数 学 年 刊
1 (2) 1980

# 一般离散序列模型信源编码定理

沈 世 镒
(南开大学)

## 1. 引　言

信源编码问题是Shannon信息论中的一个重要问题，近几年来，对这个问题有一系列重要结果，无论在理论上与实际应用上都有一定的意义．对该问题的确切提法与若干重要结果已由C. E. Shannon[1,2]，R. G. Gallager[3]，T. Berger[4]文阐明，与信道编码问题相似，信源编码问题反映了信源内在的结构特征，它的基础问题就为信源编码定理，近几年许多工作均在有记忆信源问题上进行探讨．

六十年代初，由胡国定[5,6]文引述的信源、信道离散序列模型与结果，使我们对信息传输中的若干特征问题有进一步的了解，因此我们在研究信源编码定理时，采用序列模型的描述办法，同样可以帮助我们认识该问题中的若干基本特征的意义与联系．本文在一般的情形下获得了信源编码定理的充要条件，作为本文的特殊情形，就是为[2,3,4,8,9]等文中关于平稳、无记忆、遍历等信源的编码定理．由本文的结果可直接确立有关非平稳、有记忆信源的模型与编码定理．本文对度量测度的要求也是十分宽广的．因此本文的结果在一定程度上刻划了信源编码定理的特征要求．

本文在写作过程中曾与胡国定先生、章照止同志研究商讨过，部份内容在他们的意见下写出特此表示感谢．

## 2. 序列模型与记号

关于信源、信道序列模型同[5]文所述，记$[X^{(n)}, p^{(n)}(x^{(n)})]$ $n=1, 2, \cdots$，为一信源序列，$X^{(n)}$为一有限集，$p^{(n)}(x^{(n)})$为$X^{(n)}$上概率分布．为了简单起见，在2、3、5节中，如无特别声明，各字母右上角均有$(n)$，但我们都省略不写．另外引进记号为：

**(1) 度量测度**

记$Y$为一列有限集合，为复制消息，$\rho(x, y)$为$X\otimes Y$上的函数，称之为度量函数，记

$$\rho(x, B)=\min_{y\in B}\rho(x, y), \ B\subset Y, \tag{2.1}$$

$$H(B, d)=\{x: x\in X, \ \rho(x, B)\leqslant d\}. \tag{2.2}$$

$$p(R, d)=\min_{|B|<2^R} p(H(B, d)). \tag{2.3}$$

本文1978年10月24日收到．

其中 $R$, $d$ 为确定数列, $|B|$ 为 $B$ 中元的个数.

**(2) 条件分布序列**

记 $Q(y/x)$ 为 $Y$ 关于 $X$ 的条件概率, 记

$$p(x, y)=p(x)Q(y/x),\ Q(y)=\sum_x p(x, y)$$

$$p(x/y)=\frac{p(x, y)}{Q(y)}, \tag{2.4}$$

$p(x)$ 为信源序列, 如信源序列给定, 即 $p(x)$ 确定, 则公息

$$I(Q)=\sum_{x, y} p(x)Q(y/x)\log\frac{Q(y/x)}{Q(y)}, \tag{2.5}$$

由 $Q(y/x)$ 确定而确定. 记 $i(x, y)=\log\dfrac{Q(y/x)}{Q(y)}$.

**(3) 信源序列的 $(R, d)$——压缩与 $(\mathscr{F}, d)$——信息有界**

对上述 $[X, p(x)]$, $Y$, $\rho(x, y)$ 我们总记为 $\mathscr{S}$, 也称之为信源序列. 对数列向量 $(R, d)$, 如

$$\lim_{n\to\infty} p(R, d)=0 \tag{2.6}$$

成立, 则称 $\mathscr{S}$ 为 $(R, d)$——压缩. 也就是存在一列 $B\subset Y$, $|B|\leqslant 2^R$, 使

$$\lim_{n\to\infty} p\{x: x\in X,\ \rho(x, B)>d\}=0, \tag{2.7}$$

成立. 信源编码问题就为讨论 $(R, d)$——压缩的等价条件. 以下定义 $(\mathscr{F}, d)$——信息有界:

对上述 $\mathscr{S}$, 如存在一条件分布 $Q(y/x)$, 对联合分布 $p(x, y)$ 同时有

$$p\{(x, y): i(x, y)\leqslant\mathscr{F}(1+\varepsilon)\}>1-\varepsilon, \tag{2.8}$$

$$p\{(x, y): \rho(x, y)\leqslant d\}>1-\varepsilon, \tag{2.9}$$

成立. 其中 $\varepsilon\to 0\ (\varepsilon>0)$, 则称 $\mathscr{S}$ 为 $(\mathscr{F}, d)$——信息有界. 本文极限均为当 $n\to\infty$ 时极限.

**(4) 信号在一定失真许可条件下的信号传输问题**

关于信源序列 $\mathscr{S}$ 的 $(R, d)$——压缩概念在信号传输中的作用问题. 在[1]文中已经阐明, [5]文在序列模型下作了严格叙述, 有关记号与结论为:

1) 通讯系统

一个通讯系统指一个信源序列 $\mathscr{S}$ 与一个信道序列 $\mathscr{C}=[U, p(v/u), V]$, 记之为 $\mathscr{E}$, 因此

$$\mathscr{E}=\{\mathscr{S}, \mathscr{C}\}=\{[X, p(x)],\ [U, p(v/u), V],\ Y,\ \rho(x, y)\} \tag{2.10}$$

$Y$ 在这里为接收消息 ($X$ 为拍发消息, $U$、$V$ 分别为拍发、接收信号).

记 $f$, $g$ 为 $X\to U$, $V\to Y$ 的单值映像, 分别称之为翻码、译码, $(f, g)$ 给定的 $\mathscr{E}$ 记为 $\mathscr{E}(f, g)$, 称之为编码确定的通讯系统, 对 $\mathscr{E}(f, g)$ 知 $X\otimes U\otimes V\otimes Y$ 上的联合分布确定, 而

$$\begin{aligned} &p(x, y)=p(x)\cdot p(g^{-1}(y)/f(x)),\\ &g^{-1}(y)=\{v: g(v)=y\} \end{aligned} \tag{2.11}$$

2) 可靠性准则

关于通讯系统的可靠性准则形式很多, 如概率准则、信息准则、度量准则, 对度量准则因度量函数不同, 形式也很多, 我们在此讨论二种情形:

(a)
$$\sum_{xy} p(x, y)\rho(x, y) \leqslant d,$$

(b)
$$p\{(x, y): \rho(x, y) \leqslant d\} \to 1.$$

$(a)$为平均失真不超过 $d$, $(b)$ 为大部分消息失真不超过 $d$. 这二个可靠性准则在一定条件下可以互推.

3) 通讯系统的 Shannon 问题

这个问题系指，对已给的 $\mathscr{E}$ 问在什么条件下存在适当的$(f, g)$使$(a)$或$(b)$成立. 归纳[1, 5]文结论为:

对上述 $\mathscr{E}$, 如

(i) 信源序列 $\mathscr{S}$ 为$(R, d)$——压缩;

(ii) 信道序列 $\mathscr{C}$ 为 $\mathscr{I}$——信息稳定;

(iii)
$$\lim_{n\to\infty} \frac{\mathscr{T}}{R} > 1,$$

则存在$(f, g)$使$(b)$成立.

相应的逆命题也类似成立. 因此$(R, d)$——压缩在 Shanuon 问题中起重要作用，信源编码问题就是寻找$(R, d)$——压缩的等价条件与计算公式，这与信道编码定理中找 $\varepsilon$——专线数目 $N$ 与信道 $\mathscr{I}$——信息稳定与信道容量计算的过程十分相似.

## 3. 信源编码定理

**定理 1** (信源编码正定理)

对上述信源序列 $\mathscr{S}$, 如$(\mathscr{F}, d)$——信息有界，而且有

$$\lim_{n\to\infty} \frac{R}{\mathscr{F}} > 1$$

成立(以后简记为 $R > \mathscr{F}$), 则 $\mathscr{S}$ 为$(R, d)$——压缩.

证　因 $R > \mathscr{F}$, 则存在常数 $\varepsilon_0 > 0$, 当 $n$ 充分大时 $R > \mathscr{F}(1+\varepsilon_0)$ 成立，又因 $\mathscr{S}(\mathscr{F}, d)$——信息有界，则存在一条件分布序列 $Q(y/x)$, 使(2.8), (2.9)成立，我们适当放慢 $\varepsilon\to 0$ 的速度，使 $-\dfrac{1}{\mathscr{F}}\log\varepsilon \to 0$, 记

$$A = \{(x, y): i(x, y) \leqslant \mathscr{F}(1+\varepsilon)\}, \tag{3.1}$$

$$E = \{(x, y): \rho(x, y) \leqslant d\}, \tag{3.2}$$

作 $G = A \cap E$, 与

$$G_y = \{x: (x, y) \in G\}, \tag{3.3}$$

显然有

$$\sum_y Q(y) \cdot p(G_y/y) = \sum_y p(G_y \otimes y) = p(G) > 1 - 2\varepsilon, \tag{3.4}$$

成立. 这样就可选出 $G_1, \cdots, G_N$ 为

$$G_1 = G_{y_1} \quad \text{使 } p(G_1/y_1) > \varepsilon \text{ 成立,}$$

$$G_2 = G_{y_2} - G_1 \quad \text{使 } p(G_2/y_2) > \varepsilon \text{ 成立,}$$

$$\cdots\cdots$$

$$G_N = G_{y_N} - \sum_{i=1}^{N-1} G_i \quad \text{使 } p(G_N/y_N) > \varepsilon \text{ 成立,}$$

使 $N$ 为最大数，也就是对任何 $y$ 有

$$p\left(G_y-\sum_{i=1}^{N}G_i/y\right)<\varepsilon$$

成立，这样就有，对任何 $y$ 有

$$p(G_y/y)\leqslant p\left(\sum_{i=1}^{N}G_i/y\right)+p\left(G_y-\sum_{i=1}^{N}G_i/y\right)\leqslant p\left(\sum_{i=1}^{N}G_i/y\right)+\varepsilon$$

成立．两边乘 $Q(y)$ 相加得

$$\sum_{y}Q(y)p(G_y/y)\leqslant p\left(\sum_{i=1}^{N}G_i\right)+\varepsilon \tag{3.5}$$

把(3.4)代入左式得

$$p\left(\sum_{i=1}^{N}G_i\right)>1-3\varepsilon \tag{3.6}$$

成立，另一方面，对任何 $x\in G_y$ 有

$$p(x/y)\leqslant 2^{\mathscr{F}(1+\varepsilon)}p(x),$$

因此有

$$\varepsilon<p(G_i/y_i)\leqslant 2^{\mathscr{F}(1+\varepsilon)}p(G_i) \tag{3.7}$$

成立，得

$$p(G_i)>\varepsilon\cdot 2^{-\mathscr{F}(1+\varepsilon)},$$

$$p\left(\sum_{i=1}^{N}G_i\right)>\varepsilon N\cdot 2^{-\mathscr{F}(1+\varepsilon)}$$

成立，因 $\dfrac{1}{\mathscr{F}}\log\dfrac{1}{\varepsilon}\to 0$，对上述 $\varepsilon_0>0$，当 $n$ 充分大时，有

$$N\leqslant\frac{1}{\varepsilon}2^{\mathscr{F}(1+\varepsilon)}\leqslant 2^{\mathscr{F}(1+\varepsilon_0)}$$

成立．现在取

$$B=\{y_1,\ y_2,\ \cdots,\ y_N\},$$

因 $R\geqslant\mathscr{F}(1+\varepsilon_0)$，因此当 $n$ 充分大时，有

$$N\leqslant 2^{\mathscr{F}(1+\varepsilon_0)}\leqslant 2^R$$

成立．另一方面，因 $G_i\otimes y_i\subset G\subset E$，因此

$$\{x:\rho(x,\ B)\leqslant d\}\supset\sum_{i=1}^{N}G_i$$

$$p\{x:\rho(x,\ B)>d\}\leqslant 1-p\left(\sum_{i=1}^{N}G_i\right)\leqslant 3\varepsilon,$$

则有
$$p(R,\ d)\leqslant p\{x:\rho(x,\ B)>d\}\leqslant 3\varepsilon\to 0,$$

成立此即 $\mathscr{S}$ 为 $(R,\ d)$——压缩．

证毕．

由此可见，信源编码定理与信道编码定理的证明过程也十分对偶．下面讨论信源编码的反定理．

**引理 1** 任一信源序列 $[X,\ p(x)]$，如 $|X|<2^R$，$R\to\infty$，则对任何常数 $\varepsilon_0>0$，有

$$p\{x:-\log p(x)\leqslant R(1+\varepsilon_0)\}\to 1 \tag{3.8}$$

成立

证 如(3.8)不成立，则存在二个常数 $\varepsilon_0>0$，$\alpha_0>0$，与一个子列 $n_i$，$i=1,\ 2,\ \cdots$，$n_{i+1}>n_i$，而且有

$$p^{(n_i)}\{x^{(n_i)}: -\log p^{(n_i)}(x^{(n_i)})>R(1+\varepsilon_0)\}>\alpha_0$$

成立. 如记大括号中集合为 $A^{(n_i)}$, 如 $x^{(n_i)}\in A^{(n_i)}$, 就有

$$p^{(n_i)}(x^{(n_i)})\leqslant 2^{-R(n_i)(1+\varepsilon_0)}$$

成立, 得

$$\alpha_0<p^{(n_i)}(A^{(n_i)})\leqslant |A^{(n_i)}|2^{-R(n_i)(1+\varepsilon_0)}$$

$$|A^{(n_i)}|>\alpha_0\cdot 2^{R(n_i)(1+\varepsilon_0)}$$

因 $R\to\infty$, $A^{(n_i)}$ 为 $X^{(n_i)}$ 的子集, 则与 $|X|\leqslant 2^R$ 矛盾.

证毕.

**定理 2** (信源编码反定理)

对上述信源序列 $\mathscr{S}$, 如 $(R, d)$——压缩, 则对任何 $\mathscr{F}>R$ 的 $\mathscr{F}$ 有 $\mathscr{S}$ 为 $(\mathscr{F}, d)$——信息有界.

证　因 $\mathscr{S}(R, d)$——压缩, 则在 $Y$ 上相应地有一列子集 $B$, $|B|\leqslant 2^R$, 而且

$$p\{x:\rho(x, B)>d\}\to 0 \tag{3.9}$$

成立, 记 $B$ 的元为 $\{y_1, \cdots, y_N\}$, $N<2^R$, 记

$$G_i=\{x:\rho(x, y_i)\leqslant d\}, \tag{3.10}$$

$$A_i=G_i-\sum_{k=1}^{i-1}G_k \tag{3.11}$$

因此 $A_i$ 互不相交, 作 $Q(y/x)$ 为

$$Q(y/x)=\begin{cases}1 & \text{如 } x\in A_i,\ y=y_i\\ 0 & \text{如 } x\in A_i,\ y\neq y_i\end{cases} \tag{3.12}$$

其余的 $x$ 任取一 $y_i$, 使 $Q(y_i/x)=1$, 其余 $y$ 为 0, 则有

$$\{(x, y):\rho(x, y)\leqslant d\}\supset\sum_{i=1}^{N}(A_i\otimes y_i),$$

其中 $A_i\otimes y_i=\{(x, y_i), x\in A_i\}$, 则

$$p\{(x, y):\rho(x, y)\leqslant d\}\geqslant p\left(\sum_{i=1}^{N}A_i\otimes y_i\right)=p\left\{x:x\in\sum_{i=1}^{N}A_i\right\}$$

$$=p\{x:\rho(x, B)\leqslant d\}$$

由 (3.9) 知右边趋向于 1, 因此

$$p\{(x, y):\rho(x, y)\leqslant d\}\to 1 \tag{3.13}$$

成立. 另一方面, 记

$$E_i=\{x:Q(y_i/x)=1\} \tag{3.14}$$

则有

$$p(x, y)=p(x)Q(y/x)=\begin{cases}p(x), & x\in E_i,\ y=y_i,\\ 0 & \text{其它情况},\end{cases}$$

因此当 $x\in E_i$ 时

$$i(x, y_i)=\log\frac{p(x, y_i)}{p(x)Q(y_i)}=-\log Q(y_i)$$

成立. 如 $y\in B$, $Q(y)=0$, 则由引理 1 得

$$Q\{y:-\log Q(y)\leqslant R(1+\varepsilon_0)\}\to 1$$

成立,

因 $\varepsilon_0$ 为任何大于零的常数, 而 $\mathscr{F}>R$, 则得

$$p\{(x, y): i(x, y)\leqslant \mathscr{F}\}=Q\{y: -\log Q(y)\leqslant \mathscr{F}\}\to 1, \tag{3.15}$$

成立. 由(3.13), (3.15)即知 $\mathscr{S}$ 为$(\mathscr{F}, d)$——信息有界.

证毕.

这样我们证得了一般信源序列的正、反编码定理, 得到 $\mathscr{S}$ 的$(R, d)$——压缩与$(\mathscr{F}, d)$——信息有界的等价关系.

## 4. 若干结果的比较

### (1) 平稳无记忆信源正、反编码定理

现在我们说明本文结果在特殊情形下与[2, 3, 4, 7, 8]等结果一致. 为了简单起见, 我们重点讨论无记忆情形.

无记忆信源描述如[2, 3, 4]文, 设 $X$, $Y$ 为两有限集合, 这时

$$\left.\begin{aligned} X^{(n)}&=\prod_{i=1}^{n}\otimes X_i,\ X_i=X,\ i=1, \cdots, n;\\ Y^{(n)}&=\prod_{i=1}^{n}\otimes Y_i,\ Y_i=Y,\quad i=1, \cdots, n;\end{aligned}\right\} \tag{4.1}$$

相应的元分别为

$$\left.\begin{aligned} x^{(n)}&=(x_1, \cdots, x_n),\ x_i\in X\quad i=1, \cdots, n;\\ y^{(n)}&=(y_1, \cdots, y_n),\ y_i\in Y,\quad i=1, \cdots, n;\end{aligned}\right\} \tag{4.2}$$

记 $p(x)$ 为 $X$ 上概率分布, 平稳无记忆信源序列为

$$p^{(n)}(x^{(n)})=\prod_{i=1}^{n}p(x_i),\ x^{(n)}=(x_1, \cdots, x_n); \tag{4.3}$$

记 $Q(y/x)$ 为 $Y$ 对 $X$ 的条件概率, 记

$$Q_0^{(n)}(y^{(n)}/x^{(n)})=\prod_{i=1}^{n}Q(y_i/x_i), \tag{4.4}$$

其中 $x^{(n)}=(x_1, \cdots, x_n)$, $y^{(n)}=(y_1, \cdots, y_n)$; 如 $X\otimes Y$ 上度量函数为 $\rho(x, y)$, 则 $X^{(n)}\otimes Y^{(n)}$ 上度量函数为

$$\rho^{(n)}(x^{(n)}, y^{(n)})=\frac{1}{n}\sum_{i=1}^{n}\rho(x_i, y_i) \tag{4.5}$$

$x^{(n)}$, $y^{(n)}$ 意义同上.

**引理2** 在上述各记号下, 如有

$$\sum_{xy}p(x, y)\rho(x, y)<d \tag{4.6}$$

成立, 则一定有

$$p_0^{(n)}\{(x^{(n)}, y^{(n)}),\ \rho^{(n)}(x^{(n)}, y^{(n)})>d\}\to 0, \tag{4.7}$$

成立. 反之, 如(4.7)成立, 则(4.6)中 $<$ 号改为 $\leqslant$ 号即成立.

上述

$$p(x, y)=p(x)Q(y/x)$$

$$p_0^{(n)}(x^{(n)}, y^{(n)})=p^{(n)}(x^{(n)})Q_0^{(n)}(y^{(n)}/x^{(n)}), \tag{4.8}$$

证 由(4.8)知

$$p_0^{(n)}(x^{(n)}, y^{(n)})=\prod_{i=1}^{n}p(x_i, y_i), \tag{4.9}$$

为 $(X\otimes Y)^{(n)}$ 上的乘积分布，而 $\rho^{(n)}(x^{(n)}, y^{(n)})$ 为独立随机变数的和. 由大数定律即得引理成立.

定义平稳、无记忆信源序列率失真函数为

$$R(d)=\min_{Q\in Q_d} I(Q), \tag{4.10}$$

$I(Q)$ 由 (2.5) 定义，而

$$Q_d=\{Q(y/x):\sum_{xy} p(x)Q(y/x)\rho(x, y)\leqslant d\}. \tag{4.11}$$

我们记上述平稳、无记忆信源序列为 $\mathscr{S}_0^{(n)}$，其中 $Y^{(n)}$, $\rho^{(n)}$ 如 (4.1), (4.5) 所述.

**定理 3** 对上述 $\mathscr{S}^{(n)}$，如 $R>R(d)$，则 $\mathscr{S}^{(n)}(nR, d)$——压缩. 反之，如 $\mathscr{S}^{(n)}(nR, d)$——压缩，则 $R\geqslant R(d)$.

证 如 $R>R(d)$，由 $R(d)$ 的连续性(见[4])知存在 $\varepsilon_0>0$，使 $R>R(d-\varepsilon_0)(1+\varepsilon_0)^2$ 成立.

由 $R(d-\varepsilon_0)$ 定义知，有一 $Q(y/x)\in Q_{d-\varepsilon_0}$ 使

$$I(Q)<R(d-\varepsilon_0)(1+\varepsilon_0), \tag{4.12}$$

成立，对这个 $Q(y/x)$ 相应的 $p^{(n)}(x^{(n)}, y^{(n)})$ 有

$$p^{(n)}\{(x^{(n)}, y^{(n)}):\rho^{(n)}(x^{(n)}, y^{(n)})>d\}\to 0$$

$$p^{(n)}\left\{(x^{(n)}, y^{(n)}):\left|\frac{1}{n}i^{(n)}(x^{(n)}, y^{(n)})-I(Q)\right|>\varepsilon_0\right\}\to 0$$

成立，此即 $\mathscr{S}^{(n)}$ 为 $(nI(Q)(1+\varepsilon_0), d)$——信息有界，因

$$R>R(d-\varepsilon_0)(1+\varepsilon_0)^2>I(Q)(1+\varepsilon_0).$$

因此 $\mathscr{S}^{(n)}$ 为 $(nR, d)$——压缩，由定理 1 即可推出. 证得正定理.

对逆命题，如 $\mathscr{S}_0^{(n)}(nR, d)$——压缩，则对任何 $\varepsilon_0>0$，由定理 2 知 $\mathscr{S}_0^{(n)}(n(R+\varepsilon_0), d)$——信息有界，这时有一列 $Q^{(n)}(y^{(n)}/x^{(n)})$ 与 $\varepsilon^{(n)}\to 0\,(\varepsilon^{(n)}>0)$ 使

$$p^{(n)}\{(x^{(n)}, y^{(n)}), \rho^{(n)}(x^{(n)}, y^{(n)})\leqslant d\}>1-\varepsilon^{(n)} \tag{4.14}$$

成立. 记 $p^{(n)}(x^{(n)}, y^{(n)})$ 在 $X_i\otimes Y_i$ 上的边际分布为 $p_i(x_i, y_i)$，记

$$Q_i(y_i/x_i)=\frac{p_i(x_i, y_i)}{p(x_i)},$$

记

$$Q_*^{(n)}(y^{(n)}/x^{(n)})=\prod_{i=1}^{n} Q_i(y_i/x_i) \tag{4.15}$$

$$p_*^{(n)}(x^{(n)}, y^{(n)})=p^{(n)}(x^{(n)})Q_*^{(n)}(y^{(n)}/x^{(n)})=\prod_{i=1}^{n} p_i(x_i, y_i),$$

这时有

$$E\{\rho^{(n)}(x^{(n)}, y^{(n)})\}=\frac{1}{n}\sum_{i=1}^{n} E_i\{\rho(x_iy_i)\}=E_*\{\rho^{(n)}(x^{(n)}, y^{(n)})\} \tag{4.16}$$

其中 $E$, $E_*$, $E_i$ 分别为分布 $p^{(n)}(x^{(n)}, y^{(n)})$, $p_*^{(n)}(x^{(n)}, y^{(n)})$, $p_i(x_i, y_i)$ 的数学期望，因为 $\rho(x, y)$ 只取有限多个值(有界)，由 (4.14) 即得，当 $n$ 充分大时，有

$$E\{\rho^{(n)}(x^{(n)}, y^{(n)})\}\leqslant d+\varepsilon_0 \tag{4.17}$$

成立，另一方面由定理 2 证明 $Q^{(n)}(y^{(n)}/x^{(n)})$ 的构造知

$$E\{i^{(n)}(x^{(n)}, y^{(n)})\}=E\{-\log Q^{(n)}(y^{(n)})\}\leqslant nR, \tag{4.18}$$

因 $p^{(n)}(x^{(n)})$ 为乘积分布,因此

$$I^{(n)}(Q_*^{(n)})\leqslant I^{(n)}(Q^{(n)})\leqslant nR \tag{4.19}$$

成立,如记 $E_i\{\rho(x_i, y_i)\}=d_i$ 则由(4.19)知

$$\frac{1}{n}[R(d_1)+\cdots+R(d_n)]\leqslant\frac{1}{n}I^{(n)}(Q_*^{(n)})\leqslant R \tag{4.20}$$

成立,由 $R(d)$ 函数的凸性即得

$$R\left(\frac{d_1+\cdots+d_n}{n}\right)\leqslant\frac{1}{n}[R(d_1)+\cdots R(d_n)],$$

又由 $R(d)$ 的单调下降性与(4.16),(4.17)知

$$R(d+\varepsilon_0)\leqslant R\left(\frac{d_1+\cdots+d_n}{n}\right)\leqslant R,$$

因 $\varepsilon_0$ 任意,因此 $R(d)\leqslant R$ 成立. 证毕.

关于(4.19)不等式系由以下情况得出,如 $\xi_1\cdots\xi_n$ 为独立随机变数,则对公息有不等式

$$I((\xi_1\cdots\xi_n);\ (\eta_1\cdots\eta_n))\geqslant\sum_{i=1}^{n}I(\xi_i,\ \eta_i)$$

成立. 因为这时

$$\begin{aligned}&I((\xi_1\cdots\xi_n),\ (\eta_1\cdots\eta_n))=H(\xi_1\cdots\xi_n)-H(\xi_1\cdots\xi_n/\eta_1\cdots\eta_n)\\&\geqslant H(\xi_1)+\cdots+H(\xi_n)-H(\xi_1/\eta_1\cdots\eta_n)-\cdots-H(\xi_n/\eta_1\cdots\eta_n)\\&\geqslant\sum_{i=1}^{n}I(\xi_i,\ \eta_i).\end{aligned}$$

这样就保证了本文结果与[2, 3, 4]各文结果一致.

**(2) 非平稳,无记忆信源编码定理**

非平稳,无记忆信源描述与1)相似,我们只要将(4.3)的 $p^{(n)}(x^{(n)})$ 改为

$$p^{(n)}(x^{(n)})=\prod_{i=1}^{n}p_i(x_i),\quad x^{(n)}=(x_1,\ \cdots,\ x_n) \tag{4.21}$$

而 $Q^{(n)}(y^{(n)}/x^{(n)})$, $p^{(n)}(x^{(n)},\ y^{(n)})$, $I(Q^{(n)})$ 定义如前,记 $n$ 维平均率失真函数为

$$R^{(n)}(d)=\min_{Q^{(n)}\in Q_d^{(n)}}\frac{1}{n}I(Q^{(n)})$$

$$Q_d^{(n)}=\{Q^{(n)}(y^{(n)}/x^{(n)}):\sum_{x^{(n)}y^{(n)}}p^{(n)}(x^{(n)})Q^{(n)}(y^{(n)}/x^{(n)})\rho^{(n)}(x^{(n)},\ y^{(n)})\leqslant d\} \tag{4.22}$$

**引理3** 对上述非平稳,无记忆信源,$R^{(n)}(d)$ 的极小值在无记忆条件分布

$$Q^{(n)}(y^{(n)}/x^{(n)})=\prod_{i=1}^{n}Q_i(y_i/x_i)$$

上达到.

证 由 $p^{(n)}(x^{(n)})$ 的无记忆性及定理3的证明过程知,对任一 $Q^{(n)}(y^{(n)}/x^{(n)})$ 有一个(4.15)型的无记忆条件分布 $Q_*^{(n)}(y^{(n)}/x^{(n)})$ 使 $E_*\{\rho^{(n)}(x^{(n)},\ y^{(n)})\}=E\{\rho^{(n)}(x^{(n)}y^{(n)}\}$

$$I(Q_*^{(n)})\leqslant I(Q^{(n)})$$

成立,即得命题.

**定理4** 上述无记忆信源序列 $\mathscr{S}^{(n)}$ 关于 $R^{(n)}(d)$ 的正、反编码定理成立. 即有

① 如有在 $\varepsilon_0>0$，使 $\frac{1}{n}R^{(n)}>R^{(n)}(d)(1+\varepsilon_0)$，则 $\mathscr{S}^{(n)}(R^{(n)},d)$——压缩；

② 如 $\mathscr{S}^{(n)}(R^{(n)},d)$——压缩，则对任何 $\varepsilon_0>0$，只要 $n$ 充分大就有

$$\frac{1}{n}R^{(n)}>R^{(n)}(d+\varepsilon_0)(1-\varepsilon_0)$$

成立.

证　1）由定理 1，引理 3，仿定理 3 正定理证明就可，这时要求，当 $n$ 充分大时

$$R^{(n)}(d)>\alpha_0>0$$

成立.

2）由定理 2 知，如 $\mathscr{S}^{(n)}$，$(R^{(n)},d)$——压缩则对任何 $\varepsilon_0>0$ 有 $\mathscr{S}^{(n)}(R^{(n)}(1+\varepsilon_0),d)$——信息有界，此即存在条件分布 $Q^{(n)}(y^{(n)}/x^{(n)})$，使

$$\begin{aligned}&p^{(n)}\{(x^{(n)},y^{(n)}):i(x^{(n)},y^{(n)})\leqslant R^{(n)}(1+\varepsilon_0)\}\to 1\\&p^{(n)}\{(x^{(n)},y^{(n)}):\rho^{(n)}(x^{(n)},y^{(n)})\leqslant d\}\to 1\end{aligned}\tag{4.23}$$

成立，因为 $X$, $Y$ 为有限集，因此当 $n$ 充分大时有

$$\begin{aligned}&I(Q^{(n)})\leqslant R^{(n)}(1+2\varepsilon_0)\\&E\{\rho^{(n)}(x^{(n)},y^{(n)})\}\leqslant d+2\varepsilon_0,\end{aligned}\tag{4.24}$$

成立，得 $\frac{1}{n}R^{(n)}(1+2\varepsilon_0)\geqslant R^{(n)}(d+2\varepsilon_0)$ 成立，因为 $2\varepsilon_0$ 任意，即得命题.

**(3) 平稳遍历信源编码定理**

平稳遍历信源描述如[4, 8, 9]等文，记 $X$, $Y$ 为两有限集，而

$$X^I=\prod_{i\in I}\otimes X_i,\quad Y^I=\prod_{i\in I}\otimes Y_i,\quad (X_i=X,\ Y_i=Y),\tag{4.25}$$

上述 $I=\{\cdots-1,0,1,\cdots\}$，它们的元分别为

$$x^I=(\cdots x_{-1},x_0,x_1\cdots),\ y^I=(\cdots y_{-1},y_0,y_1\cdots),$$

其中 $x_i\in X$, $y_i\in Y$，记 $X^{(n)}$, $Y^{(n)}$ 为柱集

$$x^{(n)}=[x_1,\cdots,x_n],\ y^{(n)}=[y_1\cdots y_n]\tag{4.26}$$

全体，由 $X^I$ 全体柱集产生的 $\sigma$-代数为 $\sigma_X$，而 $Y^I$ 全体柱集产生的 $\sigma$-代数为 $\sigma_Y$.

如 $\mu(\cdot)$ 为 $[X^I,\sigma_X]$ 上概率测度，则称 $[X^I,\sigma_X,\mu(\cdot)]$ 为信源过程，而 $[X^{(n)},p^{(n)}(x^{(n)})]$ 为由这个信源过程确定的信源序列，如

$$p^{(n)}(x^{(n)})=\mu(x^{(n)}),\ x^{(n)}\in X^{(n)},\tag{4.27}$$

记 $\omega(\cdot,\cdot)$ 为 $[X^I\otimes Y^I,\sigma_X\otimes\sigma_Y]$ 上概率测度，满足条件

$$\omega(A\otimes Y^I)=\mu(A),\ A\in\sigma_X,\tag{4.28}$$

记

$$p^{(n)}(x^{(n)},y^{(n)})=\omega(x^{(n)}\otimes y^{(n)}),\tag{4.29}$$

$$Q^{(n)}(y^{(n)}/x^{(n)})=\frac{p^{(n)}(x^{(n)},y^{(n)})}{p^{(n)}(x^{(n)})},$$

而 $X^{(n)}\otimes Y^{(n)}$ 上的 $\rho^{(n)}(x^{(n)},y^{(n)})$ 定义如前，如 $\omega(\cdot,\cdot)$ 为平稳的，则

$$E\{\rho^{(n)}(x^{(n)},y^{(n)})\}=E\{\rho(x,y)\}\tag{4.30}$$

成立，数学期望分别对 $p^{(n)}(x^{(n)},y^{(n)})$ 与 $p^{(1)}(x^{(1)},y^{(1)})$ 而取，记

$$\Omega_d=\{\omega(\cdot,\cdot):E\{\rho(x,y)\}\leqslant d,\ \omega(\cdot,Y^I)=\mu(\cdot)\}\tag{4.31}$$

$$I(\omega)=\lim_{n\to\infty}\frac{1}{n}I^{(n)}(x^{(n)},\ y^{(n)}) \tag{4.32}$$

如 $\omega$ 为平移过程，则(4.32)右边极限一定存在.

如 $\mu(\cdot)$ 为平稳遍历，$\Omega_e$ 为全体平稳，遍历的 $\omega(\cdot,\ \cdot)$ 记

$$R_e(d)=\inf_{\omega\in\Omega_d\cap\Omega_e}I(\omega),$$

则平稳遍历信源 $[X^I,\sigma_X,\mu(\cdot)]$ 关于 $R_e(d)$ 的编码定理成立. 也就是对任何 $R>R_e(d)$ 的 $R$，由 $\mu(\cdot)$ 决定的 $\mathscr{S}^{(n)}$ 为 $(nR,d+\varepsilon)$— 压缩，其中 $\varepsilon>0$ 为任意常数.

证明由遍历过程的极限定理即得.

关于上述命题的反定理与非遍历信源的编码定理将涉及到平稳过程的遍历分解问题，我们在此不作详述.

## 5. 关于连续状态下序列模型的信源编码定理

### (1) 一般信源序列的正、反编码定理

以上我们讨论 $X^{(n)}$, $Y^{(n)}$ 均为有限状态，下面将上述结果推广到连续状态，为了简单起见，本段仍省去右上角$(n)$，如无特别声明，各字母右上角均带$(n)$，但省略不写.

我们记信源序列

$$\mathscr{S}=\{[X,\ \sigma_X,\ p_X(\cdot)],\ [Y,\ \sigma_Y],\ \rho(x,\ y)\},$$

其中 $\sigma_X$, $\sigma_Y$ 为 $X$, $Y$ 的子集 $\sigma$–代数 $\{y\}\in\sigma_Y$, $p_X(\cdot)$ 为 $\sigma_X$ 上的概率测度，$\rho(x,\ y)$ 为 $\sigma_X\otimes\sigma_Y$——可测函数而 $\rho(x,\ B)$, $H(B,\ d)$, $p(R,\ d)$, $\mathscr{S}$ 的$(R,\ d)$——压缩定义与(2.1—2.3), (2.6)或(2.7)定义相同.

在讨论连续状态下信源编码问题时涉及到连续状态下的信息量问题，为了避免由此而产生的困难，我们作如下处理.

因为 $Y$ 是一个无限集，因此我们不在 $Y$ 上直接定义 $Q(y/x)$，而是取 $Y$ 的一个有限子集 $Y_0$，在 $Y_0$ 上定义条件概率分布 $Q(y/x)$，则 $Q(y/x)$ 仍为离散概率分布，有公息与公息密度为：

$$I(Q)=\int_X\left[\sum_{y\in Y_0}Q(y/x)\log\frac{Q(y/x)}{Q(y)}\right]p_X(dx), \tag{5.1}$$

$$i(x,\ y)=\log\frac{Q(y/x)}{Q(y)},\ y\in Y_0,\ x\in X, \tag{5.2}$$

其中对 $y\in Y_0$ 的分布为

$$p_{XY}(A,\ y)=\int_A Q(y/x)p(dx),\ Q(y)=p_{XY}(X,\ y) \tag{5.3}$$

上述 $A\in\sigma_X$，要求 $Q(y/x)$ 对每个 $y\in Y$ 为 $\sigma_X$ 一可测函数.

称上述 $\mathscr{S}$ 为$(\mathscr{F},\ d)$——信息有界，如存在一有限集 $Y_0\in Y$，与 $Y_0$ 上条件分布 $Q(y/x)$ 使(2.8), (2.9)成立. 这样我们就可推广定理 1, 2 的结论，证明过程完全相同，只要将相应的 $p(x,\ y)$, $p(x)$ 改为 $p(dx,\ y)$, $p(dx)$; $\sum_{x\in A}$ 改为 $\int_A$ 就可.

**(2) 平稳无记忆信源编码定理**

关于连续状态下平稳，无记忆信源的描述和结论与离散情形相同，相应的 $X$, $Y$, $X^{(n)}$, $Y^{(n)}$ 改为 $(X, \sigma_X)$, $(Y, \sigma_Y)$, $(X^{(n)}, \sigma_X^{(n)})$, $(Y^{(n)}, \sigma_Y^{(n)})$，其中

$$\sigma_X^{(n)}=\prod_{i=1}^{n}\otimes\sigma_{X_i},\quad \sigma_Y^{(n)}=\prod_{i=1}^{n}\otimes\sigma_{Y_i},$$

而 $X_i=X$, $Y_i=Y$, $i=1, \cdots, n$，同样对 $Q(y/x)$ 我们作上述离散化处理，记 $Y_0$, $Y_0^{(n)}$ 为 $Y$, $Y^{(n)}$ 的有限子集．定义

$$R(d)=\inf_{Y_0\in\mathscr{Y}}\{\inf_{Q\in Q_d(Y_0)} I(Q)\},\tag{5.4}$$

其中 $\mathscr{Y}$ 为 $Y$ 中全体有限子集集合，而 $Q_d(Y_0)$ 为全体这样的 $Q(y/x)$，使

1) $Q(y/x)$ 为 $Y_0$ 上条件概率分布；

2) 有关系式

$$\int_X[\sum_{y\in Y_0}Q(y/x)\rho(x, y)]p(dx)\leqslant d\tag{5.5}$$

成立．

这样即可得到在一定条件下关于 $R(d)$ 的正、反编码定理．这个条件只要保证(4.17)成立即可，如 $\rho^{(n)}(x, y)$ 为在 $X\otimes Y$ 上一致有界等条件就可，另外我们注意到，由本文(5.4)定义的 $R(d)$ 同样为 $d$ 的下凹函数，即对任何 $0\leqslant\lambda\leqslant1$，有

$$R(\lambda d+\bar{\lambda}d')\leqslant\lambda R(d)+\bar{\lambda}R(d')$$

成立 $(\bar{\lambda}=1-\lambda)$，证明可仿[4]文定理 2、4、1 的证明即可．我们在此不一一细述了．

本文部分结果的化简与推广参阅文[13]．

## 参　考　文　献


[1] C. E. Shannon, *B. S. T. J.*, **27** (1948), 379—423; 623—658.

[2] C. E. Shannon, *Nat. Conv. Rec., Pt.* **4** (1959), 142—163.

[3] R. G. Gallager, «Information Theory and Reliable Communication», *New York, Wiley*, (1968).

[4] T. Berger, «Rate Distortion Theory» Englewood Cliffs, N. J., *Prentice-Hall*, (1971).

[5] 胡国定，**数学学报**，**11**: 3 (1961), 260—294.

[6] Hu Guo-ding, Trans Third prague Conforence on Inform. Theory etc. (1962), 285—332.

[7] K. Morton, IEEE Trans. Inform. Theory, **IT**: 20 (1974), 192—200.

[8] T. M. Cover, IEEE Trans. Inform. Theory, **IT**: 21 (1975), 226—228.

[9] L. D. Davisson and M. B. Pursley, IEEE Trans. Inform. Theory, **IT**: 21 (1975), 310—318.

[10] R. M. Gray and L. D. Davisson, IEEE Trans. Inform. Theory, **IT**: 20 (1974), 520—517.

[11] R. M. Gray, D. L. Neuhoff and J. K. Omvra, IEEE Trans. Inform, Theory, **IT**: 21 (1975), 524—533.

[12] M. B. Pursley and L. D. Davisson, IEEE Trans. Inform. Theory, **IT**: 22 (1976), 324—337.

[13] 沈世镒，具有道信息的信源编码定理(未发表).

# SOURCE CODING THEOREMS FOR GENERAL SEQUENCE MODEL

SHEN SHIYI
(*Nankai University*)

## ABSTRACT

In this paper, we give a coding theorem for general source sequence. A source sequence

$$\mathscr{S}^{(n)}=\{[X^{(n)}, p^{(n)}(X^{(n)})], [X^{(n)}\otimes Y^{(n)}, \rho^{(n)}(X^{(n)}, Y^{(n)})]\}$$

is said to be $(R^{(n)}, d^{(n)})$-compress, if (i) $R^{(n)}$, $d^{(n)}$ are two sequences of real numbers $R^{(n)}\to\infty$; (ii) there exist $\varepsilon^{(n)}>0$ $(\varepsilon^{(n)}\to 0)$ and set $B^{(n)}\subset Y^{(n)}$, so that $|B^{(n)}|\leqslant 2^{R^{(n)}(1+\varepsilon^{(n)})}$ and

$$p^{(n)}\{X^{(n)}:\rho^{(n)}(X^{(n)}, B^{(n)})\leqslant d^{(n)}\}\geqslant 1-\varepsilon^{(n)},$$

where

$$\rho^{(n)}(X^{(n)}, B^{(n)})=\min_{y^{(n)}\in B^{(n)}}\rho^{(n)}(X^{(n)}, Y^{(n)}).$$

A $\mathscr{S}^{(n)}$ is said to be $(\mathscr{F}^{(n)}, D^{(n)})$-information bounded, if (i) $\mathscr{F}^{(n)}\to\infty$; (ii) there exist $\varepsilon^{(n)}\to 0$ and conditional probability $Q^{(n)}(Y^{(n)}/X^{(n)})$ so that probability distribution $p^{(n)}(X^{(n)}, Y^{(n)})=p^{(n)}(X^{(n)})Q^{(n)}(Y^{(n)}/X^{(n)})$ is satisfied by

$$p^{(n)}\{(X^{(n)}, Y^{(n)}):i(X^{(n)}, Y^{(n)})\leqslant\mathscr{F}^{(n)}(1+\varepsilon^{(n)}),$$
$$\rho^{(n)}(X^{(n)}, Y^{(n)})\leqslant\rho^{(n)}\}\geqslant 1-\varepsilon^{(n)},$$

where

$$i(X^{(n)}, Y^{(n)})=\log\frac{p^{(n)}(X^{(n)}, Y^{(n)})}{p^{(n)}(X^{(n)})p^{(n)}(Y^{(n)})}.$$

**Theorem.** The necessary and sufficient conditions for a source sequence $\mathscr{S}^{(n)}$ to be $(R^{(n)}, \alpha^{(n)})$-compress is that $\mathscr{S}^{(n)}$ must be $(R^{(n)}, d^{(n)})$-information bounded.

From the theorem we obtain immediately the coding theorem and its converse for stationary and unstationary sources with memory.