# 信息论第五次小测解答

中国科学技术大学《信息论 A》006125.01 班助教组

2024 年 5 月 10 日

## 第 1 题

*Prove Corollary 6.1.*

$$C(\Gamma) = \max_{P_X} I(X;Y), \tag{1}$$

$$s.t.\ \mathbf{E}[c(X)] \le \Gamma. \tag{2}$$

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

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

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

*d):If $C(\infty) > 0$, then $C(\Gamma)$ 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：

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

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 $C(\Gamma) = 0$.

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：

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

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

c): 对于任意的 $\Gamma_{\min} \le \Gamma_1 < \Gamma_2$, 记满足 (2) 的分布集合为 $\mathcal{P}_\Gamma$, 那么满足 $\Gamma_1$ 约束时一定满足 $\Gamma_2$ 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.

我们希望进一步证明代价容量函数是凹 (concave) 的, 即对于 $\Gamma_{\min} \le \Gamma_1 < \Gamma_2$ 以及 $0 \le \lambda \le 1$ 有：

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

为此, 我们取在 $\Gamma_1$ 和 $\Gamma_2$ 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：

$$\begin{aligned}
\mathbf{E}_{X\sim P_{\lambda,X}}[c(X)] &= \sum_{x\in\mathcal{X}} p_\lambda(x)c(x) \\
&= (1-\lambda)\sum_{x\in\mathcal{X}} p_1(x)c(x) + \lambda \sum_{x\in\mathcal{X}} p_2(x)c(x) \\
&\le (1-\lambda)\Gamma_1 + \lambda\Gamma_2,
\end{aligned}$$

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：

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

进一步, 藉由互信息关于 $P_X$ 的凹性 (concave), 我们又有：

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

由 (3)(4)：

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

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 $C(\Gamma_1) = C(\Gamma_2) = C(\Gamma)$, 其中 $\Gamma_1 \le \Gamma \le \Gamma_2$, 由我们上面证明的凹性：

$$\Gamma_2 = \frac{\Gamma_{max} - \Gamma_2}{\Gamma_{max} - \Gamma_1}\Gamma_1 + \frac{\Gamma_2 - \Gamma_1}{\Gamma_{max} - \Gamma_1}\Gamma_{\max}$$
$$C(\Gamma_2) \ge \frac{\Gamma_{max} - \Gamma_2}{\Gamma_{max} - \Gamma_1}C(\Gamma_1) + \frac{\Gamma_2 - \Gamma_1}{\Gamma_{max} - \Gamma_1}C(\Gamma_{\max})$$

但是我们又假定 $C(\Gamma_1) = C(\Gamma_2)$, 那么只有 $C(\Gamma_1) = C(\Gamma_2) = C(\Gamma_{\max})$, 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 (2) 约束下 (1) 取到极值时, 有 $\mathbf{E}[c(X)] = \Gamma' < \Gamma_1$, 那么我们可以断言 $C(\Gamma') \ge C(\Gamma_1)$, 但是这又和严格递增矛盾, 命题得证. $\square$