证明推论6.1。
a) 对于容量 - 成本函数 $C_I(\Gamma)$ ，当 $\Gamma < \Gamma_{\min} := \min_{x \in \mathcal{X}} c(x)$ 时，优化问题 (1) 不可行，此时我们简单地令 $C_I(\Gamma) = 0$ 。
b) 存在一个阈值 $\Gamma_{\max}$ ，使得对于任意 $\Gamma \geq \Gamma_{\max}$ ，有 $C_I(\Gamma) = C_I(\infty)$ ，$C_I(\infty)$ 是无成本约束时的容量 。
c) 容量 - 成本函数 $C_I(\Gamma)$ 是关于 Γ 的非递减凹函数。
d) 若 $C_I(\infty) > 0$ ，则对于 $\Gamma \in [\Gamma_{\min}, \Gamma_{\max}]$ ，$C_I(\Gamma)$ 严格递增，并且在优化问题 (1) 中，不等式约束可以替换为等式约束。

信息论第六讲作业解答

中国科学技术大学《信息论 A》006125.01 班助教组

2025 年 5 月 7 日

第 1 题

Prove Corollary 6.1.

$$C_I(\Gamma) = \max_{P_X} I(X;Y), \quad s.t.\ \mathbf{E}[c(X)] \leq \Gamma. \tag{1}$$

a):For capacity-cost function $C_I(\Gamma)$, the optimization problem (1) is infeasible for any $\Gamma < \Gamma_{\min} := \min_{x\in\mathcal{X}} c(x)$ and we simply set $C_I(\Gamma) = 0$ therein.

b):There exists a threshold $\Gamma_{\max}$ such that for any $\Gamma \geq \Gamma_{\max}, C_I(\Gamma) = C_I(\infty)$ which is the capacity without cost constraint.

c):The capacity-cost function $C_I(\Gamma)$ is a non-decreasing concave function in Γ.

d):If $C_I(\infty) > 0$, then $C_I(\Gamma)$ is strictly increasing for $\Gamma \in [\Gamma_{\min}, \Gamma_{\max}]$, and in the optimization problem (1) the inequality constraint can be replaced by an equality constraint.

证明: a): 对于任意的 P_X, 我们均有：

$$\mathbf{E}[c(X)] = \sum_{x\in\mathcal{X}} p(x)c(x) \geq \Gamma_{\min} := \min_{x\in\mathcal{X}} c(x),$$

从而对于所有的 $\Gamma < \Gamma_{\min}$, 这样的信源分布是不可能存在的, 认为此时无法按要求传递信息, 视 $C_I(\Gamma) = 0$.

b): 当我们在无约束下求解 $\max_{P_X} I(X;Y)$ 得到一组 $\mathcal{P}^*$, 从中取得 $p^*(X)$ 使得 $\mathbf{E}[c(X)]$ 最小, 此时便有：

$$\Gamma = \sum p^*(x)c(x) := \Gamma_{\max} \leq \max_{x\in\mathcal{X}} c(x) \tag{2}$$

那么对于任意 $\Gamma \geq \Gamma_{\max}$, 我们仍然能在约束 $\Gamma_{\max}$ 下取到 $C_I(\Gamma)$ 的极值, 这也便意味着 $C_I(\Gamma) = C_I(\infty)$, 等价于无约束.

c): 对于任意的 $\Gamma_{\min} \leq \Gamma_1 < \Gamma_2$, 记满足优化问题中约束的分布集合为 $\mathcal{P}_\Gamma$, 那么满足 Γ_1 约束时一定满足 Γ_2 约束, 即 $\mathcal{P}_{\Gamma_1} \subseteq \mathcal{P}_{\Gamma_2}$, 那么 $\mathcal{P}_{\Gamma_1}$ 下的最优解一定也属于 $\mathcal{P}_{\Gamma_2}$, 因为我们求的是极大值, 所以容量代价函数是非减的.

下面进一步证明代价容量函数是凹的, 即对于 $\Gamma_{\min} \le \Gamma_1 < \Gamma_2$ 以及 $0 \le \lambda \le 1$ 有：

$$C_I((1-\lambda)\Gamma_1 + \lambda\Gamma_2) \ge (1-\lambda)C_I(\Gamma_1) + \lambda C_I(\Gamma_2),$$

为此, 我们取在 Γ_1 和 Γ_2 约束下达到极值的概率密度函数 $P_{1,X}, P_{2,X}$, 定义 $P_{\lambda,X} = (1-\lambda)P_{1,X} + \lambda P_{2,X}$, 我们有：

$$\begin{aligned}
\mathbf{E}_{X\sim P_{\lambda,X}}[c(X)] &= \sum_{x\in\mathcal{X}} P_{\lambda,X}(x)c(x) \\
&= (1-\lambda)\sum_{x\in\mathcal{X}} P_{1,X}(x)c(x) + \lambda \sum_{x\in\mathcal{X}} P_{2,X}(x)c(x) \\
&\le (1-\lambda)\Gamma_1 + \lambda\Gamma_2,
\end{aligned}$$

从而在约束 $(1-\lambda)\Gamma_1 + \lambda\Gamma_2$ 下我们至少拥有了一个分布 $P_{\lambda,X} = (1-\lambda)P_{1,X} + \lambda P_{2,X}$, 又因为我们求的是极大值：

$$C_I\big((1-\lambda)\Gamma_1 + \lambda\Gamma_2\big) \ge I(X;Y)|_{P_{\lambda,X}}. \tag{3}$$

进一步, 由互信息关于 P_X 的凹性 (concave), 我们又有：

$$I(X;Y)|_{P_{\lambda,X}} \ge (1-\lambda)I(X;Y)|_{P_{1,X}} + \lambda I(X;Y)|_{P_{2,X}} = (1-\lambda)C_I(\Gamma_1) + \lambda C_I(\Gamma_2). \tag{4}$$

由 (3) 和 (4)：

$$C_I\big((1-\lambda)\Gamma_1 + \lambda\Gamma_2\big) \ge (1-\lambda)C_I(\Gamma_1) + \lambda C_I(\Gamma_2).$$

d): 若不是严格增长的, 我们一定可以找到 $\Gamma_{\min} \le \Gamma_1 < \Gamma_2 \le \Gamma_{\max}$, 使得 $C_I(\Gamma_1) = C_I(\Gamma_2) = C_I(\Gamma)$, 其中 $\Gamma_1 \le \Gamma \le \Gamma_2$, 由我们上面证明的凹性：

$$\begin{aligned}
\Gamma_2 &= \frac{\Gamma_{max} - \Gamma_2}{\Gamma_{max} - \Gamma_1}\Gamma_1 + \frac{\Gamma_2 - \Gamma_1}{\Gamma_{max} - \Gamma_1}\Gamma_{\max} \\
C_I(\Gamma_2) &\ge \frac{\Gamma_{max} - \Gamma_2}{\Gamma_{max} - \Gamma_1}C_I(\Gamma_1) + \frac{\Gamma_2 - \Gamma_1}{\Gamma_{max} - \Gamma_1}C_I(\Gamma_{\max})
\end{aligned}$$

但是我们又假定 $C_I(\Gamma_1) = C_I(\Gamma_2)$, 那么只有 $C_I(\Gamma_1) = C_I(\Gamma_2) = C_I(\Gamma_{\max})$, 但是这又和我们对于 $\Gamma_{\max}$ 的定义矛盾, 所以容量代价函数是严格递增的.

进一步, 若存在 $\Gamma_{\min} \le \Gamma_1 \le \Gamma_{\max}$ 使得在约束下优化问题 (1) 取到极值时, 有 $\mathbf{E}[c(X)] = \Gamma' < \Gamma_1$, 那么我们可以断言 $C_I(\Gamma') \ge C_I(\Gamma_1)$, 但是这又和严格递增矛盾, 命题得证. $\square$

第 2 题

计算以下离散无记忆信道（DMC）的容量 - 成本函数：
a) 二元对称信道（BSC），其中 $c(0) = 0$ 且 $c(1) = 1$ ；
b) 二元删除信道（BEC），其中 $c(0) = 0$ 且 $c(1) = 1$ 。

Calculate the capacity-cost function for the following DMC:

a) BSC with $c(0) = 0$ and $c(1) = 1$;

b) BEC with $c(0) = 0$ *and* $c(1) = 1$.

a) 解: 首先可得 $\Gamma_{\min} = \min_x c(x) = 0$, 根据讲义 Example 6.1, 可知无约束下 BSC 信道的容量为 $1 - h_2(\delta)$, 并且此时最优分布为 $P_X^*(0) = P_X^*(1) = \frac{1}{2}$. 根据第 1 题 b) 中的 (2) 式, 计算可得 $\Gamma_{\max} = \frac{1}{2}$, 所以当 $\Gamma > \Gamma_{\max}$ 时, 有$C(\Gamma) = 1 - h_2(\delta)$ bits.

接下来我们考虑 $\Gamma_{\min} \leq \Gamma \leq \Gamma_{\max}$, 假设信源 X 的分布为 $P_X(1) = p$, 分析互信息

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H(Y) - \sum_{x=0,1} P_X(x) h_2(\delta) \\ &= h_2\left((1-p)\delta + p(1-\delta)\right) - h_2(\delta) \\ &= h_2\left((1-2\delta)\, p + \delta\right) - h_2(\delta), \end{aligned} \tag{5}$$

此时根据约束条件我们有 $\mathbf{E}[c(X)] = p \leq \Gamma \leq \Gamma_{\max}$, 显然 $(1-2\delta)\, p + \delta \leq \frac{1}{2}$, 所以 (5) 式最大值为 $h_2\left((1-2\delta)\,\Gamma + \delta\right) - h_2(\delta)$.

综上，容量-代价函数为:

$$C(\Gamma) = \begin{cases} h_2\left((1-2\delta)\,\Gamma + \delta\right) - h_2(\delta), & 0 \leq \Gamma \leq \frac{1}{2} \\ 1 - h_2(\delta), & \Gamma > \frac{1}{2} \end{cases}.$$

□

b) 解: 首先可得 $\Gamma_{\min} = \min_x c(x) = 0$, 根据讲义 Example 6.2, 可知无约束下 BEC 信道的容量为 $1 - \delta$, 并且此时最优分布为 $P_X^*(0) = P_X^*(1) = \frac{1}{2}$, 可得 $\Gamma_{\max} = \frac{1}{2}$, 所以当 $\Gamma > \Gamma_{\max}$ 时, 有$C(\Gamma) = 1 - \delta$ bits.

接下来我们考虑 $\Gamma_{\min} \leq \Gamma \leq \Gamma_{\max}$, 假设信源 X 的分布为 $P_X(1) = p$, 分析互信息

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= h_2(p) - [(1-p)(1-\delta)h_2(0) + p(1-\delta)h_2(1) + \delta h_2(p)] \\ &= (1-\delta) h_2(p), \end{aligned} \tag{6}$$

约束为 $\mathbf{E}[c(X)] = p \leq \Gamma \leq \Gamma_{\max}$, 所以 (6) 最大值为 $(1-\delta)h_2(\Gamma)$.

综上，容量-代价函数为:

$$C(\Gamma) = \begin{cases} (1-\delta) h_2(\Gamma), & 0 \leq \Gamma \leq \frac{1}{2} \\ 1 - \delta, & \Gamma > \frac{1}{2} \end{cases}.$$

□

3 计算以下离散无记忆信道（DMC）的容量：a) 具有 “ 无用 ” 输入的离散无记忆信道，其信道转移概率矩阵 \(P_{Y|X}\) 为：（提示：对称性很有用，但你需要解释为什么利用对称性进行简化不会损失最优性。）b) 非对称二元删除信道（BEC），其信道转移概率矩阵 \(P_{Y|X}\) 为：c)其信道转移概率矩阵 \(P_{Y|X}\) 为：

第 3 题

Calculate the capacity for the following DMC:

a) A DMC with a “useless” input,

$$P_{Y|X} = \begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1 \end{bmatrix}.$$

(Hint: Symmetry is helpful, but you need to explain why your simplification using symmetry does not lose optimality.)

b) An asymmetric BEC,

$$P_{Y|X} = \begin{bmatrix} 1-\alpha & \alpha & 0 \\ 0 & \beta & 1-\beta \end{bmatrix}.$$

c)

$$P_{Y|X} = \begin{bmatrix} 1 & 0 \\ \alpha & \beta \\ 0 & 1 \end{bmatrix}.$$

a) 解: 可以验证对所有 $a \in (0,1) \cup (1,\infty)$ 和 $p \in (0,1)$ 有

$$\frac{d}{dp}(-p\log_a(p) - (1-p)\log_a(1-p)) = \log_a\left(\frac{1-p}{p}\right). \tag{7}$$

在求信道容量和率失真函数时我们可以直接用这个结论.

方法一: 设 $Y|X \sim P_{Y|X}$. 定义随机变量 Y_2 在 $Y=2$ 时取 1, 否则取 0.

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H(Y) + H(Y_2|Y) - H(Y|X) \\ &= H(Y_2) + H(Y|Y_2) - \log_2(3)P_X(2) \\ &= H(Y_2) + P_{Y_2}(0)H(Y|Y_2=0) - \log_2(3)P_X(2) \\ &\le H(Y_2) + P_{Y_2}(0) - \log_2(3)P_X(2) \\ &= h_2\left(\frac{P_X(2)}{3}\right) + 1 - \frac{P_X(2)}{3} - \log_2(3)P_X(2), \end{aligned}$$

其中第二行的等号是因为 $H(Y_2|Y)=0$, 第四行的等号是因为如果 $P_{Y_2}(1)>0$ 则 $H(Y|Y_2=1)=0$, 最后一行的等号是因为 Y_2 以概率 $P_X(2)/3$ 取 1, 以概率 $1-P_X(2)/3$ 取 0. 用

$g(t)=h_2(t/3)+1-t/3-\log_2(3)t$ 定义 $g:[0,1]\to\mathbb{R}$. 利用 (7) 式的结论, 对每个 $t\in(0,1)$ 有

$$g'(t)=\frac{1}{3}\log_2\left(\frac{3-t}{54t}\right),$$

所以 g 在 $[0,3/55]$ 上递增, 在 $[3/55,1]$ 上递减. 因此

$$I(X;Y)\le g\left(\frac{3}{55}\right)=\log_2\left(\frac{55}{27}\right).$$

如果 X 分别以概率 26/55, 3/55 和 26/55 取 1, 2 和 3, 则 $H(Y|Y_2=0)=H(Y|Y\neq 2)=1$,

$$h_2\left(\frac{P_X(2)}{3}\right)+1-\frac{P_X(2)}{3}-\log_2(3)P_X(2)=g\left(\frac{3}{55}\right),$$

所以 $I(X;Y)=\log_2(55/27)$. 作为 $Y|X\sim P_{Y|X}$ 的条件下 $I(X;Y)$ 的最大值, 信道容量等于 $\log_2(55/27)$ 比特. □

方法一不涉及 "simplification using symmetry". 方法二说明了 "simplification using symmetry does not lose optimality" 的意思. □

方法二: 设 $Y|X\sim P_{Y|X}$. 定义随机变量 X_1 和 Y_1 满足 $P_{X_1}(1)=P_X(3)$, $P_{X_1}(2)=P_X(2)$, $P_{X_1}(3)=P_X(1)$ 和 $Y_1|X_1\sim P_{Y|X}$. 再定义随机变量 X_2 和 Y_2 满足 $P_{X_2}=(P_X+P_{X_1})/2$ 和 $Y_2|X_2\sim P_{Y|X}$. 可以看出 $I(X;Y)=I(X_1;Y_1)$. 根据互信息的关于 P_X 凹性,

$$\frac{1}{2}I(X;Y)+\frac{1}{2}I(X_1;Y_1)\le I(X_2;Y_2).$$

所以 $I(X;Y)\le I(X_2;Y_2)$. 可以算出

$$P_{X_2}(1)=P_{X_2}(3)=\frac{1-P_X(2)}{2},P_{X_2}(2)=P_X(2),$$
$$I(X_2;Y_2)=h_2\left(\frac{P_X(2)}{3}\right)+1-\frac{P_X(2)}{3}-\log_2(3)P_X(2)=g(P_X(2)).$$

这里 g 的定义和方法一中一样. 所以我们还是只需要求 g 在 $[0,1]$ 上的最大值. □

b) 解:

引理 1. 一个输入分布 P_X 是离散无记忆信道 $P_{Y|X}$ 的最优分布（即达到信道容量）, 当且仅当存在常数 C 使得

$$\begin{cases} I(x;Y)=C & \forall x\in\mathcal{X} \text{ with } P_X(x)>0;\\ I(x;Y)\le C & \forall x\in\mathcal{X} \text{ with } P_X(x)=0.\end{cases}$$

其中 $I(x;Y)=\sum_{y\in\mathcal{Y}}P_{Y|X}(y|x)\log\frac{P_{Y|X}(y|x)}{P_Y(y)}$, 那么该常数 C 就是信道容量 *(证明见注 1)*.

考虑本题的非对称 BEC 信道 $P_{Y|X}$, 若最优输入分布为 $P_X(0)=0$ 或1, 显然可算得 $I(X;Y)=0$. 那么可以假设最优分布为 $P_X(0)=p_0^*\in(0,1)$ 和 $P_X(1)=p_1^*\in(0,1)$, 则根据上述引理, 可得 $I(0;Y)=I(1;Y)$, 即

$$p_0^*+p_1^*=1,$$
$$(1-\alpha)\log\frac{(1-\alpha)}{(1-\alpha)p_0^*}+\alpha\log\frac{\alpha}{\alpha p_0^*+\beta p_1^*}=\beta\log\frac{\beta}{\alpha p_0^*+\beta p_1^*}+(1-\beta)\log\frac{(1-\beta)}{(1-\beta)p_1^*}.$$

我们记 $x=\frac{p_0^*}{p_1^*}$, 将上式化简可得

$$\alpha\log\alpha-\beta\log\beta=(1-\beta)\log(\alpha x+\beta)-(1-\alpha)\log(\alpha+\tfrac{\beta}{x}). \tag{8}$$

(8) 式左侧是常数，右侧是关于 x 的单调增函数，使得正好存在一个 x^* 使得 (8) 成立.

最后，我们可以通过求解 x^* 来得到最优输入分布 $P_X(0)=\frac{x^*}{1+x^*}$ 和 $P_X(1)=\frac{1}{1+x^*}$, 进而计算出该信道的信道容量. □

c) 解: 假定 $\mathcal{X}=\{0,e,1\},\mathcal{Y}=\{0,1\}$, 输入分布为 $P_X(0)=p_0,P_X(e)=p_e,P_X(1)=1-p_0-p_e$, 则可以 $P_Y(0)=p_0+\alpha p_e$, 接下来分析最大化 $I(X;Y)$ 问题.

$$\begin{aligned}I(X;Y)&=H(Y)-H(Y|X)\\&=H(Y)-\sum_x P_X(x)H(Y|X=x)\\&=h_2(p_0+\alpha p_e)-p_e h_2(\alpha)\end{aligned} \tag{9}$$

记 $f(p_0,p_e)=h_2(p_0+\alpha p_e)-p_e h_2(\alpha)$, 考虑优化问题 $\max_{p_0,p_e} f(p_0,p_e)$, 利用分块坐标下降的思想，首先固定 p_e, 更新 p_0, 利用 (7) 式的结论，则有

$$\frac{\partial f}{\partial p_0}=\log_2\frac{1-p_0-\alpha p_e}{p_0+\alpha p_e}=0$$

所以有 $p_0=\frac{1}{2}-\alpha p_e$, 代入到 (9) 式，可得此时只关于 p_e 的优化问题为

$$\max_{p_e} -p_e h_2(\alpha) \tag{10}$$

显然最优解为 $p_e^*=0$, 进而有 $p_0^*=\frac{1}{2}$. 最后将 p_e^*,p_0^* 代入 (9) 式算得该信道容量 $C=1$ bit.

我们可以用引理 1 来验证该结果. 根据输入分布 $P_X=[\frac{1}{2},0,\frac{1}{2}]$, 可得输出分布为 $P_Y=[\frac{1}{2},\frac{1}{2}]$, 进而可分别计算得

$$\begin{aligned}I(0;Y)&=\log_2 2=1\\I(e;Y)&=\alpha\log_2 2\alpha+(1-\alpha)\log_2 2(1-\alpha)=1-h_2(\alpha)\le 1\\I(1;Y)&=\log_2 2=1\end{aligned}$$

所以该信道的容量为 1 bit. □

注 1 (引理 1 的证明). 通过 *KKT (Karush-Kuhn-Tucker)* 条件 *[1, pp.243]* 可以证明该引理. 容量的最大化问题等价于下列优化问题

$$\begin{aligned} -C = \underset{P_X}{\text{minimize}} \quad & -I(X;Y) \\ \textit{subject to} \quad & \sum_x P_X(x) = 1 \\ & -P_X(x) \le 0, \forall x \in \mathcal{X} \end{aligned} \tag{11}$$

该优化问题的拉格朗日函数为

$$\mathcal{L}(\mathbf{x}, \lambda, \mathbf{u}) = -I(X;Y) + \lambda\left(\sum_x P_X(x) - 1\right) - \sum_x u_x P_X(x) \tag{12}$$

其中 λ 和 $u_x, \forall x$ 分别是拉格朗日乘子. 考察 *KKT* 条件, 则有

$$\frac{\partial \mathcal{L}}{\partial P_X(x)} = -I(x;Y) + \lambda + 1 - u_x = 0 \tag{13}$$

$$\sum_x P_X(x) = 1, P_X(x) \ge 0 \tag{14}$$

$$u_x \ge 0, u_x P_X(x) = 0 \tag{15}$$

根据 *(15)* 式, 若 $u_x = 0$, 则 $I(x;Y) = \lambda + 1$; 若 $P_X(x) = 0$, 则 $I(x;Y) = \lambda + 1 - u_x \le \lambda + 1$. 所以信道容量为 $C = I(X;Y) = \sum_x P_X(x) I(x;Y) = \lambda + 1$, 我们记常数 $C = \lambda + 1$, 即可完成引理的证明.

注 2. 关于 *(13)* 式详细推导如下, 我们假设对数的底是 e.

$$\begin{aligned}
&\frac{\partial \mathcal{L}}{\partial P_X(x)} = \\
&\frac{\partial}{\partial P_X(x)}\left\{-\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_X(x) P_{Y|X}(y|x) \log \frac{P_{Y|X}(y|x)}{P_Y(y)} + \lambda\left(\sum_x P_X(x) - 1\right) - \sum_x u_x P_X(x)\right\} \\
&= -\frac{\partial}{\partial P_X(x)}\left\{\sum_{x\in\mathcal{X}, y\in\mathcal{Y}} P_X(x) P_{Y|X}(y|x)\left(\log P_{Y|X}(y|x) - \log \sum_{x'\in\mathcal{X}} P_Y(y)\right)\right\} + \lambda - u_x \\
&= -\frac{\partial}{\partial P_X(x)}\left\{\sum_{x\in\mathcal{X}, y\in\mathcal{Y}} P_X(x) P_{Y|X}(y|x) \log P_{Y|X}(y|x)\right. \\
&\qquad\left. - \sum_{x\in\mathcal{X}, y\in\mathcal{Y}} P_X(x) P_{Y|X}(y|x) \log \sum_{x'\in\mathcal{X}} P_X(x') P_{Y|X}(y|x')\right\} + \lambda - u_x \\
&= -\left\{\sum_{y\in\mathcal{Y}} P_{Y|X}(y|x) \log P_{Y|X}(y|x)\right.
\end{aligned}$$

$$
\begin{aligned}
&- \left(\sum_{y\in\mathcal{Y}} P_{Y|X}(y|x) \log P_Y(y) + \sum_{x\in\mathcal{X}, y\in\mathcal{Y}} P_X(x) P_{Y|X}(y|x) \frac{P_{Y|X}(y|x')}{\sum_{x'\in\mathcal{X}} P_X(x') P_{Y|X}(y|x')} \right) \Bigg\} \\
&+ \lambda - u_x \\
= &- \left\{ \sum_{y\in\mathcal{Y}} P_{Y|X}(y|x) \log \frac{P_{Y|X}(y|x)}{P_Y(y)} - \sum_{y\in\mathcal{Y}} \frac{P_{Y|X}(y|x') \sum_{x\in\mathcal{X}} P_X(x) P_{Y|X}(y|x)}{\sum_{x'\in\mathcal{X}} P_X(x') P_{Y|X}(y|x')} \right\} + \lambda - u_x \\
= &- I(x;Y) + \sum_{y\in\mathcal{Y}} P_{Y|X}(y|x') + \lambda - u_x \\
= &- I(x;Y) + \lambda + 1 - u_x
\end{aligned}
$$

第 4 题

定义对称离散无记忆信道（DMC）的方式有多种。其中一种定义如下：若离散无记忆信道 $P_{Y|X} = [p_{i,j}]_{i = 1,\ldots,|\mathcal{X}|, j = 1,\ldots,|\mathcal{Y}|}$ 的每一行都是其他行的一个排列，且每一列都是其他列的一个排列，则该信道是对称的。求这种对称离散无记忆信道的容量，并证明当输入服从均匀概率分布时可以达到该容量。

There can be several ways of defining symmetric DMC. One of them is as follows: a DMC $P_{Y|X} = [p_{i,j}]_{i=1,\ldots,|\mathscr{X}|, j=1,\ldots,|\mathscr{Y}|}$ is symmetric if every row is a permutation of every other row, and every column is a permutation of every other column. Find the capacity of such a symmetric DMC and show that it can be achieved when the input follows a uniform probability distribution.

解: 用 $\mathbf{q} = (q_1, q_2, \ldots, q_{|\mathcal{Y}|})$ 表示信道转移矩阵的任一行。根据对称性，可以得到

$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y|X) \\
&\overset{(a)}{=} H(Y) - \sum_x P_X(x) H(q_1, q_2, \ldots, q_{|\mathcal{Y}|}) \\
&= H(Y) - H(\mathbf{q}) \\
&\le \log|\mathcal{Y}| - H(\mathbf{q}),
\end{aligned} \tag{16}
$$

其中 (a) 式成立的原因是对称信道任一行为其他每行的排列．此时当输入分布为均分分布时，对 $\forall y$ 可得

$$
\begin{aligned}
P_Y(y) &= \sum_x P_X(x) P_{Y|X}(y|x) \\
&= \frac{1}{|\mathcal{X}|} \sum_x P_{Y|X}(y|x) \overset{(b)}{=} \frac{A}{|\mathcal{X}|}
\end{aligned} \tag{17}
$$

其中 (b) 式成立的原因是对称信道任一列也为其他每列的排列，所以每列之和相等．又根据 $\sum_y P_Y(y) = 1$, 可得

$$
\sum_y \frac{A}{|\mathcal{X}|} = 1 \Longrightarrow \frac{A|\mathcal{Y}|}{|\mathcal{X}|} = 1 \tag{18}
$$

所以可得 $\forall y, P_Y(y) = \frac{A}{|\mathcal{X}|} = \frac{1}{|\mathcal{Y}|}$, 即此时 (16) 式中等号成立.

综上可得，对于一个对称 DMC, 其容量为

$$C = \log|\mathcal{Y}| - H(\mathbf{q}), \tag{19}$$

并且当输入为均匀分布时可以达到上述容量. □

第 5 题 我们可以适当放宽上一题中对对称离散无记忆信道（DMC）的要求，将弱对称离散无记忆信道 $P_{Y|X} = [p_{i,j}]_{i = 1,\ldots,|\mathcal{X}|, j = 1,\ldots,|\mathcal{Y}|}$ 定义如下：矩阵 $P_{Y|X}$ 的列可以划分为若干子集，使得对于每个子集，由其对应列所构成的子矩阵满足上一题中的对称条件 。求这种弱对称离散无记忆信道的容量，并证明当输入服从均匀概率分布时可以达到该容量。

We can slightly relax the requirement on a symmetric DMC in the previous exercise and define a weakly symmetric DMC $P_{Y|X} = [p_{i,j}]_{i=1,\ldots,|\mathscr{X}|, j=1,\ldots,|\mathscr{Y}|}$ as follows: the columns of matrix $P_{Y|X}$ can be partitioned into several subsets, such that for each subset, the submatrix formed by its corresponding columns satisfies the symmetric condition in the previous exercise. Find the capacity of such a weakly symmetric DMC and show that it can be achieved when the input follows a uniform probability distribution.

解: 将 $P_{Y|X}$ 的列划分为 k 个子集，形成 k 个大小为 $|\mathcal{X}| \times |\mathcal{Y}_i|$ 的对称子矩阵 $P_{Y_1|X}, P_{Y_2|X}, \ldots, P_{Y_k|X}$，其中 $\mathcal{Y}_1 \cup \cdots \cup \mathcal{Y}_k = \mathcal{Y}, \mathcal{Y}_i \cap \mathcal{Y}_j = \emptyset, \forall i \neq j, i, j \in \{1, \ldots, k\}$.

记 $a_i = \sum_{y \in \mathcal{Y}_i} P_{Y_i|X}(y|x)$, 定义

$$\tilde{P}_{Y_i|X} = \frac{1}{a_i} P_{Y_i|X}, \quad \forall i \in \{1, \ldots, k\}. \tag{20}$$

所以 $\tilde{P}_{Y_i|X}$ 是一个合法的对称 DMC. 此时利用上一题的结论，我们有

$$C_i = \max_{P_X} I(X; Y_i) = \log|\mathcal{Y}_i| - H(\mathbf{q}), \tag{21}$$

其中 $\mathbf{q}^{(i)}$ 是 $\tilde{P}_{Y_i|X}$ 的任一一行, 依然在输入均匀分布的条件下达到信道容量. 考虑整个信道的互信息

$$\begin{aligned} I(X;Y) &= \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} P_X(x) P_{Y|X}(y|x) \log \frac{P_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} P_X(x') P_{Y|X}(y|x')} \\ &= \sum_{i=1}^{k} a_i \sum_{y \in \mathcal{Y}_i} \sum_{x \in \mathcal{X}} P_X(x) \frac{P_{Y|X}(y|x)}{a_i} \log \frac{P_{Y|X}(y|x)/a_i}{\sum_{x'} P_X(x') P_{Y|X}(y|x')/a_i} \\ &= \sum_{i=1}^{k} a_i \sum_{y \in \mathcal{Y}_i} \sum_{x \in \mathcal{X}} P_X(x) \tilde{P}_{Y_i|X}(y|x) \log \frac{\tilde{P}_{Y_i|X}(y|x)}{\sum_{x'} P_X(x') \tilde{P}_{Y_i|X}(y|x')} \\ &= \sum_{i=1}^{k} a_i I(X; Y_i). \end{aligned} \tag{22}$$

因此该信道的容量为

$$\begin{aligned} C = \max_{P_X} I(X;Y) &= \max_{P_X} \sum_{i=1}^{k} a_i I(X;Y_i) = \sum_{i=1}^{k} a_i C_i \\ &= \sum_{i=1}^{k} a_i \big(\log |\mathcal{Y}_i| - H(\mathbf{q}^{(i)}) \big). \end{aligned} \tag{23}$$

并且当输入为均匀分布时可以达到上述容量. □

第 6 题

考虑一个离散无记忆信道（DMC），其信道转移概率矩阵为：。。a) 除了 $\mathcal{Y}$ 中的输出外，增加一个额外的删除输出 e，具体如下：对于任意输入 $x \in \mathcal{X}$，信道以概率 α 输出 e，否则根据 $P_{Y|X}$ 输出 $y \in \{1, \ldots, |\mathcal{Y}|\}$。证明这样一个信道的容量为 $(1 - \alpha)C$，其中 C 是原始离散无记忆信道 $P_{Y|X}$ 的容量。

Consider a DMC with $P_{Y|X} = [p_{i,j}]_{i=1,\ldots,|\mathscr{X}|, j=1,\ldots,|\mathscr{Y}|}$.

a) *Add an additional erasure output e besides $\mathscr{Y}$ as follows: for any input $x \in \mathscr{X}$, the channel outputs e with probability α, otherwise outputs $y \in \{1, \ldots, |\mathscr{Y}|\}$ according to $P_{Y|X}$. Prove that the capacity of such a channel is $(1-\alpha)C$, where C is the capacity of the original DMC $P_{Y|X}$.*

b) *Instead, add an additional erasure output e besides $\mathscr{Y}$ as follows: first pass an input $x \in \mathscr{X}$ through $P_{Y|X}$ to get an output $y \in \mathscr{Y}$, and then change this output to e with probability α. What is the capacity of such a channel?*

证明: a): 我们记原先的转移概率矩阵为 $P_{Y|X}$, 之后的为 $\tilde{P}_{Y|X}$.

$$\begin{aligned} I(X;\tilde{Y}) &= H(X) - H(X|\tilde{Y}) \\ &= H(X) - P(\tilde{Y} = e)H(X|\tilde{Y} = e) - P(\tilde{Y} \neq e)H(X|\tilde{Y} \neq e) \\ &= H(X) - \alpha H(X) - (1-\alpha) \sum_{y \neq e} p(\tilde{Y} = y_j | \tilde{Y} \neq e) H(X|\tilde{Y} = y_j) \\ &= (1-\alpha)H(X) - (1-\alpha)H(X|Y) \\ &= (1-\alpha)I(X;Y). \end{aligned}$$

b) 换一种方式
,除了 $\mathcal{Y}$ 中的输出外,增加一个额外的删除输出 e
,具体如下:首先将输入 $x \in \mathcal{X}$ 通过 $P_{Y|X}$ 得到输出 $y \in \mathcal{Y}$,然后以概率 α 将这个输出变为 e 。这样一个信道的容量是多少?

我们提供使用另一种互信息计算方式的解法:

$$\begin{aligned} I(X;\tilde{Y}) &= H(\tilde{Y}) - H(\tilde{Y}|X) \\ &= H(\alpha, 1-\alpha) + (1-\alpha)H(\tfrac{p(\tilde{y_1})}{1-\alpha}, ..., \tfrac{p(\tilde{y_m})}{1-\alpha}) \\ &\quad - \sum_x P_X(x)[H(\alpha, 1-\alpha) + (1-\alpha)H(\tfrac{p(\tilde{y_1}|x)}{1-\alpha}, ..., \tfrac{p(\tilde{y_m}|x)}{1-\alpha})] \\ &= (1-\alpha)H(Y) - (1-\alpha)H(Y|X) \\ &= (1-\alpha)I(X;Y). \end{aligned}$$

综上所述, 对于任意一个分布 $P_{Y|X}$, 我们均有 $I(X;\tilde{Y}) = (1-\alpha)I(X;Y)$, 从而:

$$\max_{P_X} I(X;\tilde{Y}) = \max_{P_X}(1-\alpha)I(X;Y) = (1-\alpha)C.$$

b): 我们只需说明两个信道的等价性, 或者说, 重点关注 X 到 $\tilde{Y}$ 的转移概率矩阵是否一致. 此问中设计了两个连续的信道, $X \to Y \to \tilde{Y}$, 对于任意 x:

$$\begin{aligned} p(\tilde{y}_j|x) &= \sum_{y\in\mathscr{Y}} p(y|x)p(\tilde{y}_j|y) \\ &= \begin{cases} (1-\alpha)p(y_j|x), & \tilde{y}_j \neq e \\ \sum_y p(y|x)\alpha = \alpha, & \tilde{y}_j = e \end{cases} \end{aligned}$$

很容易看出这和 a) 问中描述是一致的. □

第 7 题 若一个离散无记忆信道（DMC）是交叉概率 $\delta < 1/2$ 的二元对称信道（BSC），在不考虑输入成本的情况下，当 P_X 达到信道容量 C 时，描述译码规则（6.45）、（6.46）和（6.50），并找出它们的差异。对二元删除信道（BEC）重复此练习。

If a DMC is BSC with crossover probability $\delta < 1/2$, without considering input cost, when P_X achieves C, describe the decoding rules (6.45) (6.46) and (6.50), and find out their differences. Repeat this exercise for BEC.

证明: 不考虑输入代价时,

- 对于 BSC 信道, 容量 $C = 1 - h_2(\delta)$, 达到此容量的分布为 $P_X(0) = P_X(1) = \frac{1}{2}$. 此时可计算得

$$\log_2 \frac{P_{Y|X}(y_i|\mathbf{C}_i(w))}{P_Y(y_i)} = \begin{cases} \log_2 2(1-\delta), & \text{if } y_i = \mathbf{C}_i(w) \\ \log_2 2\delta, & \text{if } y_i \neq \mathbf{C}_i(w) \end{cases} \tag{24}$$

  所以

$$i(\mathbf{C}(w);\underline{y}) = \frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{P_{Y|X}(y_i|\mathbf{C}_i(w))}{P_Y(y_i)} = \frac{1}{n}\big((n-d_w)\log_2 2(1-\delta) + d_w \log_2 2\delta\big) \tag{25}$$

  其中 d_w 表示 $\mathbf{C}(w)$ 和 $\underline{y}$ 之间的码距, 也就是 $d_w(\mathbf{C}(w),\underline{y}) = \sum_{i=1}^{n} \mathbb{I}(y_i \neq \mathbf{C}_i(w))$.
  首先考虑解码规则 (6.46) 式, 即 $|i(\mathbf{C}(w);\underline{y}) - I(X;Y)| \leq \epsilon$, 其中此时 $I(X;Y) = 1 - h_2(\delta)$, 最后代入 (25) 式化简得

$$\Big|\big(\frac{d_w}{n} - \delta\big)\log_2 \frac{\delta}{1-\delta}\Big| \leq \epsilon \tag{26}$$

  也就是找到在码本 $\mathbf{C}$ 中找到唯一一个 w 使得上式成立, 我们就可以认为这个 w 是 $\underline{y}$ 的解码结果.

接下来考虑解码规则 (6.50) 式, 即 $\hat{w} = \arg\max_{w=1,\dots,M_n} P_{\underline{Y}|\underline{X}}(\underline{y}|\mathbf{C}(w)$, 可得

$$\begin{aligned}\hat{w} &= \arg\max_{w=1,\dots,M_n} \prod_{i=1}^{n} P_{Y|X}(y_i|C_i(w)) &(27)\\ &= \arg\max_{w=1,\dots,M_n} (1-\delta)^{n-d_w}\delta^{d_w} &(28)\\ &= \arg\min_{w=1,\dots,M_n} d_w &(29)\end{aligned}$$

最后一个等式是因为 $\delta < \frac{1}{2}$, 也就是说在码本 $\mathbf{C}$ 中找到一个 w 使得该码字与 $\underline{y}$ 之间的码距最小, 该 w 也就是 $\underline{y}$ 的解码结果.

- 对于擦除概率为 δ 的 BEC 信道, 容量 $C = 1 - \delta$, 达到此容量的分布为 $P_X(0) = P_X(1) = \frac{1}{2}$. 此时可计算得

$$\log_2 \frac{P_{Y|X}(y_i|\mathbf{C}_i(w))}{P_Y(y_i)} = \begin{cases} 1, & \text{if } y_i \neq e \text{ and } y_i = \mathbf{C}_i(w) \\ -\infty, & \text{if } y_i \neq e \text{ and } y_i \neq \mathbf{C}_i(w) \\ 0 & \text{if } y_i = e \end{cases} \tag{30}$$

首先考虑解码规则 (6.46) 式, d_w 定义与上述问题中相同, 依然表示 $\mathbf{C}(w)$ 和 $\underline{y}$ 之间对应位置符号不相等的个数, 另外我们记 d_e 表示 $\underline{y}$ 中符号 e 的个数, 所以

$$\begin{aligned}&|i(\mathbf{C}(w);\underline{y}) - I(X;Y)| \leq \epsilon &(31)\\ \Longrightarrow &\left|\frac{1}{n}\big((n-d_w) + (d_w - d_e)(-\infty)\big) - (1-\delta)\right| \leq \epsilon, &(32)\end{aligned}$$

要满足上述规则，要使得 $d_w = d_e$, 即找到一个 w 使得 $\mathbf{C}(w)$ 与 $\underline{y}$ 中非擦除位置的符号完全相同.
接下来考虑解码规则 (6.50) 式, 由于 $P_{Y|X}(0|1) = P_{Y|X}(1|0) = 0$, 并且 $P_{Y|X}(e|0) = P_{Y|X}(e|1) = \delta$, 所以该规则也等价于即找到一个 w 使得 $\mathbf{C}(w)$ 与 $\underline{y}$ 中非擦除位置的符号完全相同.

第 8 题

考虑一个删除概率 $\alpha > 0$ 且具有无噪声无延迟反馈的二元删除信道（BEC）。设计一种分组编码方法（即所有信道码字长度均为 n，且 n 不依赖于反馈），使其在 n 趋于无穷时，能达到信道容量 $(1 - \alpha)$ 比特 / 信道使用。所设计的方法应是明确的，而非像6.4节可达性证明中使用的随机码本。

Consider a BEC with erasure probability $\alpha > 0$ and with noiseless feedback without delay. Design a block coding method (i.e., all channel codewords having the same length n which does not depend upon the feedback) that achieves the channel capacity $(1-\alpha)$ bits/channel use (as n tends to infinity). The designed method should be explicit, not the random codebook as used in the proof of the achievability part in Section 6.4.

解: 设 $E_1, E_2, \cdots$ 是一列独立同分布的随机变量, E_1 分别以 $1-\alpha$ 和 α 的概率取 0 和 1, $1/2 < \beta < 1$.

用 n 表示一个任意的正整数. 记

$$\nu(n) = \lceil \alpha n + n^\beta \rceil, \{0,1\}^{n-\nu(n)} = \{\underline{b}(1), \underline{b}(2), \cdots, \underline{b}(2^{n-\nu(n)})\}.$$

让一个编码器对每个正整数 $w \le 2^{n-\nu(n)}$ 用 $\underline{b}(w)$ 传输消息 w. 如果 $\underline{b}(w)$ 的某个符号被信道擦除, 编码器应该重传这个符号, 直到符号被正确传输或已经使用了 n 次信道. 如果 $\underline{b}(w)$ 的最后一个符号被正确传输时编码器使用信道的次数小于 n, 编码器接下来可以传输 0 直到信道已被使用 n 次. 再设计以下译码器. 如果 $\underline{y} \in \{0,1,e\}^n$ 中 e 的个数小于等于 $\nu(n)$ 则存在唯一的正整数 $w \le 2^{n-\nu(n)}$ 使 $\underline{b}(w)$ 是 $\underline{y}$ 中不等于 e 的部分的前缀, 译码器收到 $\underline{y}$ 后应该输出这个 w. 如果 $\underline{y} \in \{0,1,e\}^n$ 中 e 的个数大于 $\nu(n)$, 译码器收到 $\underline{y}$ 后应该输出 1. 这个码的码率为

$$\frac{\log_2(2^{n-\nu(n)})}{n} = \frac{n-\nu(n)}{n} > \frac{n - \alpha n - n^\beta - 1}{n} = 1 - \alpha - n^{\beta-1} - \frac{1}{n}.$$

对每个正整数 $w \le 2^{n-\nu(n)}$, 可以看出消息 w 被传错的概率 $P_{e,w}^{(n)}$ 小于等于被擦除的符号多于 $\nu(n)$ 的概率, 所以

$$P_{e,w}^{(n)} \le P\left[\sum_{i=1}^n E_i > \nu(n)\right] \le P\left[\sum_{i=1}^n E_i > \alpha n + n^\beta\right].$$

可以看出 $\sum_{i=1}^n E_i$ 的均值为 αn, 方差为 $n\alpha(1-\alpha)$. 这样最大错误概率

$$P_{e,\max}^{(n)} \le P\left[\sum_{i=1}^n E_i > \alpha n + n^\beta\right] \le \frac{n\alpha(1-\alpha)}{(n^\beta)^2} = \alpha(1-\alpha)n^{1-2\beta},$$

其中第二个不等式关系用了 Chebyshev 不等式.

因为 $n \to \infty$ 时码率的下界 $1 - \alpha - n^{\beta-1} - 1/n$ 趋于 $1-\alpha$, $P_{e,\max}^{(n)}$ 趋于 0, 所以上述编译码方法达到 $1-\alpha$ 的码率. □

第 9 题

不考虑渐近趋近于零的错误概率准则，而是考虑严格的零错误信道传输 。将零错误可达速率定义为：存在某个码字长度的码，使得错误概率严格为零的码率；将零错误容量定义为所有零错误可达速率的上确界。

a) 证明对于二元对称信道（BSC），只要交叉概率 $\delta > 0$，零错误容量就为零。

Instead of the asymptotically vanishing error probability criterion, consider strictly zero-error channel transmission. Define a zero-error achievable rate as the code rate at which there exists a code of a certain codeword length such that the probability of error is strictly zero, and define the zero-error capacity as the supremum of all zero-error achievable rates.

a) *Argue that for a BSC, the zero-error capacity is zero whenever $\delta > 0$.*

b) 求含 26 个字母的有噪打字机信道的零错误容量 。有噪打字机信道定义如下：///
c) 证明对于 \(\mathcal{X} = \mathcal{Y} = \{a, b, c, d, e\}\) 的有噪打字机信道，零错误速率 \(\frac{1}{2}\log_2{5}\) 比特 / 信道使用 是可达的。

b) Find the zero-error capacity of the noisy typewritter channel with 26 letters. The noisy typewritter channel is as follows: $\mathcal{X} = \mathcal{Y} = \{a, b, c, \ldots, z\}$, $P_{Y|X}(a|a) = P_{Y|X}(b|b) = \cdots = P_{Y|X}(y|y) = P_{Y|X}(z|z) = 1/2$, *and* $P_{Y|X}(b|a) = P_{Y|X}(c|b) = \cdots = P_{Y|X}(z|y) = P_{Y|X}(a|z) = 1/2$.

c) Show that the zero-error rate $\frac{1}{2}\log_2 5$ *bits/channel use is achievable for a noisy typewriter channel with* $\mathcal{X} = \mathcal{Y} = \{a, b, c, d, e\}$.

a) 证明: 设正整数 $M \geq 2$, f 是从 $\{1, 2, \cdots, M\}$ 到 $\{0, 1\}^n$ 的函数, g 是从 $\{0, 1\}^n$ 到 $\{1, 2, \cdots, M\}$ 的函数. 我们来证明分组码 (f, g) 的错误概率大于 0. 任取

$$\underline{y} = (y_1, y_2, \cdots, y_n) \in \{0, 1\}^n.$$

因为 $M \geq 2$, 所以存在正整数 $m \leq M$ 满足 $m \neq g(\underline{y})$. 设

$$f(m) = (x_1, x_2, \cdots, x_n).$$

用 (f, g) 发送的 m 被误认为是 $g(\underline{y})$ 的概率大于等于 $f(m)$ 被信道传成 $\underline{y}$ 的概率, 也就是 $\prod_{i=1}^{n} P_{Y|X}(y_i|x_i)$. 因为 $P_{Y|X}(y|x) > 0$ 对所有 $x, y \in \{0, 1\}$ 成立, 所以 $\prod_{i=1}^{n} P_{Y|X}(y_i|x_i) > 0$. 因此 (f, g) 的错误概率大于 0.

换句话说, 所有码率大于 0 的分组码都有大于 0 的错误概率, 所有大于 0 的码率都不可达. 作为可达码率的上确界, 零错容量等于 0. □

[2] 的 7.1.3 小节也讲到了打字机信道. 我们可以把打字机信道形象地画成图 1 的左半部分.

b) 解: 用

$$f(1) = \text{a}, f(2) = \text{c}, \cdots, f(13) = \text{y},$$
$$g(\text{a}) = g(\text{b}) = 1, g(\text{c}) = g(\text{d}) = 2, \cdots, g(\text{y}) = g(\text{z}) = 13$$

定义 $f : \{1, 2, \cdots, 13\} \to \mathcal{X}$ 和 $g : \mathcal{Y} \to \{1, 2, \cdots, 13\}$. (f, g) 是一个长度为 1 的分组码, 可以用图 1 右图表示. (f, g) 不会出错, 且有码率 $\log_2(13)/1 = \log_2(13)$, 所以 $\log_2(13)$ 是零错可达码率. 零错容量大于等于 $\log_2(13)$.

如果一列分组码的错误概率都等于 0, 则它们的错误概率当然也是趋于 0 的. 所以零错容量不会超过普通的信道容量即 $Y|X \sim P_{Y|X}$ 的条件下 $I(X;Y)$ 的最大值. 如果 $Y|X \sim P_{Y|X}$, 则 $I(X;Y) = H(Y) - H(Y|X) \leq \log_2(26) - \log_2(2) = \log_2(13)$. 所以普通的信道容量不超过 $\log_2(13)$, 零错容量也不会超过 $\log_2(13)$.

综上所述, 零错容量等于 $\log_2(13)$. □

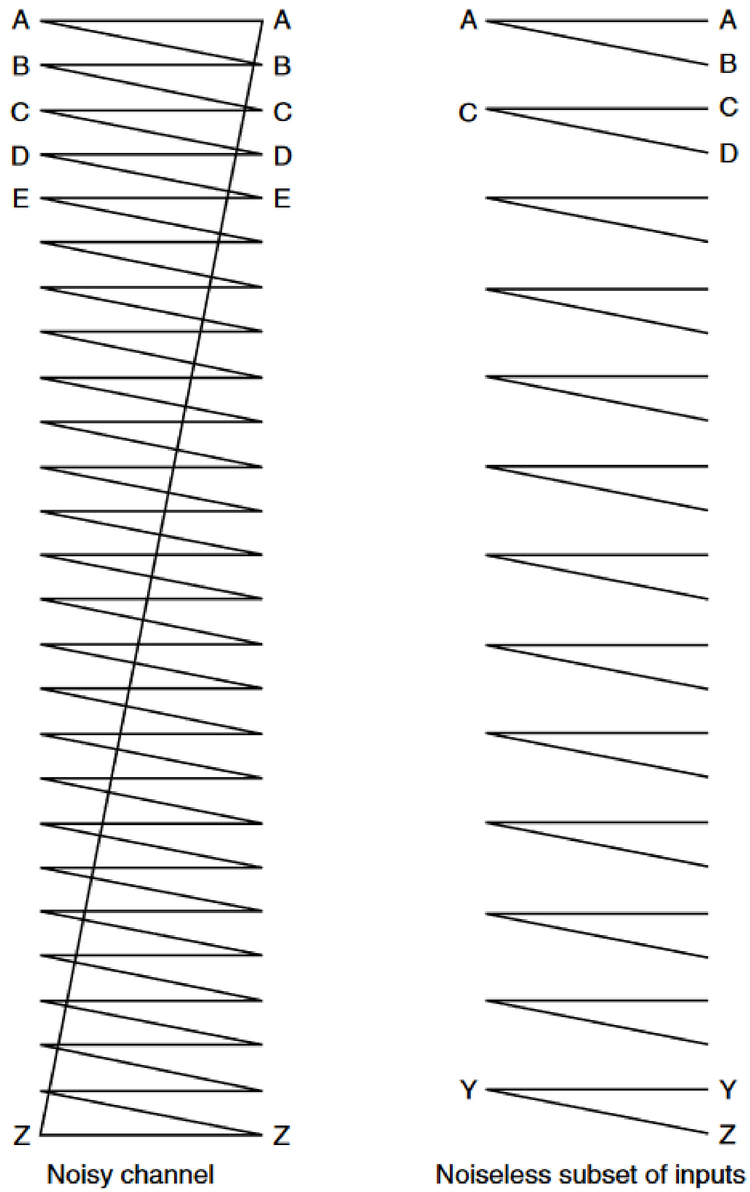

图 1: 打字机信道和它的一个错误概率为 0 的码. 图来自于 [2].

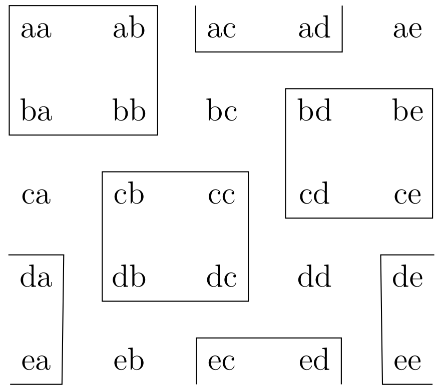

图 2: $\mathcal{X}^2 = \mathcal{Y}^2$ 中的 5 个不重叠的方块.

c) 证明: 我们可以把 $\mathcal{X}^2 = \mathcal{Y}^2$ 的所有元素排成图 2 所示的矩阵. $\mathcal{X}^2$ 中的字母对经信道传输可能不变, 也可能变成它在图 2 中右, 下或右下方的字母对. 这 4 个字母对组成一个 2×2 的方块. 这里认为矩阵是循环的, 如 ee 可能变成 ee, ea, ae 或 aa. 我们可以把图 2 中 5 个不重叠的 2×2 方块的左上角作为码字, 也就是用

$$f(1) = \mathrm{aa}, f(2) = \mathrm{cb}, f(3) = \mathrm{ec}, f(4) = \mathrm{bd}, f(5) = \mathrm{de}$$

定义 $f : \{1,2,3,4,5\} \to \mathcal{X}^2$. 这样存在 $g : \mathcal{Y}^2 \to \{1,2,3,4,5\}$ 使分组码 (f,g) 不会出错. (f,g) 的码率是 $\log_2(5)/2$, 所以 $\log_2(5)/2$ 是零错可达码率. □

第 10 题

我们可以将多个离散无记忆信道（DMC）组合成一个和信道（sum DMC）。将离散无记忆信道 $P_{Y|X}$ 写成矩阵形式，其第 i 行第 j 列元素为 $P_{Y|X}(j|i)$。给定输入字母表为 $\mathcal{X}_k$、输出字母表为 $\mathcal{Y}_k$ 的离散无记忆信道 $P_{Y|X,k}$（$k = 1, \ldots, K$），和信道 $P_{Y|X,sum}$ 的矩阵为 $P_{Y|X,sum} = \text{diag}\{P_{Y|X,1}, \ldots, P_{Y|X,K}\}$，即一个块对角矩阵，其对角子矩阵为 $P_{Y|X,1}, \ldots, P_{Y|X,K}$。计算 $P_{Y|X,sum}$ 的容量，并根据各组成离散无记忆信道的容量可达输入分布，描述和信道的容量可达输入分布。

We may combine several DMCs to create a sum DMC. Write a DMC $P_{Y|X}$ as a matrix whose i-th row j-th column element is $P_{Y|X}(j|i)$. Given DMCs $P_{Y|X,k}$ over input and output alphabets $\mathcal{X}_k$ and $\mathcal{Y}_k$, for $k = 1, \ldots, K$, the sum DMC has its matrix as $P_{Y|X,sum} = \operatorname{diag}\{P_{Y|X,1}, \ldots, P_{Y|X,K}\}$, i.e., a block-diagonal matrix whose diagonal submatrices $P_{Y|X,1}$, ..., $P_{Y|X,K}$. Calculate the capacity of $P_{Y|X,sum}$ and describe the capacity-achieving input distribution in terms of the capacity-achieving input distributions for the component DMCs.

解: 我们记结合后 X 的分布为

$$(\underbrace{p_X(x_{1,1}), p_X(x_{1,2}), ..., p_X(x_{1,n_1})}_{p_1}, ..., \underbrace{p_X(x_{k,1}), p_X(x_{k,2}), ..., p_X(x_{k,n_k})}_{p_K}),$$

n_i 为第 i 个子信道中的元素数量. 并记 $p_X(x_{i,j}) = p_i p_{X,i}(x_j), j = 1, 2, ..., n_i$, 右侧第一项即为选中第 i 个 DMC 中元素的概率, 第二项为指定该信道下取到对应元素的概率, 我们用

X_i 表示这个概率分布下的随机变量, 也就是第 i 个信道的信源分布. 在这些符号辅助下, 我们试着展开互信息的表达式:

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= H(p_1, p_2, ..., p_K) + \sum_{i=1}^{K} p_i H(p_{X,i}(x_1), ..., p_{X,i}(x_{n_i})) - H(X|Y) \\ &= H(p_1, p_2, ..., p_K) + \sum_{i=1}^{K} p_i H(X_i) - \sum_{i=1}^{K} p_i H(X_i|Y_i) \\ &= H(p_1, p_2, ..., p_K) + \sum_{i=1}^{K} p_i I(X_i;Y_i) \end{aligned} \tag{33}$$

此处唯一需要注意的是 (33) 中使用 $p(Y \in \mathcal{Y}_i)H(X|Y \in \mathcal{Y}_i) = p_i H(X_i|Y_i)$, 这是因为我们指定了 $y \in \mathcal{Y}_i$ 之后, 发现 Y_i 只能由 X_i 转移得到, $p(Y \in \mathcal{Y}_i) = p_i$, 且:

$$\begin{aligned} H(X|Y \in \mathcal{Y}_i) &= \sum_{j=1}^{m_i} p(Y = y_{i,j}|Y \in \mathcal{Y}_i) H(X|Y = y_{i,j}, Y \in \mathcal{Y}_i) \\ &= \sum_{j=1}^{m_i} p_{Y,i}(y_j) H(X|Y_j = y_j) \\ &= \sum_{j=1}^{m_i} p_{Y,i}(y_j) H(X_i|Y_j = y_j) = H(X_i|Y_i) \end{aligned}$$

从而对于原先的优化问题, 我们表示如下:

$$\begin{aligned} \max_P I(X;Y) &= \max_P \max_{P_{X,i}} H(p_1, p_2, ..., p_K) + \sum_{i=1}^{K} p_i I(X_i;Y_i) \\ &= \max_P H(p_1, p_2, ..., p_K) + \sum_{i=1}^{K} p_i C_i \\ &= \max_P \sum_{i=1}^{K} p_i (C_i - \log_2(p_i)), \qquad \text{s.t.} \sum_{i=1}^{K} p_i = 1. \end{aligned}$$

对这个问题的求解, 我们使用 Lagrange 乘子法:

$$\begin{aligned} J(P) &= \sum_{i=1}^{K} p_i (C_i - \log_2(p_i)) + \lambda p_i, \\ J'(p_i) &= C_i - \log_2(p_i) - 1 - \lambda = 0, \end{aligned}$$

$$\Rightarrow p_i = 2^{\beta + C_i}, \text{ which } 2^{\beta} = \frac{1}{\sum_i 2^{C_i}}. \text{ 整理可得:} C = \log_2(\sum_{i=1}^{k} 2^{C_i}).$$

例：$P_{Y|X,1}=\begin{bmatrix}1&0\\0&1\end{bmatrix}$,$P_{Y|X,2}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$, $C_1=1, C_2=\log_2(3)$, 带入结果可知信道容量为 $\log_2(5)$. 也可以解出 $(p_1,p_2)=(\frac{2}{5},\frac{3}{5})$, 进一步根据各个信道本身达到信道容量的分布, 得知 $P_{Y|X,sum}=(\frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5})$, 同样可以检验上述结果. □

考虑联合信源信道编码的未编码传输方案。
a) 验证在例6.3中，恒等映射能达到最优性能。
b) 在例6.3中，将离散无记忆信道（DMC）改为二元删除信道（BEC）。我们仍令 f 为恒等映射，并将 g 修改为：//
计算由此可达到的性能，并将其与信源 - 信道分离所达到的最优性能进行比较。

第 11 题

Consider the uncoded transmission scheme for joint source channel coding.

a) Verify that the identity mappings achieve the optimal performance in Example 6.3

b) In Example 6.3, change the DMC to a BEC. We still let f be an identity mapping, and modify g into

$$g(y)=y \text{ if } y\neq e,$$
$$\text{and} = U\sim Bernoulli(1/2) \text{ otherwise.}$$

Calculate the resulting achievable performance and compare it with the optimal performance achieved by source-channel separation.

a) 证明: 联合信源信道编码 (JSCC) 方案的平均失真为：

$$D_{JSCC}=\sum_{(s,\hat{s})}p(s,\hat{s})d(s,\hat{s})=\frac{1}{2}\cdot\delta+\frac{1}{2}\cdot\delta=\delta$$

当考虑分离信源信道编码 (SSCC) 方案时，参数为 $\frac{1}{2}$ 的 Bernoulli 信源的率失真函数为：$R(D_{SSCC})=h_2(\frac{1}{2})-h_2(D_{SSCC})$, 参数为 δ 的 BSC 信道容量为：$C=1-h_2(\delta)$. 根据性能界 $R(D_{SSCC})\leq C$, 则有 $D_{SSCC}\geq\delta$, 所以 JSCC 方案达到了最优性能. □

b) 解: 信源分布仍为参数为 $\frac{1}{2}$ 的 Bernoulli 分布, 设 BEC 信道参数为 α，则 JSCC 方案的平均失真为：

$$D_{JSCC}=\sum_{(s,\hat{s})}p(s,\hat{s})d(s,\hat{s})=\frac{1}{2}\cdot\alpha\cdot\frac{1}{2}+\frac{1}{2}\cdot\alpha\cdot\frac{1}{2}=\frac{\alpha}{2}$$

考虑分离信源信道编码 (SSCC) 方案, 信源的率失真函数为:$R(D_{SSCC})=h_2(\frac{1}{2})-h_2(D_{SSCC})$, BEC 信道容量为：$C=1-\alpha$. 根据性能界 $R(D_{SSCC})\leq C$, 则 $h_2(D_{SSCC})\geq\alpha$. □

参考文献

[1] S. P. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, New Jersey: John Wiley & Sons, Ltd, 2006.