第 11 讲 信道编码定理

联合典型序列

信道编码定理

零误差码

联合典型序列

信道编码定理

零误差码

定义 1.1

服从分布 $p(x,y)$ 的联合典型序列$\{\{x^n,y^n\}\}$ 所构成的集合 $A_\epsilon^{(n)}$ 是指其经验熵与真实熵 ϵ 接近的 n 长序列构成的集合，即：

$$\begin{aligned} A_\epsilon^{(n)} = \quad & \{(x^n,y^n)\in\mathcal{X}^n\times\mathcal{Y}^n: \\ & |-\tfrac{1}{n}\log p(x^n)-H(X)|<\epsilon, \\ & |-\tfrac{1}{n}\log p(y^n)-H(Y)|<\epsilon, \\ & |-\tfrac{1}{n}\log p(x^n,y^n)-H(X,Y)|<\epsilon\} \end{aligned}$$

其中

$$p(x^n,y^n)=\prod_{i=1}^{n}p(x_i,y_i).$$

定理 1.2: 联合 AEP

设 (X^n, Y^n) 为服从 $p(x^n, y^n) = \prod_{i=1}^n p(x_i, y_i)$ 的 i.i.d. 的 n 长序列，那么:

- 对任意 $\epsilon > 0$，当 n 充分大时，有 $P((X^n, Y^n) \in A_\epsilon^{(n)}) > 1 - \epsilon$.
- $|A_\epsilon^{(n)}| \leq 2^{n(H(X,Y)+\epsilon)}$.
- 如果 $(\tilde{X}^n, \tilde{Y}^n) \sim p(x^n)p(y^n)$，即 $\tilde{X}^n$ 与 $\tilde{Y}^n$ 是独立的且与 $p(x^n, y^n)$ 有相同的边际分布，那么

$$P\{(\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}) \leq 2^{-n(I(X;Y)-3\epsilon)}.$$

而且，对于充分大的 n，

$$P\{(\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}) \geq (1-\epsilon)2^{-n(I(X;Y)+3\epsilon)}.$$

证明.

1. 首先我们来证明，包含在典型集中的序列具有很高的概率. 由弱大数定律，

$$-\frac{1}{n}\log P(X^m) \to -E[\log p(X)] = H(X) \text{ 依概率.}$$

因此，给定 $\epsilon > 0$，存在 n_1，使得对于任意 $n > n_1$，使得对于任意 $n > n_1$，

$$P(|-\frac{1}{n}\log p(X^n) - H(X)| \geq \epsilon) < \frac{\epsilon}{3}.$$

类似地，由弱大数定律，我们有

$$-\frac{1}{n}\log p(Y^n) \to -E[\log p(Y)] = H(Y) \text{ 依概率,}$$

以及

$$-\frac{1}{n}\log p(X^n, Y^n) \to -E[\log p(X, Y)] = H(X, Y) \text{ 依概率.}$$

证明.

从而，存在 n_2 和 n_3，使得对任意 $n \geq n_2$，

$$P(|-\frac{1}{n}\log p(Y^n) - H(Y)| \geq \epsilon) < \frac{\epsilon}{3}.$$

以及对任意 $n \geq n_3$，

$$P(|-\frac{1}{n}\log p(X^n, Y^n) - H(X,Y)| \geq \epsilon) < \frac{\epsilon}{3}.$$

选取 $n > \max(n_1, n_2, n_3)$，则我们知

$$\begin{aligned} P((X^n, Y^n) \in A_\epsilon^{(n)}) \geq & 1 - P(|-\frac{1}{n}\log p(X^n) - H(X)| \geq \epsilon | \geq \epsilon) \\ & - P(|-\frac{1}{n}\log p(Y^n) - H(Y)| \geq \epsilon | \geq \epsilon) \\ & - P(|-\frac{1}{n}\log p(X^n, Y^n) - H(X,Y)| \geq \epsilon | \geq \epsilon) \\ > & 1 - \epsilon \end{aligned}$$

从而定理的第一部分得证.

证明.

2. 为证明定理的第二部分，我们注意到

$$\begin{aligned} 1 &= \sum p(x^n, y^n) \\ &\geq \sum_{A_\epsilon^{(n)}} p(x^n, y^n) \\ &\geq |A_\epsilon^{(n)}| 2^{-n(H(X,Y)+\epsilon)}. \end{aligned}$$

因此，$|A_\epsilon^{(n)}| \leq 2^{n(H(X,Y)+\epsilon)}$.

证明.

3. 现在，如果 $\tilde{X}^n$ 和 $\tilde{Y}^n$ 相互独立，但是与 X^n 和 Y^n 分别具有相同的边际分布，那么

$$\begin{aligned} P\{(\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}\} &= \sum_{(x^n, y^n) \in A_\epsilon^{(n)}} p(x^n) p(y^n) \\ &\leq 2^{n(H(X,Y)+\epsilon)} 2^{-n(H(X)-\epsilon)} 2^{-n(H(Y)-\epsilon)} \\ &= 2^{-n(I(X;Y)-3\epsilon)}. \end{aligned}$$

对充分大的 n，$P(A_\epsilon^{(n)}) \geq 1-\epsilon$，因此

$$\begin{aligned} 1-\epsilon &\leq \sum_{(x^n, y^n) \in A_\epsilon^{(n)}} p(x^n, y^n) \\ &\leq |A_\epsilon^{(n)}| 2^{-n(H(X,Y)-\epsilon)} \end{aligned}$$

证明.

以及

$$|A_\epsilon^{(n)}| \geq (1-\epsilon)2^{n(H(X,Y)-\epsilon)}.$$

和上界讨论的估计类似，我们可以证明，对于充分大的 n，有

$$\begin{aligned} P((\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}) &= \sum_{(x^n,y^n)\in A_\epsilon^{(n)}} p(x^n)p(y^n) \\ &\geq (1-\epsilon)2^{n(H(X,Y)-\epsilon)}2^{-n(H(X)+\epsilon)}2^{-n(H(Y)+\epsilon)} \\ &= (1-\epsilon)2^{-n(I(X;Y)+3\epsilon)}. \quad \square \end{aligned}$$

定理 2.1: 信道编码定理

对于离散无记忆信道，小于信道容量 C 的所有码率都是可达的. 具体来说，对任意码率 $R < C$，存在一个 $(2^{nR}, n)$ 码序列，它的最大误差概率为 $\lambda^{(n)} \to 0$.
反之，任何满足 $\lambda^{(n)} \to 0$ 的 $(2^{nR}, n)$ 码序列必定有 $R \leq C$.

证明.

我们在这一节证明可达性，逆定理的证明放在后面.

固定 $p(x)$，根据分布 $p(x)$ 随机生成一个 $(2^{nR}, n)$ 码. 具体来说，根据分布

$$p(x^n) = \prod_{i=1}^{n} p(x_i)$$

独立生成 2^{nR} 个码字，将它们作为信息 $\{1, \cdots, M\}$ 的编码. 将 2^{nR} 个码字展开为矩阵的行：

$$\mathcal{C} = \begin{pmatrix} x_1(1) & x_2(1) & \cdots & x_n(1) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(2^{nR}) & x_2(2^{nR}) & \cdots & x_n(2^{nR}) \end{pmatrix}$$

证明.

该矩阵的每一项都是依据 i.i.d. 服从 $p(x)$ 生成的，即 $P(X^n(w) = x^n) = \prod_{i=1}^n p(x_i)$ 且 $X^n(w)$ 是独立的，其中 $w \in \{1, \cdots, M\}$. 依次我们生成一个特定码 $\mathcal{C}$ 的概率就是

$$P(\mathcal{C}) = P(X^n(w) = x_1(w)x_2(w)\cdots x_n(w),\ w \in \{1, 2, \cdots, M\}) \tag{2.1}$$

$$= \prod_{w=1}^{2^{nR}} \prod_{i=1}^{n} p(x_i(w)). \tag{2.2}$$

证明.

我们考虑下面的系列事件：

1. 如2.1所示，按分布 $p(x)$ 生成一个码簿 $\mathcal{C}$.
2. 我们将码簿 $\mathcal{C}$ 告诉发送者和接收者，并且假定两者都知道该信道的信道转移矩阵 $p(y|x)$.
3. 依如下的均匀分布选取一条消息 W

$$P(W = w) = 2^{-nR},\ w = 1, 2, \cdots, 2^{nR}.$$

4. 第 w 个码字 $X^n(w)$ 是 $\mathcal{C}$ 的第 w 行，通过该信道被发送.

证明.

5. 接收者收到的序列 Y^n 服从分布

$$p(y^n|x^n(w)) = \prod_{i=1}^{n} p(y_i|x_i(w)).$$

6. 接收者猜测所发送的消息是什么. 如果满足下面两个条件，则接收者认为 $\hat{W}$ 就是所发送的下标.
   - $(X(\hat{W}), Y^n)$ 是联合典型的.
   - 不存在其他的下标 $W' \neq \hat{W}$ 满足 $(X^n(W'), Y^n) \in A_\epsilon^{(n)}$.

   如果这样的 $\hat{W}$ 不存在，或者有超过一个这样的 $\hat{W}$，则断言发生了错误 (在这种情况下，假定接收者给出一个哑下标，例如 0).
7. 如果 $\hat{W} \neq W$，则说明译码错误，我们记事件 $\{\hat{W}(Y^n) \neq W\}$ 为 $\mathcal{E}$.

证明.

下面我们来计算误差概率. 设 W 服从 $\{1,2,\cdots,2^{nR}$ 上的均匀分布，并用上面的译码方法得到 $\hat{W}(y^n)$. 设 $\mathcal{E}=\{\hat{W}(Y^n)\neq W\}$ 表示误差事件. 现在计算平均误差概率，也就是

$$\begin{aligned} P(\mathcal{E}) &= \sum_{\mathcal{C}} P(\mathcal{C}) P_e^{(n)}(\mathcal{C}) \\ &= \sum_{\mathcal{C}} P(\mathcal{C}) \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \lambda_w(\mathcal{C}) \\ &= \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \sum_{\mathcal{C}} P(\mathcal{C}) \lambda_w(\mathcal{C}). \end{aligned}$$

证明.

注意到对每个编码 $\mathcal{C}$，我们可以将 $\mathcal{C}$ 的第 1 行和第 i 行对换，得到一个新码字 $\mathcal{C}'$. 这时 $P(\mathcal{C}) = P(\mathcal{C}')$，而 $\lambda_1(\mathcal{C}) = \lambda_w(\mathcal{C}')$，对 $\mathcal{C}$ 求和，则 $\mathcal{C}'$ 也遍历所有码簿，从而

$$\sum_{\mathcal{C}} P(\mathcal{C})\lambda_1(\mathcal{C}) = \sum_{\mathcal{C}'} P(\mathcal{C}')\lambda_w(\mathcal{C}').$$

从而我们有

$$P(\mathcal{E}) = \sum_{\mathcal{C}} P(\mathcal{C})\lambda_1(\mathcal{C}) = P(\mathcal{E}|W = 1).$$

定义下列事件：

$$E_i = \{(X^n(i), Y^n)\text{在}A_\epsilon^{(n)}\text{中}\},\ i \in \{1, 2, \cdots, 2^{nR}\}.$$

E_i 表示第 i 个码字与 Y^n 为联合典型的这一事件.

证明.

如果 E_1^C 发生，或者 $E_2 \cup E_3 \cup \cdots \cup E_{2^n}$ 发生，则译码会发生错误. 反过来，若译码错误，则 $E_1^C, E_2, E_3, \cdots, E_{2^n}$ 中必有一个发生. 于是

$$\begin{aligned} P(\mathcal{E}) &= P(\mathcal{E}|W=1) \\ &= P(E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_{2^{nR}}|W=1) \\ &= P(E_1^c|W=1) + \sum_{i=2}^{2^{nR}} P(E_i|W=1). \end{aligned}$$

由联合 AEP 的性质，我们知当 n 充分大时，

$$P(E_1^c|W=1) \leq \epsilon.$$

证明.

从编码的过程我们知 $X^n(1)$ 和 $X^n(i)(i \neq 1)$ 是独立的，所以 Y^n 与 $X^n(i)$ 也是独立的. 因此，根据联合 AEP 的性质，$X^n(i)$ 与 Y^n 是联合典型的概率 $\leq 2^{-n(I(X;Y)-3\epsilon)}$. 从而，如果 n 充分大且 $R < I(X;Y) - 3\epsilon$ 时，

$$\begin{aligned} P(\mathcal{E}) &= P(\mathcal{E}|W=1) \leq P(E_1^c|W=1) + \sum_{i=2}^{2^{nR}} P(E_i|W=1) \\ &\leq \epsilon + \sum_{i=2}^{2^{nR}} 2^{-n(I(X;Y)-3\epsilon)} \\ &= \epsilon + (2^{nR}-1)2^{-n(I(X;Y)-3\epsilon)} \\ &\leq \epsilon + 2^{3n\epsilon}2^{-n(I(X;Y)-R)} \\ &\leq 2\epsilon. \end{aligned}$$

因此，如果 $R < I(X;Y)$，对于任意 $\epsilon > 0$，可以适当选取 n，使得平均误差概率不超过 2ϵ.

证明.

为完成证明，我们需要进行下面一系列操作：

- 我们将证明中的 $p(x)$ 换成 $p^*(x)$，即达到信道容量时关于 X 的分布. 此时，条件 $R < I(X;Y)$ 可以由可达性条件 $R < C$ 所替代.
- 去除码簿上的平均. 由于所有的码簿上的平均误差概率比较小，所以存在一个码簿 $\mathcal{C}^*$ 具有较小的平均误差概率. 于是 $P_e^{(n)}(\mathcal{E}|C^*) \leq 2\epsilon$. 同时我们注意到

$$P(\mathcal{E}|\mathcal{C}^*) = \frac{1}{2^{nR}} \sum_{i=1}^{2^{nR}} \lambda_i(\mathcal{C}^*).$$

- 抛弃最佳码簿 $\mathcal{C}^*$ 中最差的一半码字. 由于这个码字的算术平均误差概率 $P_e^n(\mathcal{C}^*)$ 小于 2ϵ，我们有

$$P(\mathcal{E}|\mathcal{C}^*) \leq \frac{1}{2^{nR}} \sum \lambda_i(\mathcal{C}^*) \leq 2\epsilon.$$

  这说明至少有一半的下标 i 及其对应的码字 $X^n(i)$ 的条件误差概率 λ_i 小于 4ϵ(否则，这些码字本身的和就将大于 2ϵ). 因此，所有码字中最佳的一半的最大误差概率必定小于 4ϵ. 码字总数有 2^{nR-1} 个. 抛弃一半码字使得码率由 R 变为 $R - \frac{1}{n}$，当 n 充分大时，这是可以忽略的.

结合所有这些改进，我们构造了一个码率为 $R' = R - \frac{1}{n}$ 的码，它的最大误差概率 $\lambda^{(n)} \leq 4\epsilon$. 这就证明了任何小于信道容量的码率是可达的.

我们首先来证明 $P_e^{(n)} = 0$ 蕴含结论 $R \leq C$. 假定有一个零误差概率的 $(2^{nR}, n)$ 码，也就是说，译码器输出的 $g(Y^n)$ 依概率 1 等于输入的下标 W. 那么，输入下标 W 完全由输出序列决定 (即 $H(W|Y^n) = 0$). 为了获得更强的界，随意假定 W 服从 $\{1, 2, \cdots, 2^{nR}\}$ 上的均匀分布，于是，$H(W) = nR$. 从而我们有如下的一串不等式：

$$\begin{aligned} nR = H(W) &= H(W|Y^n) + I(W; Y^n) \\ &= I(W; Y^n) \\ &\leq I(X^n, Y^n) \\ &\leq \sum_{i=1}^{n} I(X_i; Y_i) \\ &\leq nC. \end{aligned}$$

其中第一个不等式是由数据处理不等式可得，第二个不等式和第三个不等式我们将在下一讲中给出证明. 因此，对于任何零误差的 $(2^{nR}, n)$ 码以及所有的 n，

$$R \leq C.$$