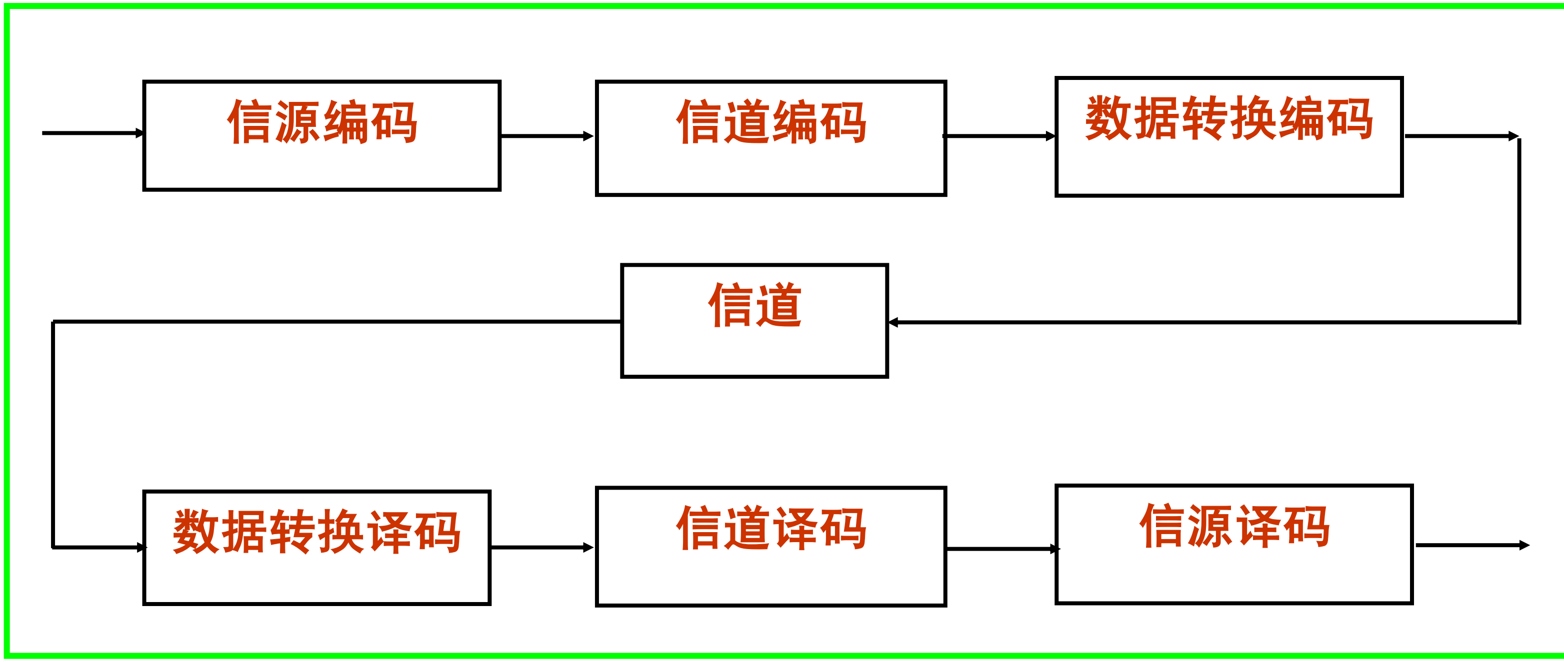
第九章 基本的信道编码技术

1、现代数字通信的两个基本理论基础：信息论和纠错编码；

2、通信中信源编码，信道编码和数据转换编码常常同时使用；



本章介绍信道编码的基本的概念、也介绍最为常用的纠错编码，即分组循环码和卷积编码。

§ 9.1 分组纠错编码的基本概念

9.1.1 用于纠错和检错的信道编码

分组信道编码器的输入是一列长度为 k 的字符序列 $\boldsymbol{m}$，其中字符是从信源字符表 M 中取值，

$$\boldsymbol{m} = (m_0, m_1, \cdots, m_{k-1}), \qquad m_i \in \mathrm{M}, \qquad i = 0, 1, 2, \cdots, k-1$$

信道编码器把输入消息序列映射成由n个信道字符组成的码字，

$$\boldsymbol{c} = (c_0, c_1, c_2, \cdots, c_{n-1}), \qquad c_i \in \mathrm{M}, \qquad i = 0, 1, 2, \cdots, n-1$$

n 一码字长度，

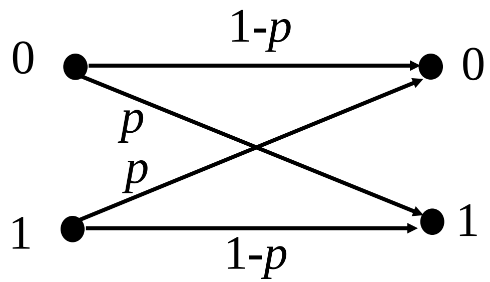
k 一信息位长度，

$r=n-k$ 一冗余位长度或称校验位长度，

码率

$$R = \frac{k}{n}$$

9.1.2 二元对称信道的差错概率和差错分布



$$P = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

信道容量 $C = 1 - H(p) = 1 + p\log p + (1-p)\log(1-p)$

T 表示码字中符号错误数目:

$$P(T=t) = C_n^t p^t (1-p)^{n-t}$$

$$P(T<t) = \sum_{j=0}^{t-1} C_n^j p^j (1-p)^{n-j}, \quad P\{T \ge t\} = 1 - P\{T<t\}$$

$$\sigma_t^2 = E\{(T-\overline{T})^2\} = np(1-p),$$

9.1.3 检错和纠错

检错是指当码字在信道上传输发生错误时，译码器能发现传输有误；纠错则是指译码器能自动纠正这个错误的能力。

检测两位错误

[例9.1.1] 3次重复编码，（$n = 3, k = 1, r = 2$ ）

“0”→“0 0 0”， “1”→“1 1 1”

接收到序列	译出的数据
0 0 0	0
0 0 1	?
0 1 0	?
1 0 0	?
0 1 1	?
1 0 1	?
1 1 0	?
1 1 1	1

纠正一位错误

接收到序列	译出的数据
0　0　0	0
0　0　1	0
0　1　0	0
0　1　1	1
1　0　0	0
1　0　1	1
1　1　0	1
1　1　1	1

[例9.1.2] $r = 3$ **的重复码，即把**

"0"→"0 0 0 0"

"1"→"1 1 1 1"

纠正一位
检测两位

接收序列	译出数据	接收序列	译出数据
0 0 0 0	0	1 0 0 0	0
0 0 0 1	0	1 0 0 1	?
0 0 1 0	0	1 0 1 0	?
0 0 1 1	?	1 0 1 1	1
0 1 0 0	0	1 1 0 0	?
0 1 0 1	?	1 1 0 1	1
0 1 1 0	?	1 1 1 0	1
0 1 1 1	1	1 1 1 1	1

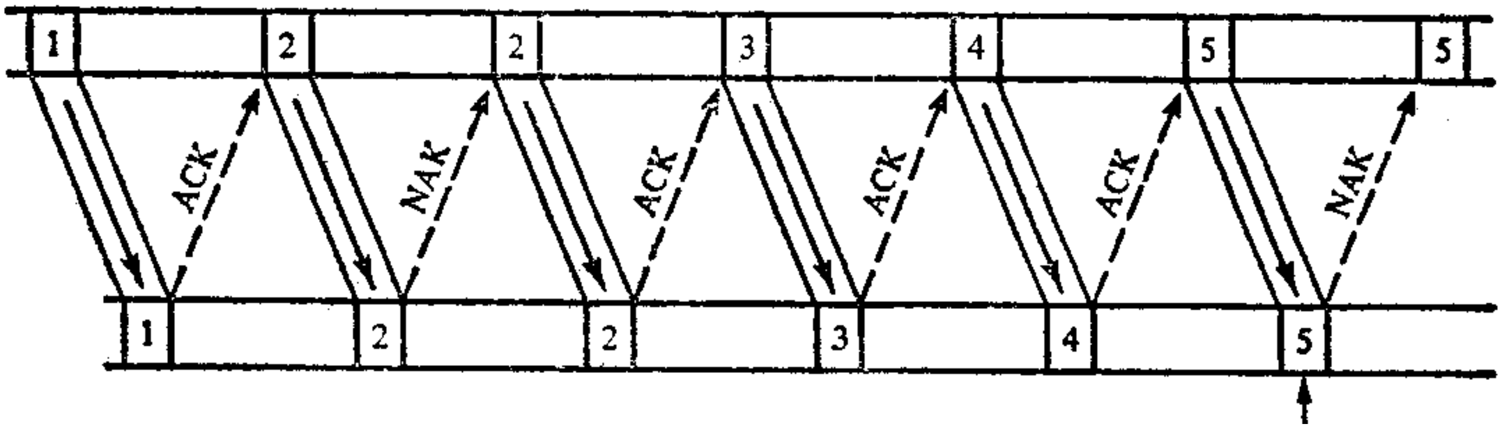
9.1.4 自动重发请求（ARQ）编码

在半双工或双工情况下，收端发现有误时，可以通过反向信道去请求对方重发一次，直到正确接收到为止。这种通过检测错误，发现错误而且自动请求重发的通信方式称为ARQ方式。

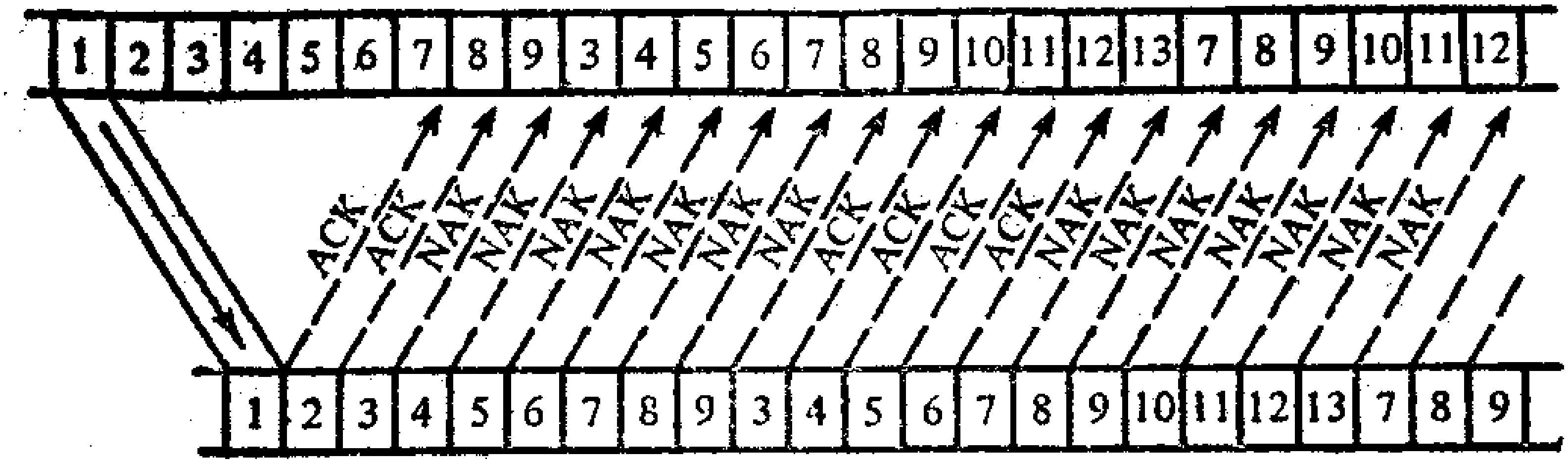
1、等待式ARQ

码字出错概率为 p，要成功传送码字，发方平均要发码字次数为：

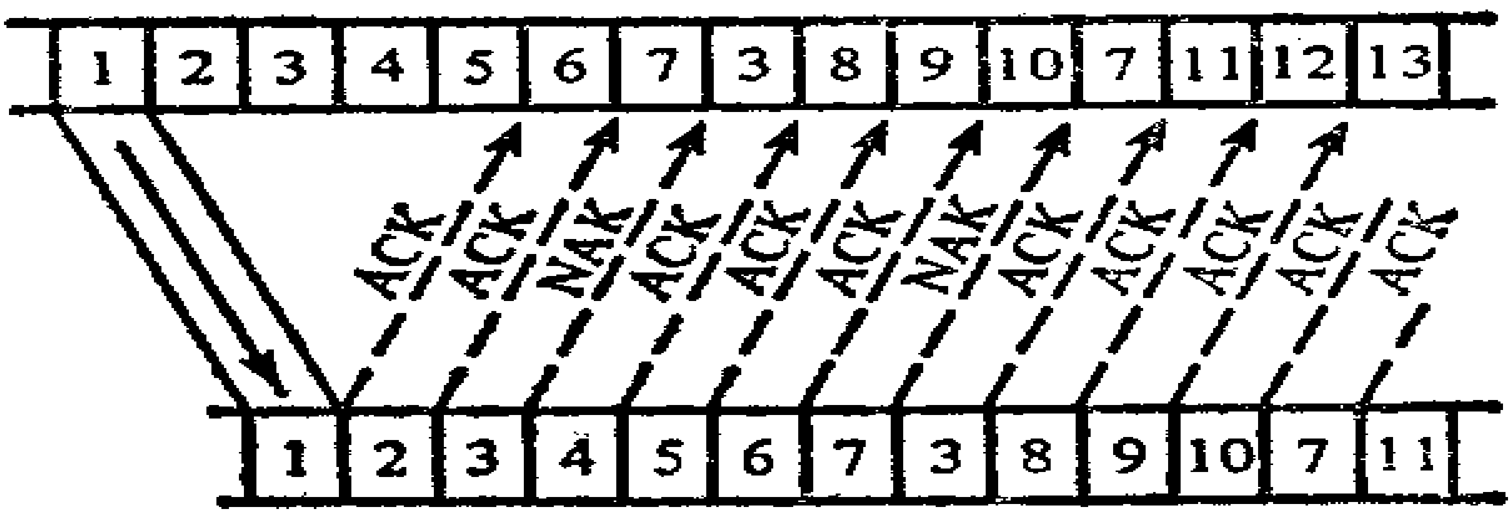
$$\overline{N} = \frac{1}{1-p}$$



2、退 *N* 步ARQ



3、选择性重发ARQ



9.1.5 最大似然译码和最小Hamming距离译码

发送码字： $\boldsymbol{c} = (c_0, c_1, \cdots, c_{n-1})$

接收序列： $\boldsymbol{r} = (r_0, r_1, \cdots, r_{n-1})$

错误矢量：

$$\boldsymbol{e} = \boldsymbol{r} - \boldsymbol{c} = (e_1, e_2, \cdots, e_{n-1})$$

$$e_i = r_i - c_i, \quad i = 0, 1, 2, \cdots, n-1$$

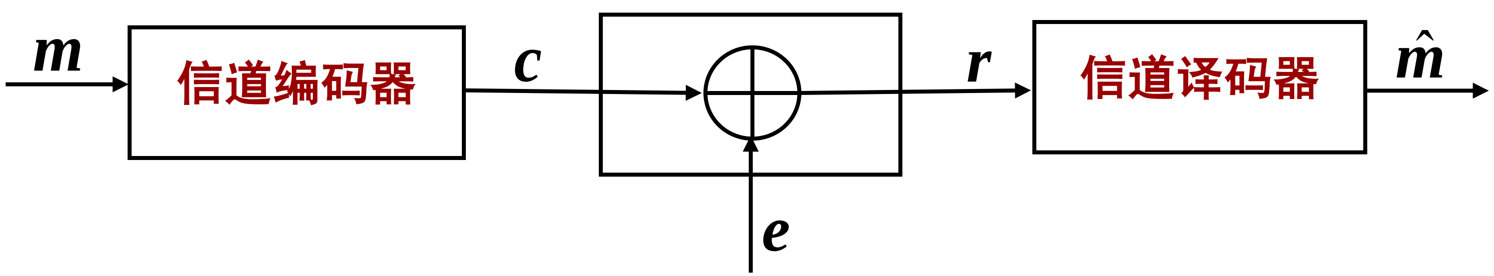
$$\boldsymbol{e} = \boldsymbol{r} + \boldsymbol{c} = (e_0, e_1, \cdots, e_{n-1})$$

二元布尔运算中，减法等同于加法

二元对称信道：

$$P(e_i = 0) = 1 - p$$

$$P(e_i = 1) = p$$



最大后验概率译码： $\hat{c} = \arg\max_{\boldsymbol{c}_i \in \mathrm{C}} P(\boldsymbol{c}_i \mid \boldsymbol{r})$

最大似然译码准则 ： $\boldsymbol{c} = \arg\max_{\boldsymbol{c}_i \in \mathrm{C}} P(\boldsymbol{r} \mid \boldsymbol{c}_i)$

先验
等概

$$P(\boldsymbol{c}_i \mid \boldsymbol{r}) = \frac{P(c_i)P(\boldsymbol{r} \mid \boldsymbol{c}_i)}{P(\boldsymbol{r})}$$

→ **最大后验概率译码＝最大似然译码**

两个序列 $\boldsymbol{r}$ **和** c **的Hamming距离为** d **指这两个序列有** d **位不同。**

$$d_H(\boldsymbol{r},\boldsymbol{c}_i) = \text{序列}\boldsymbol{r}\text{和}\boldsymbol{c}_i\text{对应不同分量位数}$$

对于差错概率为 p **的二元对称信道来说**

$$P(\boldsymbol{r} \mid \boldsymbol{c}) = p^d (1-p)^{n-d} \quad \longrightarrow \quad \log P(\boldsymbol{r} \mid \boldsymbol{c}) = n\log(1-p) + d\log\frac{p}{1-p}$$

$p < 0.5$

$$\boldsymbol{c} = \arg\max_{\boldsymbol{c}_i \in \mathrm{C}} P(\boldsymbol{r} \mid \boldsymbol{c}_i) \Leftrightarrow \boldsymbol{c} = \arg\min_{\boldsymbol{c}_i \in \mathrm{C}} d_H(\boldsymbol{r},\boldsymbol{c}_i)$$

9.1.6 最小Hamming距离与检错、纠错能力的关系

把二个长度为n 的序列$\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间的Hamming距离 $d_H(\boldsymbol{u},\boldsymbol{v})$ 定义为 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间对应分量取不同值的位数。

定义9.1.1 长度为n的分组码C 的最小Hamming距离d为

$$d = \min_{\boldsymbol{c}_i, \boldsymbol{c}_j \in \mathrm{C}, i \neq j} d_H(\boldsymbol{c}_i, \boldsymbol{c}_j)$$

分组码的三个最重要参数:

码字长度n，信息位数目k，最小Hamming距离d;

对于一个（$\boldsymbol{n}, \boldsymbol{k}$）分组码来说，**最小Hamming距离** d 与纠错、检错能力有如下关系：

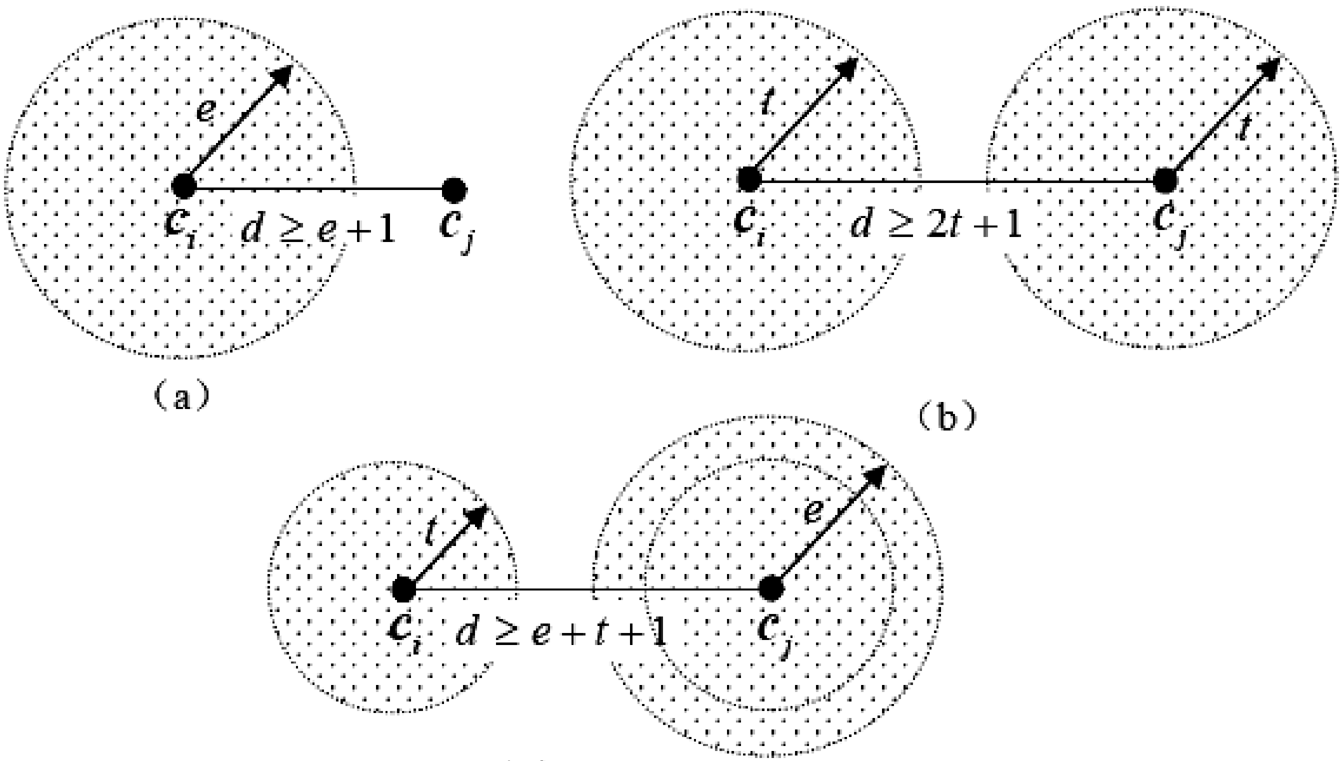
定理9.1.1 任何一个（n, k）分组码，若要在任何码字内

a. 能检测e个随机错误，则要求最小Hamming距离$d \geq e+1$。

b. 能纠正t个随机错误，则要求$d \geq 2t+1$。

c. 能纠正t个随机错误，同时检测出e ($\geq t$)个错误，则要求$d \geq t+e+1$。

[证明]



（a） （b） （c）

§ 9.2 线性分组纠错编码

9.2.1 线性分组编码的生成矩阵和校验矩阵

线性分组码的基本特征是它具有“线性”的结构， 即两个码字的线性组

合仍是码字，对于二元码，即两个码字之和仍为码字。可以利用数学

工具“线性空间”来研究线性分组码，使得编码器和译码器的实现更为简单。

定义9.2.1 一个速率为 $R=k/n$ 的线性分组码（n, k），把 k 比特的消息

矢量 $\boldsymbol{m}=(m_0, m_1, \cdots, m_{k-1})$ 线性地映射成 n 比特的码字 $\boldsymbol{c}=(c_0, c_1, \cdots, c_{n-1})$

其中 $m_i \in \{0,\ 1\}, i=0,\ 1,\cdots,k-1,\quad c_j \in \{0,\ 1\}, j=0,\ 1,\cdots,n-1$。

线性映射：

$$\begin{matrix} \boldsymbol{m}_1 \rightarrow \boldsymbol{c}_1 \\ \boldsymbol{m}_2 \rightarrow \boldsymbol{c}_2 \end{matrix} \quad \Longrightarrow \quad \boldsymbol{m}_1+\boldsymbol{m}_2 \rightarrow \boldsymbol{c}_1+\boldsymbol{c}_2$$

全体码字集合称为码字空间 C，它是 n 维空间中的一个 k 维子空间。

$$\left.\begin{aligned} \boldsymbol{g}_0 &= (g_{00}, g_{01}, \cdots, g_{0,n-1}) \\ \boldsymbol{g}_1 &= (g_{10}, g_{11}, \cdots, g_{1,n-1}) \\ &\cdots\cdots \\ \boldsymbol{g}_{k-1} &= (g_{k-10}, g_{k-11}, \cdots, g_{k-1,n-1}) \end{aligned}\right\}$$ k个线性独立的二元n维矢量

消息矢量

所以 $\boldsymbol{c} = m_0\boldsymbol{g}_0 + m_1\boldsymbol{g}_1 + \cdots + m_{k-1}\boldsymbol{g}_{n-1} = \boldsymbol{m}\cdot G,\ \forall \boldsymbol{c} \in \mathrm{C}$

$$\boldsymbol{m} = (m_0, m_1, \cdots, m_{k-1})$$

$$G = \begin{pmatrix} \boldsymbol{g}_0 \\ \boldsymbol{g}_1 \\ \vdots \\ \boldsymbol{g}_{k-1} \end{pmatrix}$$

生成矩阵

[例9.2.1] G 生成的（6, 3）线性码

$$G = \begin{array}{c} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array} \begin{array}{l} \cdots \boldsymbol{g}_0 \\ \cdots \boldsymbol{g}_1 \\ \cdots \boldsymbol{g}_2 \end{array}$$

$$\boldsymbol{c} = m_0\boldsymbol{g}_0 + m_1\boldsymbol{g}_1 + m_2\boldsymbol{g}_2,\ m_i \in \{0,1\},\ i = 1,2,3$$

$$\boldsymbol{m} = (0\ \ 1\ \ 1) \Longrightarrow \boldsymbol{c} = \boldsymbol{g}_1 + \boldsymbol{g}_2 = (110110)$$

行变换
列交换

系统生
成矩阵

$$G \Rightarrow G = [P_{k\times r} \;\vdots\; I_{k\times k}]$$

系统生成矩阵生成系统线性码：

$$\boldsymbol{c} = \boldsymbol{m}\cdot G = (\underbrace{c_0, c_1, \cdots c_{r-1}}_{r=n\text{-}k\ \text{校验位}}, \underbrace{m_0, m_1 \cdots, m_{k-1}}_{k\ \text{信息位}})$$

系统线性码的校验矩阵：

$$H = [I_{r\times r} \;\vdots\; -(P_{k\times r})^T]$$

$$\boldsymbol{c}\cdot H^T = \boldsymbol{m}\cdot G\cdot H^T = (\underbrace{0, 0, \cdots, 0}_{r\text{位}}),\ \ \forall \boldsymbol{c} \in \mathrm{C}$$

[例9.2.2] **例9.2.1中的（6.3）线性码的校验矩阵为**

$$H=\left[\begin{array}{c:c} 100 & 110 \\ 010 & 101 \\ 001 & 011 \end{array}\right]$$

$$\boldsymbol{m}=(m_0,m_1,m_2) \Longrightarrow \boldsymbol{c}=(c_0,c_1,c_2,m_0,m_1,m_2)$$

$$\boldsymbol{c}\cdot H^T=0 \Longrightarrow \begin{cases} c_0+m_0+m_1=0 \\ c_1+m_0+m_2=0 \\ c_2+m_1+m_2=0 \end{cases}$$

9.2.2 对偶码

$\boldsymbol{u}=(u_0,u_1,\cdots,u_{n-1})$ 和 $\boldsymbol{v}=(v_p,v_r,\cdots,v_{n-1})$内积：

$$\boldsymbol{u}\cdot\boldsymbol{v}=u_0v_0+u_1v_1+\cdots+u_{n-1}v_{n-1}$$

*u*和*v*正交： $\boldsymbol{u}\cdot\boldsymbol{v}=0$

n 维空间中与一个k维子空间C 正交的所有矢量的全体构成一个n-k 维子空间，它称为C 的对偶空间，或称C 的正交补，用$C^{\perp}$来表示。

定义9.2.2 生成矩阵为G，校验矩阵为H 的（n, k）线性分组码C 的对偶码$C^{\perp}$是一个生成矩阵为H，校验矩阵为G的（n, n–k）线性分组码。

9.2.3 线性分组码的最小Hamming距离和最小Hamming重量

定义9.2.3 一个n维矢量的Hamming重量$w_H(\boldsymbol{v})$定义为该矢量中非零分量的个数，对于二元矢量也就是矢量中“1”分量的个数。

因为 $\boldsymbol{c}_1, \boldsymbol{c}_2 \in \mathrm{C} \Rightarrow \boldsymbol{c}_1 + \boldsymbol{c}_2 \in \mathrm{C}$

所以
$$d_{\min} = \min_{\substack{\boldsymbol{c}_i \neq \boldsymbol{c}_j \\ \boldsymbol{c}_i, \boldsymbol{c}_j \in \mathrm{C}}} d_H(\boldsymbol{c}_i, \boldsymbol{c}_j) = \min_{\substack{\boldsymbol{c}_i \neq \boldsymbol{c}_j \\ \boldsymbol{c}_i, \boldsymbol{c}_j \in \mathrm{C}}} w_H(\boldsymbol{c}_i + \boldsymbol{c}_j)$$
$$= \min_{\substack{\boldsymbol{c} \neq 0 \\ \boldsymbol{c} \in \mathrm{C}}} w_H(\boldsymbol{c})$$

线性分组码的最小Hamming距离等于该码中非零码字的最小重量。

[例9.2.5] 由例 9.2.3中生成矩阵所生成的线性分组码总共有8个码字

（110100），（101010），（011001），（000000），
（011110），（110011），（101101），（000111）

设 $\boldsymbol{c}=(c_0,c_1,\cdots,c_{n-1}),\ \boldsymbol{c}\in \mathrm{C} \Longrightarrow \boldsymbol{c}\cdot H^T=0$

$$\sum_{j=0}^{n-1} c_i \boldsymbol{h}_i = 0,\ h_i\text{为}H\text{列矢量}$$

$$H=\begin{pmatrix} h_{00} & h_{01} & h_{02} & \cdots & h_{0i} & \cdots & h_{0,n-1} \\ h_{10} & h_{11} & h_{12} & \cdots & h_{1i} & \cdots & h_{1,n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{k-1,0} & h_{k-1,1} & h_{k-1,2} & \cdots & h_{k-1,i} & \cdots & h_{k-1,n-1} \\ \uparrow & \uparrow & \uparrow & & \uparrow & & \uparrow \\ c_0 & c_1 & c_2 & & c_i & & c_{n-1} \end{pmatrix}$$

若码字重量为 w，则相应 w 列的和为零。

如果一个线性分组码的最小Hamming距离为d，也就是说该码的最小Hamming重量为d，则它的校验矩阵H中任意$d-1$个列矢量是线性独立的。

9.2.4 线性分组码的译码

发送的二元码字矢量为 $\boldsymbol{c}$，从信道接收到的二元矢量为 $\boldsymbol{v}$，错误矢量为

$$\boldsymbol{e}=\boldsymbol{v}-\boldsymbol{c}=(e_0,e_1,\cdots,e_{n-1})$$

最大似然译码法则要求把 $\boldsymbol{v}$ 译成与之距离最近的码字。

完全译码器把接收到的二元矢量$\boldsymbol{v}$译成与它最近码字$\boldsymbol{c}$；

限定距离t 译码器，选与$\boldsymbol{v}$最近的码字 $\boldsymbol{c}$，当 $d_H(\boldsymbol{c},\boldsymbol{v})\le t$ 时，则译码器就译成$\boldsymbol{c}$；当$d_H(\boldsymbol{c},\boldsymbol{v})>t$ 时，译码器声称纠错失败。

一、标准阵列译码法

陪集首项：n 维矢量中除了前面行矢量外最轻重量者

$$\begin{array}{lllll}
\boldsymbol{c}_0=(0,\ 0,\ \cdots,\ 0), & \boldsymbol{c}_1, & \boldsymbol{c}_2, & \cdots, & \boldsymbol{c}_{2^k-1} \\
\boldsymbol{e}_1 & \boldsymbol{e}_1+\boldsymbol{c}_1, & \boldsymbol{e}_1+\boldsymbol{c}_2, & \cdots, & \boldsymbol{e}_1+\boldsymbol{c}_{2^k-1} \\
\boldsymbol{e}_2 & \boldsymbol{e}_2+\boldsymbol{c}_1, & \boldsymbol{e}_2+\boldsymbol{c}_2, & \cdots, & \boldsymbol{e}_2+\boldsymbol{c}_{2^k-1} \\
\cdots & \cdots & \cdots & \cdots, & \cdots \\
\boldsymbol{e}_{2^r-1} & \boldsymbol{e}_{2^r-1}+\boldsymbol{c}_1, & \boldsymbol{e}_{2^r-1}+\boldsymbol{c}_2, & \cdots, & \boldsymbol{e}_{2^r-1}+\boldsymbol{c}_{2^k-1}
\end{array}$$

二元（n, k）线性码C，若 a 是任意一个非码字n维矢量，则称集合 $\boldsymbol{a}+\mathrm{C}=\{\boldsymbol{a}+\boldsymbol{c};\boldsymbol{c}\in\mathrm{C}\}$为$C$ 的一个陪集，其中a 称为陪集首项。任意二个倍集或者不相交或者重合，所以**标准阵列是线性码的完全陪集分解。**

对于任何接收到的矢量$\boldsymbol{v}$，若它落在标准阵列中的第j行，则**可能的错误形式是该行中的所有矢量**，最大似然译码准则要求把与接收矢量最近的码字译为发送码字，所以相当于在第j行中寻找最轻重量的矢量作为错误形式，即把**该行的首项作为错误形式。**

对于**限定矩离 t 译码**来说，不需要构造出完整的标准阵列，只需要构造重量不大于 t 的陪集首项所对应的陪集。若接收到的矢量没有出现在这个不完整阵列表中，则说明这时发生了不可纠正的错误。

[例9.2.5] 一个具有4个码字，能纠错一位的（6，2）线性码，

$$C=\{(000000), (010101), (101010), (111111)\}$$

陪集首项重量不大于1的不完全阵列表

0 0 0 0 0 0	0 1 0 1 0 1	1 0 1 0 1 0	1 1 1 1 1 1
0 0 0 0 0 1	0 1 0 1 0 0	1 0 1 0 1 1	1 1 1 1 1 0
0 0 0 0 1 0	0 1 0 1 1 1	1 0 1 0 0 0	1 1 1 1 0 1
0 0 0 1 0 0	0 1 0 0 0 1	1 0 1 1 1 0	1 1 1 0 1 1
0 0 1 0 0 0	0 1 1 1 0 1	1 0 0 0 1 0	1 1 0 1 1 1
0 1 0 0 0 0	0 0 0 1 0 1	1 1 1 0 1 0	1 0 1 1 1 1
1 0 0 0 0 0	1 1 0 1 0 1	0 0 1 0 1 0	0 1 1 1 1 1
… … … …	… … … …	… … … …	… … … …

如果一个线性码的标准阵列中的陪集首项正好是所有重量不大于 t 的二元矢量，该线性码称为完备码。

二、伴随式译码

接收矢量 $\boldsymbol{v}$ 所对应的伴随式 $\boldsymbol{s}$ 是一个 $r=n-k$ 维矢量 $\boldsymbol{s}=\boldsymbol{v}\cdot H^{T}$

1、与 $\boldsymbol{v}$ 相应的伴随式 $\boldsymbol{s}$ 为零矢量的充要条件是 $\boldsymbol{v}$ 为一个码字；

2、若 $$\boldsymbol{e}=0,\ \cdots,\ 0,\ \underset{\substack{\uparrow\\ \text{第a位}}}{1},\ 0,\ \cdots,\ \underset{\substack{\uparrow\\ \text{第b位}}}{1},\ \cdots,\ \underset{\substack{\uparrow\\ \text{第c位}}}{1},\ \underset{\cdots}{\cdots},\ 0$$

则 $$\boldsymbol{s}=\sum_{i} e_i\cdot\boldsymbol{h}_i=\boldsymbol{h}_a+\boldsymbol{h}_b+\boldsymbol{h}_c+\cdots$$

3、二个矢量出现在C 的同一陪集中的充要条件是他们具有相同的伴随式；所以伴随式与陪集一一对应。

如果接收到矢量为$\boldsymbol{v}$，首先计算出它的伴随式$\boldsymbol{s}$，如果 $\boldsymbol{s}=\boldsymbol{0}$，则表示接收到的是码字，没有错。如果不为$\boldsymbol{0}$，则根据 $\boldsymbol{s}$ 查出对应的错误形式。

9.2.5 译码错误概率计算

误码字率

$$P_{EB}=\frac{1}{M}\sum_{i=1}^{M}P\{\text{译码输出}\neq \boldsymbol{c}_i\,|\,\boldsymbol{c}_i\text{被发送}\}$$
$$=\frac{1}{M}\sum_{i=1}^{M}P\{\text{错误形式}\neq\text{陪集首项}\,|\,\boldsymbol{c}_i\text{被发送}\}$$
$$=P\{\text{错误形式}\neq\text{陪集首项}\}$$

$$P_{EB}=1-\sum_{i=0}^{n}\alpha_i p^i(1-p)^{n-i}$$

α_i为重量为i 的陪集首项数目;

误比特率

$$\frac{P_{EB}}{k}\leq P_{eb}\leq P_{EB}$$

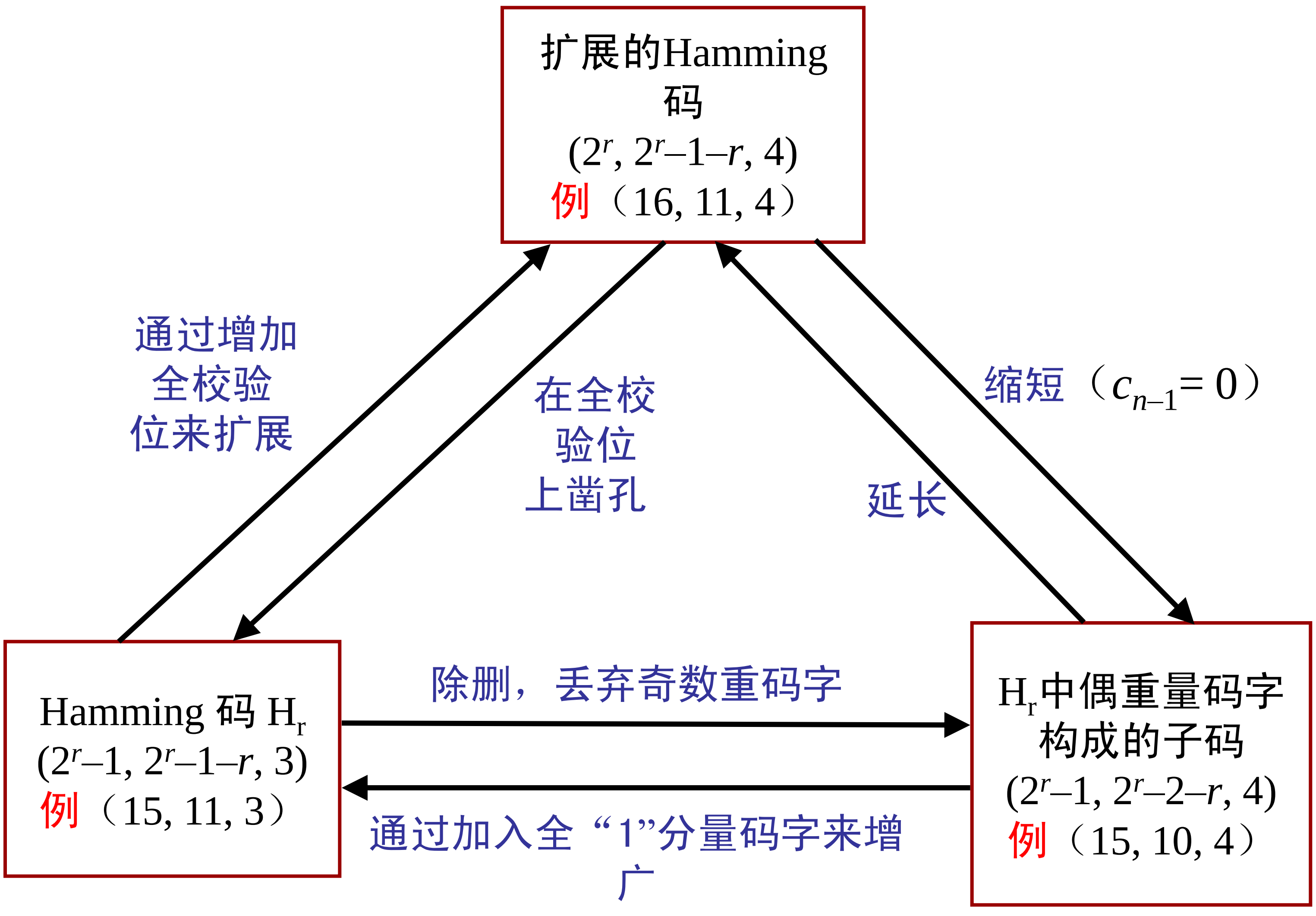
9.2.6 二元Hamming码

定义9.2.4 长度为$n = 2^r-1$（r≥2）的二元Hamming码是一个（$n=2^r-1$, $k=2^r-1-r$）线性分组码，它的最小Hamming距离为3，能纠正全部一位错误。它的校验矩阵H由全部2^r-1个长度为r的非零、相异的二元列矢量组成。

[例9.2.7] 对于系统（7.4）Hamming码，它的校验矩阵

$$H = \left[\begin{array}{ccc:cccc} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{array}\right] \qquad G = \left[\begin{array}{ccc:cccc} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

9.2.7 从一个已知线性分组码来构造一个新的线性分组码



§ 9.3 线性分组码的纠错能力

线性码的纠错能力是由给定 n, k 条件下，最小距离 d 的上，下界限来表征的。下面3个定理给出关于最小距离 d 的上限。

定理9.3.1 （Singleton限）

任何线性（n, k）码的最小Hamming距离 d 满足

$$d \le n - k + 1$$

定理9.3.2 （Hamming限）

长度为n，能纠正 t 个错误的二元分组码所含有码字数M必须满足

$$M \le 2^n / \sum_{i=0}^{t} C_n^i$$

定理9.3.3 （Plotkin限）

长度为n，码字数为M的分组码，它的最小Hamming距离d必须满足

$$d \le \frac{nM}{2(M-1)}$$

定理9.3.4 （Varsharmov-Gilbert下限）

可以构成一个最小距离为d的（n, k）线性分组码，其中参数n, k, d满足

$$n-k>\log_2\sum_{j=0}^{d-2}C_{n-1}^{j}$$