5.4 汉明距离



- 上节讨论译码时曾经提过,可将接收序列译为与 之最"相似"的输入序列(码字)。
- 如何定量描述符号序列之间的"相似"程度呢?
- 汉明(R.W.Hamming) 受距离概念的启发,在符号序列之间引入汉明距离,用来定量描述符号序列之间的"相似"程度。







1、汉明距离的定义与性质

定义:两个等长符号序列 \bar{x} 和 \bar{y} 之间的汉明距离,记为 $D(\bar{x},\bar{y})$,是 \bar{x} 与 \bar{y} 之间对应位置上不同符号的个数。

例:
$$\overline{x} = 1011111$$
 $\overline{y} = 111100$ $D(\overline{x}, \overline{y}) = 3$, $\alpha = 1320120$ $\beta = 1220320$ $D(\alpha, \beta) = 2$

用汉明距离来度量两个符号序列的"相似"程度:

 $D(\bar{x},\bar{y})$ 小: \bar{x} 与 \bar{y} 的相似程度高。

 $D(\bar{x},\bar{y})$ 大: \bar{x} 与 \bar{y} 的相似程度低。

相似程度的高低是相对而言的。

汉明距离的性质(距离公理):

- (1) 非负性: $D(\bar{x},\bar{y}) \geq 0$, 当且仅当 $\bar{x} = \bar{y}$ 时等号成立;
 - (2) 对称性: $D(\bar{x}, \bar{y}) = D(\bar{y}, \bar{x})$
- (3) 三角不等式: $D(\overline{x},\overline{z}) + D(\overline{z},\overline{y}) \ge D(\overline{x},\overline{y})$

2、二元序列的汉明距离

$$\overline{x} = x_1 x_2 \cdots x_N , \quad x_k \in \{0,1\}$$

$$\overline{y} = y_1 y_2 \cdots y_N , \quad y_k \in \{0,1\}$$

$$D(\overline{x},\overline{y}) = \sum_{k=1}^{N} x_k \oplus y_k$$

二元序列汉明重量 W(x): 二元序列 \overline{x} 中含 "1"的个数。



$$W(\overline{x}) = D(\overline{x}, O_N)$$









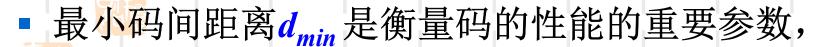


3、码的相似性

等长码:
$$C = \{c_1, c_2, \dots, c_q\}$$

码间距离: $D(c_i,c_j)$, $c_i \neq c_j$, $c_i,c_j \in C$

码
$$C$$
的最小码间距离: $d_{\min} = \min \left[D(c_i, c_j) \right]$ $c_i \neq c_j$ $c_i, c_j \in C$



- 码距小:说明有些码字受干扰后容易变为另一码字, 译码时就会出错。
- 进行信道编码时,只要条件允许,尽量选择最小码间距离大一些的码。

