5.6 线性分组码

- 信道编码的目的是为了降低平均差错率，又称纠错编码。
- 香农的有噪声信道定理的意义在于，它告诉我们什么是通过努力可以做到的事情，什么是不可能做到的事情。但香农并没有给出切实可行的实现方法。
- 香农提出的随机编码方法，是一种为避免寻找好码而采取的权宜之计，有理论意义而无实用价值。
- 有实用价值的码须用适当的数学工具来构造，使得构造出的码具有很好的结构特性，以便于译码。
- 纠错编码理论几乎与信息论同时创立，创始人是汉明。
- 纠错编码理论发展到现在，已形成以近世代数为理论基础的系统编码理论，又称代数编码理论。

纠错编码的基本概念

- 纠错编码的基本思路：引入可控冗余，即在信息序列中加入一些冗余码元（或称校验码元）。
- 译码：利用码元之间的相关性质来检测错误和纠正错误。
- 分组码：先将信息序列分成K个符号一组，称为信息组，然后在信息组中加入一些校验码元组成N长码字，由此得到的码称为（N，K）分组码。分组码中的任一码字的码长为N，所含的信息位数目为K、校验位数目为$N-K$。
- 线性码：线性码的最重要性质是线性特性，即信息位和监督位之间的关系为线性关系，码中任意两个码字的和仍为该编码的码字。否则为非线性码。
- 循环码：循环码是线性码的一个子集。循环码中任一码字循环移位后仍为该码的码字。否则为非循环码。

1、线性分组码的生成矩阵和校验矩阵

（以（5，2）分组码为例）

信息组长度 $K=2$：

$\boldsymbol{m}=[m_1 \quad m_2]; m_i \in \{0,1\}$

码字长度$N=5$:

$\boldsymbol{c}=[c_1 \quad c_2 \quad c_3 \quad c_4 \quad c_5]$

编码函数：

$$f:\begin{cases} c_1 = m_1 \\ c_2 = m_2 \\ c_3 = m_1 \oplus m_2 = c_1 \oplus c_2 \\ c_4 = m_1 = c_1 \\ c_5 = m_1 \oplus m_2 = c_1 \oplus c_2 \end{cases}$$

码字由信息元的模2线性组合生成，因此是二元线性分组码，简称为线性码。

编码函数的矩阵表示：$[c_1 \quad c_2 \quad c_3 \quad c_4 \quad c_5]=[m_1 \quad m_2]\begin{bmatrix}1&0&1&1&1\\0&1&1&0&1\end{bmatrix}$

$\boldsymbol{c}=\boldsymbol{mG}$ ← 生成矩阵$\boldsymbol{G}$

校验方程的矩阵表示：

$$\begin{cases} c_3 = m_1 \oplus m_2 = c_1 \oplus c_2 \\ c_4 = m_1 = c_1 \\ c_5 = m_1 \oplus m_2 = c_1 \oplus c_2 \end{cases} \Longrightarrow \begin{cases} c_1 \oplus c_2 \oplus c_3 = 0 \\ c_1 \oplus c_4 = 0 \\ c_1 \oplus c_2 \oplus c_5 = 0 \end{cases} \Longrightarrow \begin{cases} \begin{bmatrix}1&1&1&0&0\\1&0&0&1&0\\1&1&0&0&1\end{bmatrix}\begin{bmatrix}c_1\\c_2\\c_3\\c_4\\c_5\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix} \end{cases}$$

$\boldsymbol{Hc}^T=\boldsymbol{0}$　　$\boldsymbol{cH}^T=\boldsymbol{0}$

一致性校验矩阵$\boldsymbol{H}$

二元（N，K）线性码

信息组，K 维行阵：$\boldsymbol{m}=[m_1 \quad m_2 \quad \cdots \quad m_K]$，$m_i \in \{0,1\}$

码字，N 维行阵：$\boldsymbol{c}=[c_1 \quad c_2 \quad \cdots \quad c_N]$，$c_i \in \{0,1\}$

码字生成式：$\boldsymbol{c}=\boldsymbol{m}\boldsymbol{G}$

$\boldsymbol{G}$：$K \times N$ 生成矩阵，其元素取值于二元集合 {0，1}。

校验方程：$\boldsymbol{H}\boldsymbol{c}^T=\boldsymbol{0}$　　$\boldsymbol{c}\boldsymbol{H}^T=\boldsymbol{0}$

$\boldsymbol{H}$：$r \times N$ 校验矩阵，其元素取值于二元集合 {0，1}。

$r=N-K$：校验位数目。

二元（N，K）线性码（续一）

$$\boldsymbol{G}=\begin{bmatrix}\boldsymbol{g}_1\\ \boldsymbol{g}_2\\ \vdots\\ \boldsymbol{g}_K\end{bmatrix}=\begin{bmatrix}g_{11} & g_{12} & \cdots & g_{1N}\\ g_{21} & g_{22} & \cdots & g_{2N}\\ \vdots & \vdots & \vdots & \vdots\\ g_{K1} & g_{K1} & \cdots & g_{KN}\end{bmatrix} \qquad \boldsymbol{c}=\boldsymbol{mG}=m_1\boldsymbol{g}_1\oplus m_2\boldsymbol{g}_2\oplus\cdots\oplus m_K\boldsymbol{g}_K$$

（1）$\boldsymbol{G}$ 的每个向量都是一个码字。

例：$m_1=1,\ m_2=m_3=\cdots=m_K=0 \Longrightarrow c=g_1$

(2) 二元(*N*，*K*)线性码*C*={*c*}可看成一个N重维线性空间，*G*的*K*个相互独立的行向量是它的一组基底。

(3) 任意*K*个相互独立的*N*长码字都可作为*N*重维码空间的一组基底，用这个码字当作行向量组成生成矩阵，即可生成所有码字。

二元（N，K）线性码（续二）

系统码：码字的前（或后）K位照搬信息组的K个信息元

对于前K位为信息元的系统码，生成矩阵G可分成2块：

$$G=[I_{K\times K}\quad A_{K\times r}]=\begin{bmatrix}1&0&\cdots&0&a_{11}&a_{12}&\cdots&a_{1r}\\0&1&\cdots&0&a_{21}&a_{22}&\cdots&a_{2r}\\\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&\cdots&1&a_{K1}&a_{K2}&\cdots&a_{Kr}\end{bmatrix}=\begin{bmatrix}g_1\\g_2\\\vdots\\g_K\end{bmatrix}$$

校验方程：$\boxed{cH^T=0}$

g_i是码字，满足校验方程 。

$$g_iH^T=0,\quad i=1,2,\cdots,K$$

$$\left.\begin{aligned}GH^T&=0\\G&=[I_{K\times K}\quad A_{K\times r}]\end{aligned}\right\}\Rightarrow H=[A_{K\times r}^T\quad I_{r\times r}]$$

验证：$$GH^T=[I_{K\times K}\quad A_{K\times r}]\left[A_{K\times r}^T\quad I_{r\times r}\right]^T=[I_{K\times K}\quad A_{K\times r}]\begin{bmatrix}A_{K\times r}\\I_{r\times r}\end{bmatrix}=A_{K\times r}\oplus A_{K\times r}=0$$

2、汉明距离与码的纠、检错能力

- 检错：译码器能检测到是否有错误发生，码的检错能力用检测到的错误位数t_d描述；
- 纠错：译码器不但能检测到是否有错误发生，而且能纠正发生的错误，码的纠错能力用纠正错误的位数t_c描述。
- 无法检出或纠正的错误：码字出错而变为另一码字。这种情况最易发生在较为相似的码字之间。
- 码的纠、检错能力与码的最小汉明距离关系：

(1)一个码能够检测出t_d个错误的充要条件：$d_{min} \geq t_d + 1$

(2)一个码能够纠正t_c个错误的充要条件：$d_{min} \geq 2t_c + 1$

(3)一个码能够纠正t_c个错误，同时又能够检测出t_d个错误的充要条件：

$$d_{min} \geq t_c + t_d + 1 \quad (t_d > t_c)$$

二元线性分组码的最小汉明距离

结论：二元线性分组码的最小汉明距离等于该码非零码字的最小汉明重量。

例：C={00000,01101,10111,11010}，求最小汉明距离。

$W_{min}=3$

$d_{min}=W_{min}=3$

例　“重复2次”编码的检错和纠错能力

“重复2次”编码：0→000，1→111

码字：$C=\{000, 111\}$，$d_{min}=3$

检错：$d_{\min} \geq t_d + 1$

$d_{\min} = 3 = 2 + 1$

接收序列	译码
000	0
001	Error
010	Error
011	Error
100	Error
101	Error
110	Error
111	1

纠错：$d_{\min} \geq 2t_c + 1$

$d_{\min} = 3 = 2 \times 1 + 1$

接收序列	译码
000	0
001	0
010	0
011	1
100	0
101	1
110	1
111	1

例 比较(5，2)线性码和“重复2次”码

(5，2)线性码：$\boldsymbol{G}=\begin{bmatrix}1&0&1&1&1\\0&1&1&0&1\end{bmatrix}$

$\boldsymbol{c}=\boldsymbol{mG}$　　$d_{\min}=3$

信息组$\boldsymbol{m}$	码字$\boldsymbol{c}$
00	00000
01	01101
10	10111
11	11010

与“重复2次”码的最小汉明距离相同，因此，检、纠错能力相同：能检出2个错误或纠正1个错误。

“重复2次”码：$P_e=3\times10^{-4}$　　$\boldsymbol{R=1/3}$ **bit/符号**

(5，2)线性码：$P_e=7.86\times10^{-4}$　　$\boldsymbol{R=2/5}$ **bit/符号**