和 $\mathbf{E}[X] \le \mu$ 的条件下 $I(X;Y)$ 的最大值, $C(\mu) = \log_2(1+\mu/\lambda)$. □

寻找 X 的分布的过程和上面的推导过程是相反的. 我们希望 Y 满足 $h(Y)=\log_2(e(\mu+\lambda))$, 就先假设 Y 服从均值为 $\mu+\lambda$ 的指数分布. 这样 Y 有 **??** 式所示的矩母函数. 用 **??** 式所示的 Z 的矩母函数除 Y 的矩母函数可以得到 X 具有 **??** 式所示的矩母函数. 所以 X 应该以概率 $\lambda/(\mu+\lambda)$ 取 0, 以概率 $\mu/(\mu+\lambda)$ 服从均值为 $\mu+\lambda$ 的指数分布.

第 2题

For a continuous random variable S with mean zero and variance σ^2, consider its ratedistortion function under the squared error distortion measure, $d(s,\hat{s})=(s-\hat{s})^2$.
a) Show that $R(D)=0$ when $D\geq\sigma^2$.
b) Show that $R(D)\geq h(S)-\frac{1}{2}\log(2\pi eD)$ when $D<\sigma^2$.
c) Show that $R(D)\leq\frac{1}{2}\log\frac{\sigma^2}{D}$ when $D<\sigma^2$.

证明: a):对于$D\geq\sigma^2$, 我们可以将S只编码为$\hat{S}=\mathbf{E}[S]=0$, 便有$\mathbf{E}[X-\mathbf{E}[X]]^2=\sigma^2\leq D$. 即此时我们可以达到比$D$更小的失真且码率为0.

b):类似于高斯分布下率失真的推导, 我们保留$h(S)$并使用最大熵原理便得到:

$$\begin{aligned}I(S;\hat{S})&=h(S)-h(S|\hat{S})\\&\geq h(S)-h(S-\hat{S})\\&\geq h(S)-\frac{1}{2}\log(2\pi eD)\end{aligned}$$

c): 为证明上界, 我们试图构造一个确切的信道, 能够达到这个界, 则我们最优的率失真函数一定不差于他, 令:

$$\hat{S}=\frac{\sigma^2-D}{\sigma^2}S+Z\,,\quad Z\sim\mathcal{N}(0,\frac{D(\sigma^2-D)}{\sigma^2})\perp X$$

此时:

$$\begin{aligned}\mathbf{E}[d(S,\hat{S})]&=\mathbf{E}[(S-\hat{S})^2]=\mathbf{E}[(\frac{D}{\sigma^2}S-Z)^2]=\frac{D^2}{\sigma^2}+\frac{D(\sigma^2-D)}{\sigma^2}=D\\h(\hat{S})-h(\hat{S}|S)&=h(\frac{\sigma^2-D}{\sigma^2}S+Z)-h(Z)\\&=\frac{1}{2}\log(2\pi e(\sigma^2-D))-\frac{1}{2}\log(2\pi e\frac{D(\sigma^2-D)}{\sigma^2})=\frac{1}{2}\log\frac{\sigma^2}{D}\end{aligned}$$

由于我们所求的是极小值, 我们便有$R(D)\leq\frac{1}{2}\log\frac{\sigma^2}{D}$.

注 1. *c)中的构造受第三题b)的启发.*

$\square$

第 3题

We have seen that a useful trick for calculating rate-distortion functions is to construct suitable test channels. But in the optimization problem for solving rate-distortion functions, we need to characterize the forward channel.

a) *What is the forward channel $P_{\hat{S}|S}$ for a Bernoulli source under Hamming distortion?*
b) *What is the forward channel $f_{\hat{S}|S}$ for a Gaussian source under squared error distortion?*
c) *Calculate the rate-distortion function for a Laplace source under absoluate error distortion; that is, $f_S(s) = \frac{1}{2b}e^{-|s|/b}$, and $d(s, \hat{s}) = |s - \hat{s}|$.*

a) 解: 设信源符号服从 Bernoulli(δ), $\delta \leq 1/2$. 根据第 3 讲讲义, 如果 $D \geq \delta$ 则 $\hat{S}$ 以概率 1 取 0.

再设 $0 \leq D < \delta$. 此时 $\hat{S}$ 服从 Bernoulli($(\delta - D)/(1 - 2D)$), 存在独立于 $\hat{S}$ 的 Bernoulli(D) 随机变量 Z 使 $S = \hat{S} \oplus Z$ 以概率 1 成立. 对所有 s, $\hat{s} \in \{0, 1\}$ 有

$$P_{\hat{S}|S}(\hat{s}|s) = \frac{P_{\hat{S}}(\hat{s})P_{S|\hat{S}}(s|\hat{s})}{P_S(s)}.$$

用这个公式可以求出

$$P_{\hat{S}|S}(0|0) = \frac{(1-\delta-D)(1-D)}{(1-\delta)(1-2D)}, P_{\hat{S}|S}(1|0) = \frac{D(\delta-D)}{(1-\delta)(1-2D)},$$
$$P_{\hat{S}|S}(0|1) = \frac{D(1-\delta-D)}{\delta(1-2D)}, P_{\hat{S}|S}(1|1) = \frac{(\delta-D)(1-D)}{\delta(1-2D)}. \qquad \square$$

b) 解: 设信源符号服从 $\mathcal{N}(0, \sigma^2)$. 根据第 10 讲讲义, 如果 $D \geq \sigma^2$ 则 $\hat{S}$ 以概率 1 取 0.

再设 $0 < D < \sigma^2$. 此时 $\hat{S}$ 服从 $\mathcal{N}(0, \sigma^2 - D)$, 存在独立于 $\hat{S}$ 的 $\mathcal{N}(0, D)$ 随机变量 Z 使 $S = \hat{S} + Z$ 以概率 1 成立. 对所有实数 s 和 $\hat{s}$ 有

$$\begin{aligned} f_{\hat{S}|S}(\hat{s}|s) &= \frac{f_{\hat{S}}(\hat{s})f_{S|\hat{S}}(s|\hat{s})}{f_S(s)} = \frac{f_{\hat{S}}(\hat{s})f_Z(s-\hat{s})}{f_S(s)} \\ &= \sqrt{\frac{\sigma^2}{2\pi D(\sigma^2 - D)}} \exp\left(-\frac{\sigma^2}{2D(\sigma^2 - D)}\left(\hat{s} - \frac{\sigma^2 - D}{\sigma^2}s\right)^2\right). \end{aligned} \qquad \square$$

在解 c 问之前我们先推导一点 Laplace 分布的性质.

引理 1. *设 $b > 0$, 对所有实数 x 有*

$$f(x) = \frac{1}{2b}e^{-|x|/b},$$

f 是随机变量 X 的概率密度函数. 这样 $\mathbf{E}[|X|] = b$, $h(X) = \log_2(2eb)$, 对所有 $t \in (-1/b, 1/b)$ 有

$$\mathbf{E}[e^{tX}] = \frac{1}{1 - b^2t^2}.$$

如果 Y 是随机变量且 $\mathbf{E}[|Y|] = b$ 则 $h(Y) \leq \log_2(2eb)$.

证明:

$$\begin{aligned}\mathbf{E}[|X|] &= \int_{-\infty}^{\infty}|x|f(x)dx = 2\int_0^{\infty}x\frac{1}{2b}e^{-x/b}dx = b,\\ h(X) &= -\mathbf{E}[\log_2(f(X))] = -\mathbf{E}\left[-\log_2(2b) - \frac{|X|}{b}\log_2(e)\right]\\ &= \log_2(2b) + \frac{\mathbf{E}[|X|]}{b}\log_2(e) = \log_2(2eb).\end{aligned}$$

对每个 $t \in (-1/b, 1/b)$,

$$\mathbf{E}[e^{tX}] = \int_{-\infty}^{\infty}e^{tx}f(x)dx = \int_{-\infty}^{0}e^{tx}\frac{1}{2b}e^{x/b}dx + \int_0^{\infty}e^{tx}\frac{1}{2b}e^{-x/b}dx = \frac{1}{1-b^2t^2},$$

如果 Y 是随机变量且 $\mathbf{E}[|Y|] = b$ 则

$$\begin{aligned}h(Y) &= -\mathbf{E}[\log_2(f_Y(Y))]\\ &= -\mathbf{E}\left[\log_2\left(\frac{f_Y(Y)}{f(Y)}\right)\right] - \mathbf{E}[\log_2(f(Y))]\\ &= -D(f_Y\|f) - \mathbf{E}[\log_2(f(Y))]\\ &\le -\mathbf{E}[\log_2(f(Y))]\\ &= -\mathbf{E}\left[-\log_2(2b) - \frac{|Y|}{b}\log_2(e)\right]\\ &= \log_2(2eb).\end{aligned}$$

□

c) 解: 用 R 表示这个率失真函数. 对每个 $D \in [b, \infty)$, 因为 $\mathbf{E}[d(S,0)] = \mathbf{E}[|S|] = b \le D$, $I(S;0) = 0$, 所以 $\mathbf{E}[d(S,\hat{S})] \le D$ 的条件下 $I(S;\hat{S})$ 的最小值等于 0, 即 $R(D) = 0$.

设 $0 < D < b$, $\mathbf{E}[d(S,\hat{S})] \le D$. 对每个 $\hat{s}$ 有 $h(S|\hat{S} = \hat{s}) = h(S-\hat{s}|\hat{S} = \hat{s}) = h(S-\hat{S}|\hat{S} = \hat{s})$, 所以

$$h(S|\hat{S}) = h(S-\hat{S}|\hat{S}) \le h(S-\hat{S}). \tag{4}$$

我们已经假设了 $\mathbf{E}[|S-\hat{S}|] \le D$. 再用引理 **??** 得

$$h(S-\hat{S}) \le \log_2(2e\mathbf{E}[|S-\hat{S}|]) \le \log_2(2eD). \tag{5}$$

这样 $I(S;\hat{S}) = h(S) - h(S|\hat{S}) \ge \log_2(2eb) - \log_2(2eD) = \log_2(b/D)$.

设 $0 < D < b$, 随机变量 $\hat{S}$ 以概率 D^2/b^2 取 0, 以概率 $(b^2-D^2)/b^2$ 服从参数为 b 的 Laplace 分布, 随机变量 Z 服从参数为 D 的 Laplace 分布, $\hat{S}$ 和 Z 独立. 定义 $S = \hat{S} + Z$. 这样 $\mathbf{E}[d(S,\hat{S})] = \mathbf{E}[|Z|] = D$. 对每个 $t \in (-1/b, 1/b)$ 有

$$\mathbf{E}[e^{tS}] = \mathbf{E}[e^{t\hat{S}}]\mathbf{E}[e^{tZ}] = \left(\frac{D^2}{b^2} + \frac{b^2-D^2}{b^2}\frac{1}{1-b^2t^2}\right)\frac{1}{1-D^2t^2} = \frac{1}{1-b^2t^2},$$

所以 S 服从参数为 b 的 Laplace 分布, 即信源符号的分布. 可以看出 **??** 和 **??** 式成立等号, 所以 $I(S;\hat{S}) = \log_2(b/D)$.

对每个 $D \in (0,b)$, 作为 $\mathbf{E}[d(S,\hat{S})] \le D$ 的条件下 $I(S;\hat{S})$ 的最小值, $R(D) = \log_2(b/D)$.

综上所述,

$$R(D) = \begin{cases} \log_2\left(\dfrac{b}{D}\right), & 0 < D < b \\ 0, & D \ge b \end{cases}. \qquad \square$$

第 4题

Consider a memoryless additive noise channel $Y = X + Z$ where X has sopprt $[-1/2, 1/2]$, and the noise Z is uniform over $[-1, 1]$, independent of X. Calculate the capacity of the channel, $C = \max_{f(X)} I(X;Y)$.

解:

$$I(X;Y) = I(X;X+Z) = h(X+Z) - h(Z) = h(X+Z) - 1.$$

我们注意到$X+Z$取值于$[-3/2, 3/2]$, 并且:

$$\begin{aligned} f_{X+Z}(X+Z=t) = f_X * f_Z(t) &= \int_{-1/2}^{1/2} f_X(x) f_Z(t-x)\,dx \\ &= \begin{cases} \frac{1}{2}\int_{t-1}^{1/2} f_X(x)\,dx & t \ge \frac{1}{2} \\ \frac{1}{2}\int_{-1/2}^{1/2} f_X(x)\,dx = 1/2 & -\frac{1}{2} \le t \le \frac{1}{2} \\ \frac{1}{2}\int_{-1/2}^{t+1} f_X(x)\,dx & t \le -\frac{1}{2} \end{cases}, \end{aligned}$$

即$X+Z$一定以1/2的概率密度分布在$[-1/2, 1/2]$上, 而在$[-3/2, -1/2]$和$[1/2, 3/2]$上的分布依赖于$f_X(x)$.

我们又由最大熵原理, 在固定支撑集上的最大熵为均匀分布, 那么f_{X+Z}在支撑集$[-3/2, -1/2] \bigcup [1/2, 3/2]$上的概率密度为1/4, 这要求:

$$\lim_{t \to -1/2} \frac{1}{2}\int_{-1/2}^{t+1} f_X(x) = \lim_{t \to 1/2} \frac{1}{2}\int_{t-1}^{1/2} f_X(x) = 1/4.$$

仔细观察上面两个式子, 即在1/2和$-1/2$的邻域上累积分布函数分别达到1/2. 那么在连续意义上, 我们可以用$\frac{1}{2}\left(\delta_n(x-1/2) + \delta_n(x+1/2)\right)$来描述之, $\delta_n(x)$为狄拉克函数. 若没有连续要求, 我们也可以令$P_X(x=-1/2) = P_X(x=1/2) = 1/2$来描述, 此时我们便很容易做出$Y$的分布, 观察到分布满足最大熵. 此时

$$h(X+Z) = \int_{-3/2}^{-1/2} 1/4 \log 4\,dx + \int_{-1/2}^{1/2} 1/2 \log 2\,dx + \int_{1/2}^{3/2} 1/4 \log 4\,dx = 3/2,$$

$$\max_{f(X)} I(X;Y) = 3/2 - 1 = 1/2.$$

□

第 5题

Consider an AWGN channel $Y = X + Z$, where X has an average power constraint P, and the noise $Z \sim \mathcal{N}(0, \sigma_Z^2)$. Suppose that, besides Y, the decoder also observes a noisy version of Z, $V = Z + W$ where $W \sim \mathcal{N}(0, \sigma_W^2)$ is independent of Z and X. What is the capacity of this channel model?

解: 对每个 x, 依次用 Y 的定义, 第 9 讲 Proposition 3 和 X, (Z, V) 的独立性得 $h(Y, V|X = x) = h(x + Z, V|X = x) = h(Z, V|X = x) = h(Z, V)$. 所以

$$h(Y, V|X) = h(Z, V). \tag{6}$$

Z 和 W 独立且都是高斯的, 所以是联合高斯的. 作为 Z 和 W 的线性组合, (Z, V) 是高斯的. 所以

$$h(Z, V) = \frac{1}{2}\log_2((2\pi e)^2 \det(\mathbf{K}_{Z,V})), \tag{7}$$

其中 $\mathbf{K}_{Z,V}$ 表示 Z 和 V 的协方差矩阵. 因为 $\mathrm{Cov}[Z, V] = \mathbf{E}[Z(Z + W)] = \mathbf{E}[Z^2] + \mathbf{E}[Z]\mathbf{E}[W] = \sigma_Z^2$, $\mathrm{var[V]} = \sigma_Z^2 + \sigma_W^2$, 所以

$$\det(\mathbf{K}_{Z,V}) = \begin{vmatrix} \sigma_Z^2 & \sigma_Z^2 \\ \sigma_Z^2 & \sigma_Z^2 + \sigma_W^2 \end{vmatrix} = \sigma_Z^2\sigma_W^2. \tag{8}$$

根据第 9 讲第 3 题,

$$h(Y, V) \le \frac{1}{2}\log_2((2\pi e)^2 \det(\mathbf{K}_{Y,V})), \tag{9}$$

其中 $\mathbf{K}_{Y,V}$ 表示 Y 和 V 的协方差矩阵. 因为 $\mathrm{var[Y]} = \mathrm{var[X]} + \sigma_Z^2$,

$$\begin{aligned}
\mathrm{Cov}[Y, V] &= \mathbf{E}[YV] - \mathbf{E}[Y]\mathbf{E}[V] \\
&= \mathbf{E}[XZ + XW + Z^2 + ZW] - \mathbf{E}[Y]\mathbf{E}[V] \\
&= \mathbf{E}[X]\mathbf{E}[Z] + \mathbf{E}[X]\mathbf{E}[W] + \mathbf{E}[Z^2] + \mathbf{E}[Z]\mathbf{E}[W] - \mathbf{E}[Y]\mathbf{E}[V] \\
&= \sigma_Z^2,
\end{aligned}$$

$\mathrm{var[V]} = \sigma_Z^2 + \sigma_W^2$, 所以

$$\det(\mathbf{K}_{Y,V}) = \begin{vmatrix} \mathrm{var[X]} + \sigma_Z^2 & \sigma_Z^2 \\ \sigma_Z^2 & \sigma_Z^2 + \sigma_W^2 \end{vmatrix} = (\sigma_Z^2 + \sigma_W^2)\mathrm{var[X]} + \sigma_Z^2\sigma_W^2 \le (\sigma_Z^2 + \sigma_W^2)P + \sigma_Z^2\sigma_W^2. \tag{10}$$

综合 $I(X;Y,V)=h(Y,V)-h(Y,V|X)$ 和 **??**, **??**, **??**, **??**, **??** 式得

$$I(X;Y,V)\le \frac{1}{2}\log_2\left(1+\frac{\sigma_Z^2+\sigma_W^2}{\sigma_Z^2\sigma_W^2}P\right). \tag{11}$$

设 $X\sim\mathcal{N}(0,P)$. X, Z 和 W 独立且都是高斯的, 所以是联合高斯的. 作为 X, Z 和 W 的线性组合, (Y,V) 是高斯的. 所以 **??** 式成立等号. 因为 $\mathrm{var[X]}=P$, 所以 **??** 式成立等号. 这样 **??** 式成立等号.

作为 $\mathbf{E}[X^2]\le P$ 的条件下 $I(X;Y,V)$ 的最大值, 信道容量

$$C(P)=\frac{1}{2}\log_2\left(1+\frac{\sigma_Z^2+\sigma_W^2}{\sigma_Z^2\sigma_W^2}P\right). \qquad \square$$

第 6题

Consider independent Gaussian random variables $X\sim\mathcal{N}\left(0,\sigma_X^2\right)$ and $Z\sim\mathcal{N}\left(0,\sigma_Z^2\right)$. If V satisfies $\mathbf{E}[XV]=0$ and $\mathbf{E}\left[V^2\right]\le\sigma_Z^2$, prove that $I(X;X+V)\ge I(X;X+Z)$, and discuss when equality holds.

证明: 由于$I(X;X+V)=h(X)-h(X|X+V)$, 欲证明$I(X;X+V)\ge I(X;X+Z)$, 只需证：

$$h(X|X+V)\le h(X|X+Z).$$

我们有下面的不等式串：

$$\begin{aligned}
h(X|X+V) &= h(X-\frac{\sigma_X^2}{\sigma_X^2+\mathbf{E}[V^2]}(X+V)|X+V) && (12)\\
&= h\left(\frac{\sigma_X^2V-\mathbf{E}[V^2]X}{\sigma_X^2+\mathbf{E}[V^2]}|X+V\right)\\
&\le h\left(\frac{\sigma_X^2V-\mathbf{E}[V^2]X}{\sigma_X^2+\mathbf{E}[V^2]}\right) && (13)\\
&\le \frac{1}{2}\log\left(2\pi e\mathbf{Var}\left[\frac{\sigma_X^2V-\mathbf{E}[V^2]X}{\sigma_X^2+\mathbf{E}[V^2]}\right]\right) && (14)\\
&\le \frac{1}{2}\log\left(2\pi e\mathbf{E}[\frac{\sigma_X^2V-\mathbf{E}[V^2]X}{\sigma_X^2+\mathbf{E}[V^2]}]^2\right) && (15)\\
&= \frac{1}{2}\log\left(2\pi e\frac{\sigma_X^2\mathbf{E}[V^2]}{\sigma_X^2+\mathbf{E}[V^2]}\right) && (16)\\
&\le \frac{1}{2}\log\left(2\pi e\frac{\sigma_X^2\sigma_Z^2}{\sigma_X^2+\sigma_Z^2}\right) && (17)\\
&= h(X|X+Z) && (18)
\end{aligned}$$

(??) 基于平移条件不影响条件熵, 系数k的选取是当V是与X独立的高斯变量时（这是我们希望的取等条件）, 保证$X-k(X+V)$和$X+V$独立, 那么此处的必要条件是：

$$\mathbf{E}[((1-k)X-kV)(X+V)]=(1-k)\sigma_X^2-k\mathbf{E}[V^2]=0\Rightarrow k=\frac{\sigma_X^2}{\sigma_X^2+\mathbf{E}[V^2]}$$

(??) 基于条件减少熵.
(??) 基于方差约束下的最大熵原理.
(??) 基于$\mathbf{E}[X^2]=\mathbf{E}[X]^2+\mathbf{Var}[X]\geq\mathbf{Var}[X]$.
(??) 基于展开后使用$\mathbf{E}[XV]=0$并化简.
(??) 基于函数$\frac{\sigma_X^2\mathbf{E}[V^2]}{\sigma_X^2+\mathbf{E}[V^2]}$关于$\mathbf{E}[V^2]$单调递增.
(??) 基于以下事实：

$$\begin{aligned}h(X|X+Z)&=h(X+Z|X)+h(X)-h(X+Z)\\&=\frac{1}{2}\log(2\pi e\sigma_Z^2)+\frac{1}{2}\log(2\pi e\sigma_X^2)-\frac{1}{2}\log(2\pi e(\sigma_X^2+\sigma_Z^2))\\&=\frac{1}{2}\log\left(2\pi e\frac{\sigma_X^2\sigma_Z^2}{\sigma_X^2+\sigma_Z^2}\right)\end{aligned}$$

设 $I(X;X+V)=I(X;X+Z)$. 这样 $h(X|X+V)=h(X|X+Z)$, 所以 (??), (??), (??) 和 (??) 成立等号. 因为 (??), (??) 和 (??) 成立等号, 所以

$$\frac{\sigma_X^2V-\mathbf{E}[V^2]X}{\sigma_X^2+\mathbf{E}[V^2]}$$

独立于 $X+V$, 服从正态分布, 且有均值 0. 因为 (??) 成立等号, 所以 $\mathbf{E}[V^2]=\sigma_Z^2$.

反过来, 如果

$$\frac{\sigma_X^2V-\mathbf{E}[V^2]X}{\sigma_X^2+\mathbf{E}[V^2]}$$

独立于 $X+V$, 服从正态分布, 有均值 0, 且 $\mathbf{E}[V^2]=\sigma_Z^2$, 则可以验证 $I(X;X+V)=I(X;X+Z)$.

□

第 7题

Consider the parallel Gaussian channels in Section IV.

a) *Derive the water-filling optimal solution.*
b) *Show that as $P\to\infty$, the rate loss due to using uniform power allocation $P_i=P/k$, $i=1,\ldots,k$, instead of the water-filling optimal solution, asymptotically vanishes.*

a) 证明: 用 A 表示满足 $\sum_{i=1}^k P_i \le P$ 的所有 $(P_1, P_2, \cdots, P_k) \in [0, \infty)^k$ 组成的集合.

$$\begin{aligned}
I(\underline{X};\underline{Y}) &= h(\underline{Y}) - h(\underline{Y}|\underline{X}) \\
&= h(\underline{Y}) - h(Z_1, Z_2, \cdots, Z_k) \\
&= h(\underline{Y}) - \sum_{i=1}^k \frac{1}{2}\log_2(2\pi e \sigma_i^2) \\
&\le \sum_{i=1}^k h(Y_i) - \sum_{i=1}^k \frac{1}{2}\log_2(2\pi e \sigma_i^2) \\
&\le \sum_{i=1}^k \frac{1}{2}\log_2(2\pi e(\mathbf{E}[X_i^2] + \sigma_i^2)) - \sum_{i=1}^k \frac{1}{2}\log_2(2\pi e \sigma_i^2) \\
&= \sum_{i=1}^k \frac{1}{2}\log_2\left(1 + \frac{\mathbf{E}[X_i^2]}{\sigma_i^2}\right) \\
&\le \max_{(P_1,P_2,\cdots,P_k)\in A} \sum_{i=1}^k \frac{1}{2}\log_2\left(1 + \frac{P_i}{\sigma_i^2}\right),
\end{aligned} \tag{19}$$

其中第二行的等号用了第 9 讲 Proposition 3, 第五行的小于等于号是因为对所有正整数 $i \le k$ 有 $\mathrm{var[Y_i] = var[X_i] + var[Z_i]} \le \mathbf{E}[X_i^2] + \sigma_i^2$, 最后一行的小于等于号是因为 $\sum_{i=1}^k \mathbf{E}[X_i^2] \le P$.

设 $(P_1^*, P_2^*, \cdots, P_k^*) \in A$,

$$\sum_{i=1}^k \frac{1}{2}\log_2\left(1 + \frac{P_i^*}{\sigma_i^2}\right) = \max_{(P_1,P_2,\cdots,P_k)\in A} \sum_{i=1}^k \frac{1}{2}\log_2\left(1 + \frac{P_i}{\sigma_i^2}\right), \tag{20}$$

X_1, X_2, $\cdots$, X_k 分别服从 $\mathcal{N}(0, P_1^*)$, $\mathcal{N}(0, P_2^*)$, $\cdots$, $\mathcal{N}(0, P_k^*)$ 且独立. 这样 Y_1, Y_2, $\cdots$, Y_k 独立, 对每个正整数 $i \le k$, Y_i 服从 $\mathcal{N}(0, \mathbf{E}[X_i^2] + \sigma_i^2)$. 所以 **??** 式第四和第五行的小于等于号现在是等号. **??** 式最后一行的小于等于号现在当然也是等号.

作为 $I(\underline{X};\underline{Y})$ 的最大值, $C(P)$ 等于 **??** 式的最后一行. 设 $(P_1^*, P_2^*, \cdots, P_k^*) \in A$ 且 **??** 式成立. 我们用 KKT 条件 [**?**, 5.49 式] 来求 $(P_1^*, P_2^*, \cdots, P_k^*)$. 可以看出

$$-\sum_{i=1}^k \ln\left(1 + \frac{P_i^*}{\sigma_i^2}\right) = \min_{(P_1,P_2,\cdots,P_k)\in A}\left(-\sum_{i=1}^k \ln\left(1 + \frac{P_i}{\sigma_i^2}\right)\right).$$

如果 μ, ν_1, ν_2, $\cdots$, $\nu_k \ge 0$ 则对所有 P_1, P_2, $\cdots$, $P_k \ge 0$ 和正整数 $i \le k$ 有

$$\frac{\partial}{\partial P_i}\left(-\sum_{j=1}^k \ln\left(1 + \frac{P_j}{\sigma_j^2}\right) + \mu\left(\sum_{j=1}^k P_j - P\right) + \sum_{j=1}^k \nu_j(-P_j)\right) = -\frac{1}{\sigma_i^2 + P_i} + \mu - \nu_i.$$

所以存在非负数 μ, ν_1, ν_2, $\cdots$, ν_k 满足

$$\sum_{i=1}^k P_i^* - P \le 0,$$

$$-P_i^* \le 0, \forall i \in \{1, 2, \cdots, k\},$$

$$\mu\left(\sum_{i=1}^{k} P_i^* - P\right) = 0, \tag{21}$$

$$\nu_i(-P_i^*) = 0, \forall i \in \{1, 2, \cdots, k\}, \tag{22}$$

$$-\frac{1}{\sigma_i^2 + P_i^*} + \mu - \nu_i = 0, \forall i \in \{1, 2, \cdots, k\}, \tag{23}$$

即 KKT 条件. 由 **??** 式得 $\mu = 1/(\sigma_1^2 + P_1^*) + \nu_1 > 0$, 对每个正整数 $i \le k$ 有

$$P_i^* = \frac{1}{\mu - \nu_i} - \sigma_i^2.$$

如果正整数 $i \le k$, $\nu_i = 0$, 则 $P_i^* = 1/\mu - \sigma_i^2$, $1/\mu - \sigma_i^2 \ge 0$. 如果正整数 $i \le k$, $\nu_i > 0$, 则由 **??** 式得 $P_i^* = 0$, 所以 $1/(\mu - \nu_i) = \sigma_i^2 > 0$, $1/\mu - \sigma_i^2 < 1/(\mu - \nu_i) - \sigma_i^2 = 0$. 总之对每个正整数 $i \le k$ 有

$$P_i^* = \max\left(\frac{1}{\mu} - \sigma_i^2, 0\right). \tag{24}$$

因为 $\mu > 0$ 且 **??** 式成立, 所以 $\sum_{i=1}^{k} P_i^* = P$,

$$\sum_{i=1}^{k} \max\left(\frac{1}{\mu} - \sigma_i^2, 0\right) = P. \tag{25}$$

因此只要 $(P_1^*, P_2^*, \cdots, P_k^*) \in A$ 且 **??** 式成立就存在 $\mu > 0$ 满足 **??** 式且使 **??** 式对所有正整数 $i \le k$ 成立. 把 μ 换成 -2λ 我们就得到了讲义的 39 和 40 式. 我们在前面构造的 $(X_1, X_2, \cdots, X_k)$ 服从均值为 0_k, 协方差矩阵为

$$\begin{bmatrix} P_1^* & & & \\ & P_2^* & & \\ & & \ddots & \\ & & & P_k^* \end{bmatrix}$$

的高斯分布. □

b) 证明: 根据 a 问的结论, 对每个 $P \ge 0$ 存在 $L(P) \ge 0$ (即 a 问中的 $1/\mu$) 满足

$$\sum_{i=1}^{k} \max(L(P) - \sigma_i^2, 0) = P,$$

$$C(P) = \sum_{i=1}^{k} \frac{1}{2} \log_2\left(1 + \frac{\max(L(P) - \sigma_i^2, 0)}{\sigma_i^2}\right),$$

其中 C 表示容量-代价函数. 设正整数 $m \le k$, $\sigma_m = \max(\sigma_1, \sigma_2, \cdots, \sigma_k)$, $N = \sum_{i=1}^k \sigma_i^2$. 如果 $P \ge 0$ 且 $L(P) \le \sigma_m^2$ 则

$$P \le \sum_{i=1}^k \max(\sigma_m^2 - \sigma_i^2, 0) = k\sigma_m^2 - N,$$

所以对每个 $P > k\sigma_m^2 - N$ 有 $L(P) > \sigma_m^2$, 也就有

$$P = \sum_{i=1}^k (L(P) - \sigma_i^2) = kL(P) - N,$$

$$C(P) = \sum_{i=1}^k \frac{1}{2} \log_2 \left(1 + \frac{L(P) - \sigma_i^2}{\sigma_i^2}\right) = \sum_{i=1}^k \frac{1}{2} \log_2 \left(\frac{P+N}{k\sigma_i^2}\right).$$

对每个 $P > k\sigma_m^2 - N$, 均匀分配功率达到的码率与容量的差为

$$C(P) - \sum_{i=1}^k \frac{1}{2} \log_2 \left(1 + \frac{P}{k\sigma_i^2}\right) = \sum_{i=1}^k \frac{1}{2} \log_2 \left(\frac{P+N}{P+k\sigma_i^2}\right).$$

$P \to \infty$ 时这个差趋于 0. $\square$