第五章离散信道编码

第四节二元线性分组码

陈兴同

中国矿业大学 数学学院

2021 年 8 月

内容提要

内容提要

内容提要

内容提要

定义 5.4.1：域

为了学习研究线性分组码，现在介绍一下有限域及其上的向量空间的概念。

设 $\mathcal{F}$ 是一个非空集合，在这个集合规定二种运算，一个是加法记为 $+$，一个是乘法记为 $*$。如果这两种运算满足如下运算性质，就称集合 $\mathcal{F}$ 是一个**域**。

（1） 对于任何 $\alpha, \beta, \gamma \in \mathcal{F}$ 成立下列算律：

a） 加法封闭性即 $\alpha + \beta \in \mathcal{F}$。

b） 加法结合律即

$$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma。$$

c） 加法交换律即 $\alpha + \beta = \beta + \alpha$。

d） 乘法封闭性即 $\alpha * \beta \in \mathcal{F}$。

e） 乘法结合律即

$$\alpha * (\beta * \gamma) = (\alpha * \beta) * \gamma。$$

续定义：

f） 乘法交换律即 $\alpha * \beta = \beta * \alpha$。

g） 加法乘法分配律即：$\alpha * (\beta + \gamma) = \alpha * \beta + \alpha * \gamma$。

（2） 对于加法存在唯一零元素记为 0，使得任何 $\alpha \in \mathcal{F}$ 成立 $\alpha + 0 = \alpha$。

（3） 对于乘法存在唯一单位元素记为 1，使得任何 $\alpha \in \mathcal{F}$ 成立 $\alpha * 1 = \alpha$。

（4） 对于任何 $\alpha \in \mathcal{F}$，加法存在唯一负元素记为 $-\alpha \in \mathcal{F}$，成立 $\alpha + (-\alpha) = 0$。

（5） 对于任何 $\alpha \in \mathcal{F}, \alpha \neq 0$，乘法存在唯一逆元素记为 $\alpha^{-1} \in \mathcal{F}$，成立 $\alpha * \alpha^{-1} = 1$。

常见域：

显然全体实数 $\mathbb{R}$，关于数量加法与乘法构成一个域，称为实数域；
全体复数 $\mathbb{C}$ 关于复数加法与乘法也构成一个域，称为复数域。
全体有理数集合 $\mathbb{Q}$ 关于数量的加法与乘法也构成一个域，称为有理数域。
可以证明质数 $p \geq 2$ 的除法剩余类 $\mathcal{F}_p = \{0, 1, \cdots, p-1\}$，关于模 p 的加法与乘法也构成一个域，称为 p 元 Galois 域。
在编码学中经常用到这种有限域，下面举两个例子。

例题 5.4.1：

下面集合及其运算构成的域称为二元 **Galois 域**,$\mathcal{F}_2 = \{0, 1\}$。

$+$	0	1
0	0	1
1	1	0

,

$*$	0	1
0	0	0
1	0	1

.

5.10
(1.1)

例题 5.4.2：

（五元 Galois 域） 下面集合及其运算构成的域称为五元 Gailos 域，$\mathcal{F}_5 = \{0, 1, 2, 3, 4\}$。

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

，

$*$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

．

5.11 (12)

分组消息及其码字都用向量表示，都属于某个向量空间，现在给出域上向量空间的定义。

定义 5.4.2：向量空间定义

设 $\mathcal{V}$ 是一个非空集合，$\mathcal{F}$ 是一个域。规定 $\mathcal{V}$ 中的一个加法运算记为 $\oplus$，再规定一个 $\mathcal{F}$ 与 $\mathcal{V}$ 之间的运算称为数乘运算记为 $\otimes$，如果满足如下八条运算律则称 $\mathcal{V}$ 是域 $\mathcal{F}$ 上的**向量空间**。

（1） $\forall \alpha, \beta \in \mathcal{V}$ 成立加法交换律即 $\alpha \oplus \beta = \beta \oplus \alpha$

（2） $\forall \alpha, \beta, \gamma \in \mathcal{V}$ 成立加法结合律即 $\alpha \oplus \beta) \oplus \gamma = \alpha \oplus (\beta \oplus \gamma)$

（3） 在 $\mathcal{V}$ 中有一个元素记作 0，使得 $\forall \alpha \in V$ 成立 $\alpha \oplus 0 = \alpha$，这个称为零元素。

（4） $\forall \alpha \in \mathcal{V}$ 存在另一个 $\beta \in \mathcal{V}$ 成立 $\alpha \oplus \beta = 0$，这个 β 称为 α 的负元素。

（5） $\forall \alpha \in \mathcal{V}$ 以及域 $\mathcal{F}$ 中的单位元素 1，成立 $1 \otimes \alpha = \alpha$。

（6） $\forall k, l \in \mathcal{F}, \alpha \in \mathcal{V}$，成立 $k \otimes (l \otimes \alpha) = (kl) \otimes \alpha$。

（7） $\forall k, l \in \mathcal{F}, \alpha \in \mathcal{V}$，成立 $k + l) \otimes \alpha = k \otimes \alpha + l \otimes \alpha$。

（8） $\forall k \in \mathcal{F}, \alpha, \beta \in \mathcal{V}$ 成立 $k \otimes (\alpha \oplus \beta) = k \otimes \alpha \oplus k \otimes \beta$。

例题 5.4.3：

经常用到的是二元 Galois 域上的向量空间。

全体二元 n 长行向量的集合 $\mathcal{F}_2^n$ 关于运算 $\oplus, \otimes$ 构成域 $\mathcal{F}_2$ 上的向量空间。

$$\begin{aligned} x^{(n)} \oplus y^{(n)} &= (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n), \\ a \otimes x^{(n)} &= (a * x_1, a * x_2, \cdots, a * x_n), \end{aligned}$$

其中的加法 $+$ 与乘法 $*$ 是 (5-10) 式规定的 $\mathcal{F}_2$ 上的运算。

作为向量空间，$\mathcal{F}_2^n$ 中当然有线性组合、线性相关、线性无关，基，维数，坐标，子空间等概念。

例题 5.4.4：

全体 $m \times n$ 矩阵构成的集合 $\mathcal{F}_2^{m\times n}$ 关于如下运算 $\oplus, \otimes$ 构成域 $\mathcal{F}_2$ 上的向量空间，称为矩阵空间。

$$\begin{aligned} A \oplus B &= (a_{ij} + b_{ij}), \\ a \otimes A &= (a * a_{ij}), \end{aligned}$$

其中的加法 $+$ 与乘法 $*$ 是 (5-10) 式规定的 $\mathcal{F}_2$ 上的运算。

例 5.4.3 定义的向量空间 $\mathcal{F}_2^m$ 中向量在矩阵乘法意义下可以左乘例 5.4.4 中定义的矩阵空间 $\mathcal{F}_2^{m\times n}$ 中的矩阵。

消息空间和编码空间：

分组编码就是将消息划分成固定长度，然后再进行编码。对二元信道来说，设输入信道编码器的每组消息长度为 k，它们来自 Gailos 域 $\mathcal{F}_2 = \{0, 1\}$ 上的向量空间 $\mathcal{F}_2^k$；而信道编码器的输出是长度为 n 的码字，经过新到的传输后输出也是长度为 n 的序列（由于有干扰，这些输出未必是码字），它都是来自 Galois 域 $\mathcal{F}_2$ 上的向量空间 $\mathcal{F}_2^n$。

$$\begin{aligned} \mathcal{F}_2^k &= \{(u_1, u_2, \cdots, u_k) | u_i = 0, 1, i = 1, 2, \cdots, k\}, \\ \mathcal{F}_2^n &= \{(x_1, x_2, \cdots, x_n) | x_i = 0, 1, i = 1, 2, \cdots, n\}, \end{aligned}$$

今后其中的零向量记为 **0**。

定义 5.4.3：线性分组码定义

如果编码 $f:\mathcal{F}_2^k \to \mathcal{F}_2^n$ 对任何 $u^{(k)}, v^{(k)} \in \mathcal{F}_2^k, a, b \in \mathcal{F}_2$ 满足：

$$f(a \otimes u^{(k)} \oplus b \otimes v^{(k)}) = a \otimes f(u^{(k)}) \oplus b \otimes f(v^{(k)}),$$

则称编码 f 为 (n,k)**线性分组码**；而 $R = k/n$ 称为码率，$f(u^{(k)})$ 称为消息 $u^{(k)}$ 的码字。k 称为消息长度，n 称为码字长度，$r = n - k$ 称为校验位或冗余位长度。如果 f 是单射，则称为唯一可译码。

根据定义，可知线性码有如下特点：

（1） 任何两个码字之和仍是一个码字；

（2） 必定有分量全为 0 的码字；

（3） 全体码字构成一个向量空间称为**码字空间**，$\mathcal{C} = f(\mathcal{F}_2^k)$，它是向量空间 $\mathcal{F}_2^n$ 的一个子空间。

可以用上面三条来检验一个编码是否为线性码。

生成矩阵：

取 k 长消息空间 $\mathcal{F}_2^k$ 中自然基

$$e_1 = (\ 1\quad 0\quad \cdots\quad 0\ ), e_2 = (\ 0\quad 1\quad \cdots\quad 0\ ), \cdots, e_k = (\ 0\quad 0\quad \cdots\quad 1$$

则每个 k 长消息可由基线性表示

$$u^{(k)} = (u_1, u_2, \cdots, u_k) = u_1 \otimes e_1 \oplus u_2 \otimes e_2 \oplus \cdots \oplus u_k \otimes e_k,$$

于是

$$\begin{aligned} x^{(n)} = f(u^{(k)}) &= u_1 \otimes f(e_1) \oplus u_2 \otimes f(e_2) \oplus \cdots \oplus u_k \otimes f(e_k) \\ &= (u_1, u_2, \cdots, u_k) \begin{pmatrix} f(e_1) \\ f(e_2) \\ \vdots \\ f(e_k) \end{pmatrix} = u^{(k)} G, \end{aligned}$$

续：生成矩阵

注意到其中 $f(e_i)$ 都是行向量，并且是消息 e_i 的码字，因此叠在一起构成矩阵 G，称为线性码 f 的**生成矩阵**。利用生成矩阵 G，可以将线性码 f 表成矩阵向量乘积：

$$x^{(n)} = u^{(k)}G.$$

能生成相同线性码空间 $\mathcal{C} = f(\mathcal{F}_2^k)$ 的生成矩阵一般不唯一。如果选择线性码空间 $\mathcal{C}$ 的任意一组基向量 $x_1^{(n)}, x_2^{(n)}, \cdots, x_k^{(n)}$，则任意一个码字都可以表成基的线性组合

$$x^{(n)} = v_1 \otimes x_1^{(n)} \oplus v_2 \otimes x_2^{(n)} \oplus \cdots \oplus v_k \otimes x_k^{(n)}.$$

如果用这些基向量作为行构成矩阵 G，则有 $x^{(n)} = v^{(k)}G.$，这说明 G 是一个生成矩阵，选择不同的基向量，就可以得到不同的生成矩阵！

练习：

证明：f 是唯一可译码的充要条件是它的生成矩阵 G 行满秩。

例题 5.4.5：

求由生成矩阵 G 生成的线性码。

$$G = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

解：

这是一个 $(6,3)$ 分组码，由消息与码字之间关系 $x^{(n)} = u^{(k)}G$，可以导出如下关系 生成方程

$$\begin{cases} x_1 = u_1 + u_3 \\ x_2 = u_1 + u_2 \\ x_3 = u_1 \end{cases}, \quad \begin{cases} x_4 = u_2 + u_3 \\ x_5 = u_2 \\ x_6 = u_3 \end{cases}, \tag{2.1}$$

分别取遍所有的 3 长消息 $u_1u_2u_3$ 代入上式即得全体码字

消息	000	001	010	011	100	101	110
码字	000000	100101	010110	110011	111000	011101	10111

码字空间是

$$\mathcal{C} = \{000000, 111000, 010110, 100101, 101110, 110011, 011101, 001011\},$$

码字的第 3、5、6 位是消息位，可以验证 $\mathcal{C}$ 是向量空间 $\mathcal{F}_2^6$ 的子空间。

定义 5.4.4：系统码

系统码 如果 (n,k) 线性分组码生成矩阵 G 具有分块标准形

$$G = (I_k, A)_{k\times n}, \tag{2.2}$$

其中 I_k 为 k 阶单位阵，则由它生成的 (n,k) 线性分组码称为**系统码**。

显然，k 长消息 $u^{(k)} = (u_1, u_2, \cdots, u_k)$ 对应的系统码是 $x^{(n)} = (u_1, u_2, \cdots, u_k, *, \cdots, *)$，它的前 k 个分量恰好是消息，因此系统码就是前面 k 位为**消息位**的 (n,k) 线性分组码，系统码中其它位称为**校验位或冗余位**。系统码好处是：译码时可以直接从码字截取前面的字符得到消息。

例题 5.4.6：

如下矩阵 G 生成的线性码是 $(7,4)$ 系统码。

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} = (I_4, A).$$

非系统码：

（1）如果一个生成矩阵 G 不是（5-13）的分块标准形，那么它生成的线性码称为非系统码，比如例题 5.4.5 中生成的码就不是系统码。

（2）如果对一个生成矩阵 G_1 作有限次行初等变换后得到另一个矩阵 G_2，则这两个矩阵仍然生成相同的线性码空间；

（3）如果对一个生成矩阵 G_1 作有限次列交换后得到另一个矩阵 G_2，则这两个矩阵生成的线性码空间就不相同了（实际上是同构），但区别仅仅是码字符的排列顺序不同，而关于线性码的性质完全相同！比如：汉明距离，检错纠错能力，码率，码字个数等等。

非系统码：

（4）如果两个线性码的生成矩阵通过行初等变换和列交换后能化成相同的矩阵，则称这两个**线性码等价**。

（5）现在设生成矩阵 G 行满秩，则经过有限次行初等变换和列交换后可以将它化成标准形式 (5-13)，因此任何一个唯一可译线性码都与一个系统码等价。

（6）在编码时可以将生成矩阵化成标准形状（5-13）再进行编码，就可以编出系统码。

化成等价系统码：

不失一般性，以后用系统码作为研究对象。比如在例题 5.4.5 中的码不是系统码，但是它的生成矩阵可以经过行初等变换化成标准形状：

$$\tilde{G} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix},$$

这两个矩阵生成的线性码是等价的（实际上码字空间相同），但是 $\tilde{G}$ 生成的是系统码。

定义 5.4.5：校验矩阵

(系统码校验矩阵) *按照生成矩阵 (2.2) 构造矩阵*

$$H = (A^T, I_{n-k})_{(n-k)\times n}, \tag{2.3}$$

称为以 (2.2) 为生成矩阵的系统码的校验矩阵。

容易求出例题 2.1 中的生成矩阵 G 的校验矩阵为

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

定理 5.4.1：

校验矩阵的作用由下面定理体现。

$x^{(n)}$ 是以 (2.3) 为校验矩阵生成的系统码字充要条件是

$$x^{(n)} H^T = 0. \tag{2.4}$$

证明：

如果 $x^{(n)}$ 是以 H 为校验矩阵的系统码，则

$$x^{(n)}H^T = u^{(k)}GH^T = u^{(k)}(I_k, A)\begin{pmatrix} A \\ I_{n-k} \end{pmatrix} = u^{(k)}(A \oplus A) = 0.$$

反之，如果 n 长矢量 $x^{(n)}$ 成立 $x^{(n)}H^T = 0$，则

$$x^{(n)}H^T = (u, v)\begin{pmatrix} A \\ I_{n-k} \end{pmatrix} = uA \oplus v = 0,$$

从而 $v = uA$，并且 $x^{(n)} = (u, v) = (u, uA) = u(I_k, A) = uG$，故 $x^{(n)}$ 是由 G 生成的消息 u 的系统码。

作用：

利用这个定理可以作如下判断：对于输出的一个 n 长序列 $y^{(n)}$，如果 $y^{(n)}H^T \neq 0$，则 $y^{(n)}$ 一定不是码字。因此只要条件 (5-15) 不成立，就说明传输有错，从而译码器可用校验矩阵根据条件 (5-15) 来检错，判断一个输出序列是否为码字，但是还不能检出有几位错。

求非系统码校验矩阵：

非系统码的校验矩阵可以从生成矩阵建立的码字生成关系来求。比如例题 5.4.5 中矩阵不是系统码生成矩阵，但也有校验矩阵。根据式子（5-12）消去消息字符 u_1, u_2, u_3，可以产生码字符号 $x_1, x_2, x_3, x_4, x_5, x_6$ 之间的关系式（称为**校验关系式**）

$$\begin{cases} x_1 = x_3 + x_6 \\ x_2 = x_3 + x_5 \\ x_4 = x_5 + x_6 \end{cases} \text{或} \begin{cases} x_1 + x_3 + x_6 = 0 \\ x_2 + x_3 + x_5 = 0 \\ x_4 + x_5 + x_6 = 0 \end{cases}, \tag{5.16}$$

续：求校验矩阵

用（5-16）中 x_i 矩阵系数作为元素构造矩阵

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix},$$

则有 $x^{(6)}H^T = 0$，这说明 H 就是要求的非系统码的校验矩阵。

定义 5.4.6：对偶码

（对偶码）设 H 是生成矩阵 G 的校验矩阵，则由 H 作为生成矩阵构造的 $(n, n-k)$ 码 $\mathcal{C}^{\perp}$ 称为由 G 作为生成矩阵构造的 (n, k) 码 $\mathcal{C}$ 的对偶码。

事实上 $\mathcal{C}^{\perp}$ 与 $\mathcal{C}$ 构成向量空间 $\mathcal{F}_2^n$ 的正交直和分解。码 $\mathcal{C}^{\perp}$ 的校验矩阵正好是 G。通常线性码都是成对出现的，其中一个的生成矩阵正好是另一个的校验矩阵！它们对应的码字是互相垂直的（内积为 0）！

练习：

练习：寻找一个自共轭即 $\mathcal{C}^{\perp} = \mathcal{C}$ 的线性码。

定义 5.4.7：汉明重量

（汉明重量） 一个 n 长序列 $x^{(n)}$ 的汉明重量记为 $w(x^{(n)})$ 是这个序列中非 0 元素的个数。对于二元码来说是其中 1 的个数。

定理 5.4.2

（1） 两个二元 n 长序列 $x^{(n)}, y^{(n)}$ 的汉明距离满足 $d(x^{(n)}, y^{(n)}) = w(x^{(n)} \oplus y^{(n)})$。

（2） 二元线性码 C 的最小汉明距离恰好是非 0 码字的最小汉明重量即

$$d = \min_{x^{(n)} \neq \mathbf{0}} w(x^{(n)}). \tag{2.6}$$

5.17

证明留作练习。

定理 5.4.3：

下面两个定理说明校验矩阵与纠错检错的关系。

任何 (n,k) 二元线性分组码最小汉明距离等于 d 的充要条件是校验矩阵满足

（1） *任何 $d-1$ 列线性无关。*

（2） *总有 d 列线性相关。*

证明：

必要性：用反证法证明第（1）条。假若 H 中有 $d-1$ 个列向量 $h_{i_1}, h_{i_2}, \cdots, h_{i_{d-1}}$ 线性相关，则有不全为 0 系数 $a_{i_1} a_{i_2}, \cdots, a_{i_{d-1}}$（实际上可以都是 1）使得

$$a_{i_1} h_{i_1} \oplus a_{i_2} h_{i_2} \oplus \cdots \oplus a_{i_{d-1}} h_{i_{d-1}} = 0, \tag{2.7}$$

现在利用式（2.7）中系数构造一个 n 长序列 $y^{(n)}$，它的第 $i_1, i_2, \cdots, i_{d-1}$ 位等于 $a_{i_1} a_{i_2}, \cdots, a_{i_{d-1}}$，其它位上均为 0。对于这个序列 $y^{(n)}$，显然有

$$y^{(n)} H^T = a_{i_1} h_{i_1}^T \oplus a_{i_2} h_{i_2}^T \oplus \cdots \oplus a_{i_{d-1}} h_{i_{d-1}}^T = 0,$$

根据（5-15）它是码字，但它的重量最多为 $d-1$，这与码汉明距离为 d 矛盾。因此校验矩阵中任何 $d-1$ 个列向量线性无关。

续证明：

第（2）条证明：因为码最小汉明距一定是某个码字汉明重，故必有一个码字 $x^{(n)}$ 汉明重量为 d。设码字 $x^{(n)}$ 的非 0 位为第 $i_1, i_2, \cdots, i_d$ 位，则有

$$x^{(n)} H^T = h_{i_1}^T \oplus h_{i_2}^T \oplus \cdots \oplus h_{i_d}^T = 0,$$

这说明校验矩阵 H 的第 $i_1, i_2, \cdots, i_d$ 列线性相关。

续证明：

充分性：由第（1）条可知 H 中任何 $d-1$ 列线性无关，故任何非零码字中非 0 分量的个数至少为 d 个，从而汉明距离至少为 d；再由第（2）条，可取 H 列中 d 个线性相关的列 $h_{i_1}, h_{i_2}, \cdots, h_{i_d}$ 求不全为 0 的系数 $c_{i_1}, c_{i_2}, \cdots, c_{i_d}$（实际上只能都是 1）使

$$c_{i_1} h_{i_1} \oplus c_{i_2} h_{i_1} \oplus \cdots \oplus c_{i_d} h_i = 0. \tag{2.8}$$

现在利用式（2.8）中系数构造一个 n 长序列 $y^{(n)}$，它的第 $i_1, i_2, \cdots, i_d$ 位等于 $c_{i_1}, c_{i_2}, \cdots, c_{i_d}$，其它位上均为 0。对于这个向量 $y^{(n)}$ 显然有

$$cH^T = c_{i_1} h_{i_1}^T \oplus c_{i_2} h_{i_1}^T \oplus \cdots \oplus c_{i_d} h_{i_d}^T = 0,$$

从而它是码字，并且汉明重量为 d。综上可知这个码的汉明距为 d。

定理 5.4.4：

任何 (n,k) 二元线性分组码

（1） 能检出 s 位错误的充要条件是它的校验矩阵中任何 s 列线性无关。

（2） 能纠正 t 位错误的充要条件是它的校验矩阵中任何 $2t$ 列线性无关。

（3） 汉明距离满足 $d \leq n-k+1$。

由定理 5.3.1 及定理 5.4.3 易证。

由于校验矩阵 H 如此重要，因此经常用校验矩阵来描述线性码，很少用生成矩阵。

等价类：

对于二元线性分组码，求到最小汉明距离 d 后，就可以使用最小汉明距离译码方法来译码。设 $\mathcal{C}$ 是一个 $(n.k)$ 二元线性码，则它是 n 维向量空间 $\mathcal{F}_2^n$ 的一个 k 维子空间。现在定义向量空间 $\mathcal{F}_2^n$ 上向量之间的关系

$$y^{(n)} \sim z^{(n)} \text{当且仅当} \quad y^{(n)} \oplus z^{(n)} \in \mathcal{C}. \tag{5.1}$$

可以证明它满足反身性、对称性、传递性，因此是一个等价关系。根据这个等价关系，可以将向量空间 $\mathcal{F}_2^n$ 划分成等价类的集合，使每个类中 n 长向量彼此等价。

标准阵列：

通常按如下方法来划分等价类，并排成一个表称为**标准阵列**（表 5-1）：

Step1: 标准阵列的第一行排列所有的码字，全 0 码字排在第一个，共有 2^k 个元素。

Step2: 在除去码字后的集合 $\mathcal{F}_2^n - \mathcal{C}$ 中求具有最小汉明重量的 n 长向量记为 e_1，再把它与第一行中每个码字向量相加（按位异或和），将得到的 2^k 个向量放在标准阵列的第二行。

Step3: 去掉 $\mathcal{F}_2^n$ 中位于标准阵列第一、二行中的向量，在剩余的向量中求具有最小汉明重量的向量记为 e_2，再把它与第一行中每个码字向量相加，将得到的 2^k 个向量放在标准阵列的第三行。如此进行下去，直到划去 $\mathcal{F}_2^n$ 中所有向量为止。

陪集：

这样的作法可以求出 2^r（$r = n - k$）个等价类，每个类中有 2^k 个向量。可以证明标准阵列中的每一行作为一个集合构成了向量空间 $\mathcal{F}_2^n$ 关于等价关系（5-20）的一个等价类，称为码 C 的以 e_i 为首项的**陪集**，显然陪集是互不相交的集合。

表 5-1：标准阵列表

码 $\mathcal{C}_0$	$c_0 = 00\cdots0$	c_1	c_2	$\cdots$	c_{2^k-1}
陪集 $\mathcal{C}_1$	e_1	$e_1 \oplus c_1$	$e_1 \oplus c_2$	$\cdots$	$e_1 \oplus c_{2^k-1}$
陪集 $\mathcal{C}_2$	e_2	$e_2 \oplus c_1$	$e_2 \oplus c_2$	$\cdots$	$e_2 \oplus c_{2^k-1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
陪集 $\mathcal{C}_{2^r-1}$	e_{2^r-1}	$e_{2^r-1} \oplus c_1$	$e_{2^r-1} \oplus c_2$	$\cdots$	$e_{2^r-1} \oplus c_{2^r}$

Table: 标准阵列

定理 5.4.5:

记 $\mathcal{A}_i, i = 0, 1, \cdots, 2^k - 1$ 表示标准阵列的第 i 列，则码字空间 $\mathcal{C}$ 与列向量组 $\{\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_{2^k-1}\}$ 之间存在一一映射：$g : c_i \leftrightarrow \mathcal{A}_i$。

设 n 长输出向量 $y^{(n)}$ 属于 $\mathcal{A}_i$，则 $y^{(n)}$ 与码字 c_i 的汉明距离最小。

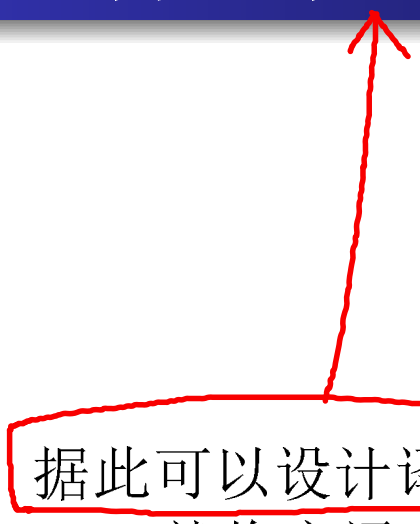
证明：

（1）如果 $y^{(n)}$ 是码字 c_i，则汉明距离为 0，直接译码。
（2）如果 $y^{(n)}$ 不是码字，设它在标准阵列第 j 行（陪集 $\mathcal{C}_j$）中，则 $y^{(n)} = e_j \oplus c_i$，可以证明 $y^{(n)}$ 与 c_i 的汉明距离是最小的，事实上：假定 c_l 与 $y^{(n)}$ 的汉明距离比 $y^{(n)}$ 与 c_i 的汉明距离还小即 $d(y^{(n)}, c_l) < d(y^{(n)}, c_i)$，则有

$$w(y^{(n)} \oplus c_l) = w(e_j \oplus c_s) < w(y^{(n)} \oplus c_i) = w(e_j \oplus c_i \oplus c_i) = w(e_j),$$

显然 $e_j \oplus c_s$ 在陪集 $\mathcal{C}_j$ 中，也在剩余集合 $\mathcal{F}_2^n - \mathcal{C}_0 - \mathcal{C}_1 - \cdots - \mathcal{C}_{j-1}$ 中，这与 e_j 是这个剩余集合中重量最小矛盾。

标准阵列译码方法：



据此可以设计译码方法：如果输出的 n 长向量 $y^{(n)}$ 属于某个 $\mathcal{A}_i$，就将它译成码字 c_i，称为**标准阵列译码方法**。这种译码方法的缺点是需要保存标准阵列，而且查找也比较费时，因此人们又将它改进提出了伴随式译码方法。

伴随式定义：

设 H 是线性码 C 的校验矩阵，定义 n 长输出向量 $y^{(n)}$ 的**伴随式**为

$$s = y^{(n)} H^T, \tag{3.2}$$

伴随式性质：

伴随式是一个 $r = n - k$ 长行向量，具有如下性质：（1）$s = \mathbf{0}$ 的充要条件是 $y^{(n)}$ 是一个码字
。（2）$y^{(n)}$ 的伴随式也可由 $y^{(n)}$ 的错误向量 e 决定。事实上

$$y^{(n)} = c_i \oplus e \Rightarrow s = y^{(n)} H^T = (c_i \oplus e) H^T = eH^T,$$

如果 e 的全部非 0 元素的下标为 $i_1, \cdots, i_t$，则

$$s = eH^T = h_{i_1}^T \oplus h_{i_2}^T \oplus \cdots \oplus h_{i_t}^T, \tag{3.3}$$

即伴随式向量 s 等于 H 中与出错位相对应的列向量之和。

求非系统码校验矩阵：

(3) 两个输出 $y^{(n)}, z^{(n)}$ 在同一陪集中的充要条件是这它们具有相同的伴随式。事实上

$$y^{(n)} \sim z^{(n)} \Leftrightarrow y^{(n)} \oplus z^{(n)} \in \mathcal{C} \Leftrightarrow (y^{(n)} \oplus z^{(n)}) H^T = 0 \Leftrightarrow y^{(n)} H^T = z^{(n)} H^T$$

因此伴随式与陪集是一一对应的，从而伴随式有 2^r 个，每个伴随式都是一个 $r = n - k$ 长的二进制向量，化成十进制，就对应从 $0 \sim 2^r - 1$ 之间的每一个整数。

伴随式表：

通常，先生成伴随式 $s = 0, 1, 2, \cdots, 2^r - 1$，并化成二进制，然后利用（3.22）来确定出它是校验矩阵的哪些列的线性组合，使错误向量 e_s 相应列元素为 1，其它元素为 0，于是就构造好了伴随式表5.2。伴随式表中第一行有些特殊，它们是码字的错误向量与伴随式，因此全为 0，即 $e = 0 \cdots 0, s_0 = 0 \cdots 0$。

错误向量 e_s	伴随式 s
e_0	0
e_1	1
e_2	2
$\vdots$	$\vdots$
e_{2^r-1}	$2^r - 1$

Table: 伴随式表

伴随式译码方法：

伴随式译码方法的基本过程是：求每个伴随式及对应的错误向量，生成一个伴随式表 2。

（1） 对信道的每个输出 $y^{(n)}$ 求出它的伴随式 $s =$ 。

（2） 查伴随式表，求出错误向量 e_s，将 $y^{(n)}$ 译成 $y^{(n)} \oplus e_s$。

例题 5.4.7：

本例说明 (6,2) 线性码的标准阵列译码与伴随式译码。该线性码生成方法为

$$\mathcal{F}_2^2 = \{00, 01, 10, 11\}, x^{(6)} = u^{(2)} \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix},$$

解：它是系统码，于是校验矩阵为

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

解：

码字空间为 $\mathcal{C} = \{000000, 010101, 101010, 111111\}$. 标准阵列总共有 16 行，按前面所说的方法可以生成标准阵列如表 5-3。易求得码距为 $d = 3$，由定理 5.3.2 它可以纠正全部 1 位错。每个码字的包含一位错的输出总共有 6 个。

码	000000	010101	101010	111111
陪集 1	000001	010100	101011	111110
陪集 2	000010	010111	101000	111101
陪集 3	000100	010001	101110	111011
陪集 4	001000	011101	100010	110111
陪集 5	010000	000101	111010	101111

Table: 表 5-3：标准阵列

续表 5-3：

陪集 6	100000	110101	001010	011111
陪集 7	000011	010110	101001	111100
陪集 8	000110	010011	101100	111001
陪集 9	001001	011100	100011	110110
陪集 10	001100	011001	100110	110011
陪集 11	010010	000111	111000	101101
陪集 12	011000	001101	110010	100111
陪集 13	100001	110100	001011	011110
陪集 14	100100	110001	001110	011011
陪集 15	110000	100101	011010	001111

Table: 表 5-3：标准阵列

当译码时，只要查输出序列 $y^{(6)}$ 所在列号，就可以进行译码。比如输出 $y^{(6)}$ 101000 时，查出它位于第 3 列，故可译成码字 101010 所对应的消息 10。

伴随式译码：

如果采用伴随式译码，则只要存储伴随式及相应的错误向量，如表 5-4。对输出 $y^{(6)}$ 译码时，只要计算它的伴随式 s，若 $s=0$ 直接译码，否则根据伴随式查找相对应的错误向量 e，从 $c=y^{(6)}\oplus e$ 则可求出相应输入码字 $x^{(n)}$。比如输出 $y^{(6)}=011001$，它的伴随式为 $s=1100$，故可译成码字 $c=y^{(6)}\oplus 001100=010101$ 所对应的消息 01。

表 5-4：伴随式表

错误向量	000000	000001	000010	000011	000100	010
伴随式（二进）	0000	0001	0010	0011	0100	01
伴随式（十进）	0	1	2	3	4	5
错误向量	001000	001001	100000	100001	001100	011
伴随式（二进）	1000	1001	1010	1011	1100	11
伴随式（十进）	8	9	10	11	12	1

Table: 伴随式译码表

5.4

定义 5.4.8：汉明码

（汉明码）二元汉明码是 (n,k) 线性分组码，其中 $n=2^r-1(r\geq 2), k=2^r-1-r$，校验矩阵由全部 $n=2^r-1$ 个 r 维非 0 向量作为列构成。

例题 5.4.8：

对于 $(7,4)$ 汉明码，它的校验矩阵由全部 7 个 3 维的非 0 列向量构成：

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} = (A^T, I_3),$$

它的生成矩阵为：

$$G = (I_4, A) = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

定理 5.4.5：汉明码性质

*（**汉明码性质**）二元汉明码最小汉明距离为 3 并且可以纠正全部 1 位错。*

证明：二元汉明码的校验矩阵 H 是由全部非 0 的 r 维向量作为列构成的，其中任何两个列向量都不相同，因此任意两列线性无关。但总有 3 个列向量线性相关，比如 $e_1, e_2, e_1 \oplus e_2$ 就是线性相关的。从而由定理 5.4.3 可知最小汉明距离为 3。再由定理 5.3.1 知它能纠正全部 1 位错。