熵的性质（1）

武汉理工大学

Information theory

and

coding

武漢理工大學

熵的性质

熵公式： $H(X)=\sum_{k=1}^{K} p_k \log \frac{1}{p_k} = -\sum_{k} p_k \log p_k$

1 对称性：

$$H(p_1, p_2, \cdots, p_K) = H(p_{m(1)}, p_{m(2)}, \cdots, p_{m(K)})$$

其中 $\{m(1), m(2), \cdots, m(K)\}$ 是 $\{1, 2, \cdots, K\}$ 的任意置换。

2 可扩展性：加入零概率事件不会改变熵。

$$H(p_1, p_2, \cdots, p_K) = H(p_1, \cdots, p_i, 0, p_{i+1}, \cdots, p_K) \quad \because 0\log 0 = 0$$

$$i = 1, 2, \cdots, K-1$$

熵的性质

3 非负性：

熵公式：$H(X)=\sum_{k=1}^{K} p_k \log \frac{1}{p_k} = -\sum_{k} p_k \log p_k$

$$H(p_1, p_2, \cdots, p_K) = H(P) \geq 0$$

确定性概率分布：$P_s = \{1, 0, \cdots, 0\}$

$$H(P_s) = 0$$

熵的性质

4 强可加性：

$$H(p_1q_{11},\cdots,p_1q_{1J},\cdots,p_Kq_{K1}\cdots,p_Kq_{KJ})$$

$$=H(p_1,p_2,\cdots,p_K)+\sum_{k=1}^{K}p_kH(q_{k1},\cdots,q_{kJ})$$

$X=\{x_1,\ x_2,\ \ldots,\ x_k\}$，$PX=\{p_1,\ p_2,\ \ldots,\ p_k\}$

$Y=\{y_1,\ y_2,\ \ldots,\ y_J\}$，$PY=\{q_1,\ q_2,\ \ldots,\ q_J\}$

q_{kj}为转移概率$q_{kj}=P(y_j|x_k)$

$$0\le p_k\le 1\quad k=1,2,\cdots,K\qquad \sum_{k=1}^{K}p_k=1$$

$$0\le q_{kj}\le 1\quad j=1,2,\cdots,J\qquad \sum_{j=1}^{J}q_{kj}=1\quad k=1,2,\cdots,K$$

“强可加性”证明

定义新函数：　$L(x) = -x\log x$

则　$L(xy) = -xy\log(xy) = -xy\log x - xy\log y = yL(x) + xL(y)$

于是　$H(p_1q_{11},\cdots,p_1q_{1J},\cdots,p_Kq_{K1}\cdots,p_Kq_{KJ})$

$$= \sum_{k=1}^{K}\sum_{j=1}^{J} L(p_K q_{Kj}) = \sum_{k=1}^{K}\sum_{j=1}^{J}[q_{kj}L(p_k) + p_k L(q_{kj})]$$

$$= \sum_{k=1}^{K}\left(L(p_k)\sum_{j=1}^{J} q_{kj} \right) + \sum_{k=1}^{K}\left(p_k \sum_{j=1}^{J} L(q_{kj}) \right)$$

$$= H(p_1, p_2, \cdots, p_M) + \sum_{k=1}^{K} p_k H(q_{k1}, q_{k2}\cdots, q_{kJ})$$

$$H(XY) = H(X) + H(Y|X)$$

$\sum_{j=1}^{J} q_{kj} = 1$

$k = 1, 2, \cdots, K$

“强可加性”证明

5 可加性：

$$H(p_1q_1,\cdots,p_1q_J,\cdots,p_Kq_1\cdots,p_Kq_J)$$
$$=H(p_1,p_2,\cdots,p_K)+H(q_1,q_2\cdots,q_J)$$

$$0\le p_k\le 1 \quad k=1,2,\cdots,K \quad \sum_{k=1}^{K}p_k=1$$

$$0\le q_j\le 1 \quad j=1,2,\cdots,J \quad \sum_{j=1}^{J}q_j=1$$

“可加性”是“强可加性”的特殊情况，在“强可加性”中，令

$$q_{1j}=q_{2j}=\cdots=q_{Kj}=q_j \ ,j=1,\cdots,J$$

就可得出可加性。

“可加性”证明

强可加性：

$$H(p_1 q_{11}, \cdots, p_1 q_{1J}, \cdots, p_K q_{K1} \cdots, p_K q_{KJ})$$

$$= H(p_1, p_2, \cdots, p_K) + \sum_{k=1}^{K} p_k H(q_{k1}, \cdots, q_{kJ})$$

$$0 \le p_k \le 1 \quad k = 1，2，\cdots, K \qquad \sum_{k=1}^{K} p_k = 1$$

$$0 \le q_{kj} \le 1 \quad j = 1，2，\cdots, J \qquad \sum_{j=1}^{J} q_{kj} = 1 \quad k = 1，2，\cdots, K$$

令 $q_{1j} = q_{2j} = \cdots = q_{Kj} = q_j, \ j = 1, \cdots, J$

则 $H(p_1 q_1, \cdots, p_1 q_J, \cdots, p_K q_1 \cdots, p_K q_J)$

$$= H(p_1, p_2, \cdots, p_K) + \sum_{k=1}^{K} p_k H(q_1, \cdots, q_J)$$

$$= H(p_1, p_2, \cdots, p_K) + H(q_1, \cdots, q_J) \sum_{k=1}^{K} p_k$$

$$= H(p_1, p_2, \cdots, p_K) + H(q_1, \cdots, q_J)$$

$H(XY)=H(X)+H(Y)$

感谢观看！

Information theory

and

coding

武漢理工大學