信息论与编码

Information Theory and Coding

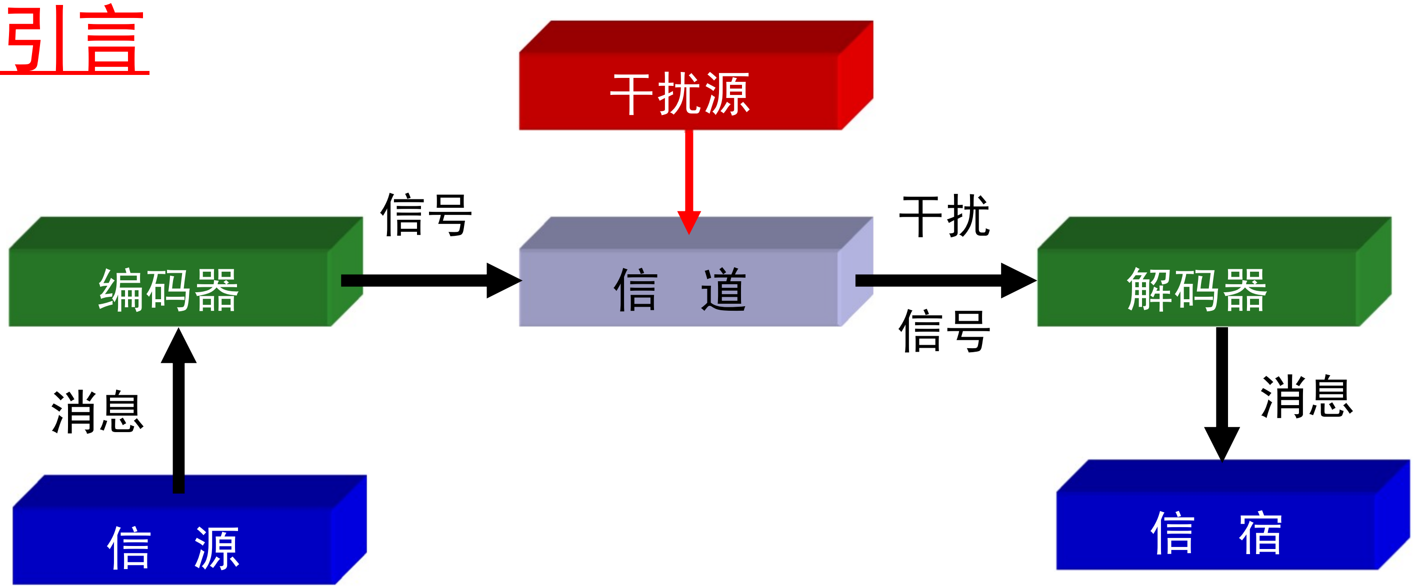
西南交通大学

电子信息工程专业

2020

第5章 信道编码

5-1 引言

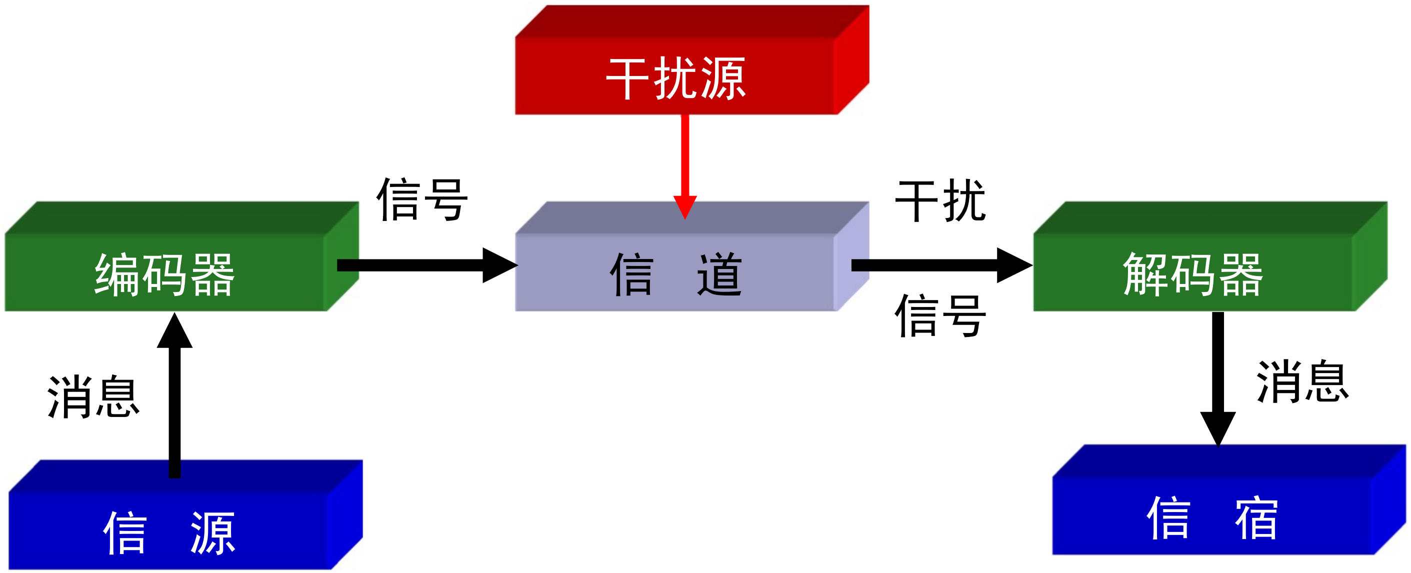


问题:

(1) 怎样使消息通过有噪声信道后发生最少的错误

(2) 在有噪声信道中无错误传输的可达的最大信息传输率是多少

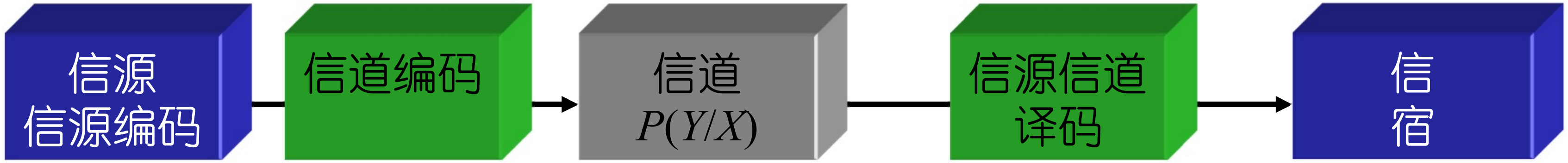
5-1 引言



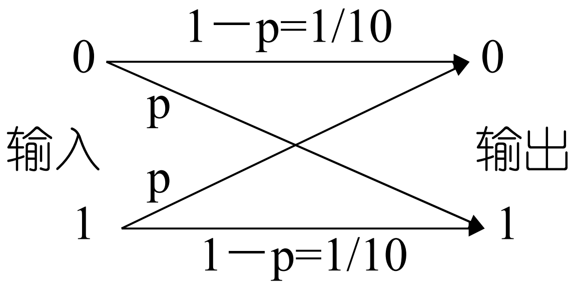
(1) 对无失真信源编码的码字，用有噪声信道的输入符号集作为码符号集，再进行一次编码提高其抗干扰能力

(2) 利用和挖掘信道的统计特性，保持一定**有效性**的基础上，提高其抗干扰的**可靠性**（有噪信道编码定理）

5-2 译码规则和错误概率



图示：二进制对称信道



译码规则1:(输入端先验等概条件下)

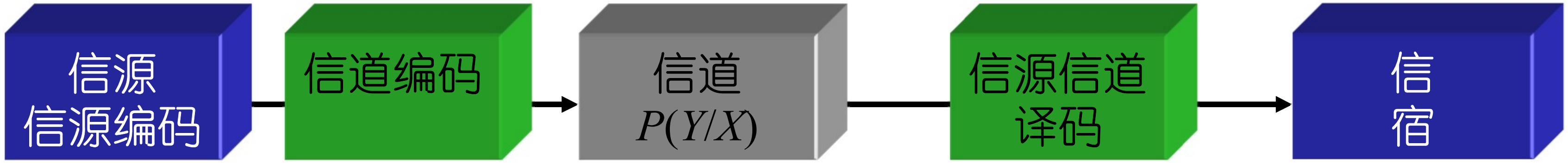
输出端“0”，认为输入端输入“0”，译码为“0”

正确译码概率：$P(X=0/Y=0)=\dfrac{P(Y=0/X=0)P(X=0)}{P(Y=0)}=1/10$

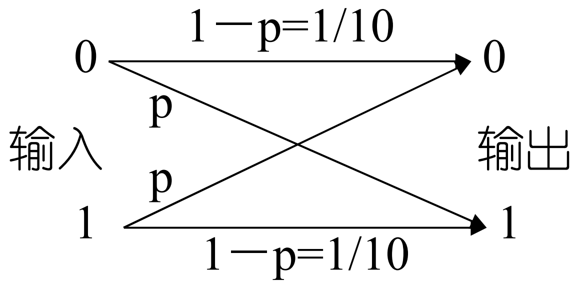
输出端“1”，认为输入端输入“1”，译码为“1”

正确译码概率：$P(X=1/Y=1)=\dfrac{P(Y=1/X=1)P(X=1)}{P(Y=1)}=1/10$

5-2 译码规则和错误概率



图示：二进制对称信道



译码规则2:(输入端先验等概条件下)

输出端“0”，认为输入端输入“1”，译码为“1”

正确译码概率：$P(X=1/Y=0)=\dfrac{P(Y=0/X=1)P(X=1)}{P(Y=0)}=9/10$

输出端“1”，认为输入端输入“0”，译码为“0”

正确译码概率：$P(X=0/Y=1)=\dfrac{P(Y=1/X=0)P(X=0)}{P(Y=1)}=9/10$

5-2 译码规则和错误概率

结论：

错误概率既与信道的统计特性（传递概率）有关，又与译码规则有关；

因此在一个完整的通信过程中，译码规则对通信的可靠性有重大的影响

5-2 译码规则和错误概率

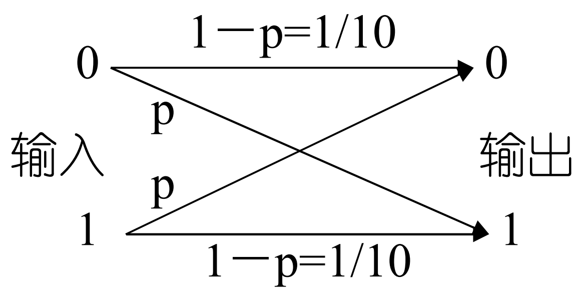
5.2.1 译码规则

定义：设有噪声信道，输入符号集$X:\{a_1,a_2,\cdots,a_r\}$，输出符号集$Y:\{b_1,b_2,\cdots,b_s\}$，信道传递概率$P(Y/X):\{p(b_j/a_i),i=1,2\cdots,r;j=1,2,\cdots,s\}$，译码规则是译码判决中的法规，它是由$s$个单值函数$F(b_j)$组成，而$F(b_j)$对于每一个输出符号$b_j$能确定一个唯一的输入符号$a_i$

即　$F(b_j)=a_i,\quad i=1,2,\cdots,r\quad j=1,2,\cdots,s$

5-2 译码规则和错误概率

例如： 图示BSC信道，输入符号集X:{0,1},输出符号集Y:{0,1},可以组成 r^2=4种译码规则:



$$\begin{cases} F(0)=0 \\ F(1)=0 \end{cases} \quad \begin{cases} F(0)=1 \\ F(1)=0 \end{cases} \quad \begin{cases} F(0)=0 \\ F(1)=1 \end{cases} \quad \begin{cases} F(0)=1 \\ F(1)=1 \end{cases}$$

问题：面对 r^2 (r^s) 种译码规则，应该怎样选取？

5-2 译码规则和错误概率

5.2.2 错误概率

定义： 在信道的输出端收到某符号$b_j(j=1,2,\cdots,s)$后，正确译码的概率p_{rj}就是信道输出端出现b_j的前提下，推测信道的输入符号是$a_i(i=1,2\cdots,r)$的后验概率，即$p_{rj}=p(X=a_i/Y=b_j)$；同样，在信道的输出端收到某符号b_j后，错误译码的概率p_{ej}就是信道输出端出现b_j的前提下，推测信道的输入符号是a_i以外的其它输入符号的后验概率，即

$$\begin{aligned}p_{ej}&=p(X\neq a_i/Y=b_j)=1-p_{rj}\\&=1-p(X=a_i/Y=b_j)\end{aligned}$$

5-2 译码规则和错误概率

定义：

在信道的输出端每收到一个符号Y的平均正确译码的概率P_r就是，

$$P_r=\sum_{j=1}^{s} p(b_j)p_{rj}=\sum_{j=1}^{s} p(b_j)p(X=a_i/Y=b_j)$$

$$=\sum_{j=1}^{s} p(b_j)p(X=F(b_j)=a_i/Y=b_j)$$

即对于每一个符号b_j正确译码概率在信道输出随机变量Y的概率空间 $P(Y):\{p(b_1),p(b_2),\ldots,p(b_s)\}$中的统计平均值.

5-2 译码规则和错误概率

定义： 在信道的输出端每收到一个符号Y的平均错误译码的概率P_e就是，

$$P_e=\sum_{j=1}^{s} p(b_j)(1-p_{rj})=\sum_{j=1}^{s} p(b_j)[1-p(X=a_i/Y=b_j)]$$

$$=\sum_{j=1}^{s} p(b_j)-\sum_{j=1}^{s} p(b_j)p(X=F(b_j)=a_i/Y=b_j)=1-P_r$$

平均错误译码概率表示，在信道的输出端，平均每收到一个符号，产生错误译码的可能性的大小。通常将**平均错误译码概率**作为衡量通信**可靠性**的标准。

5-2 译码规则和错误概率

回顾：

信道输入随机变量X的概率空间
$P(X):\{p(a_1),\ldots,p(a_s)\}$

信道传递概率固定
$P(Y/X):\{p(b_j/a_i)\}$

信道后验概率分布
$P(X/Y):\{p(a_i/b_j)\}$

信道输出随机变量Y的概率空间
$P(Y):\{p(b_1),\ldots,p(b_s)\}$

选择合适的译码规则，成为降低平均错误译码概率，提高通信有效性的一种可控制的有效手段
$F(b_j)=a_i$

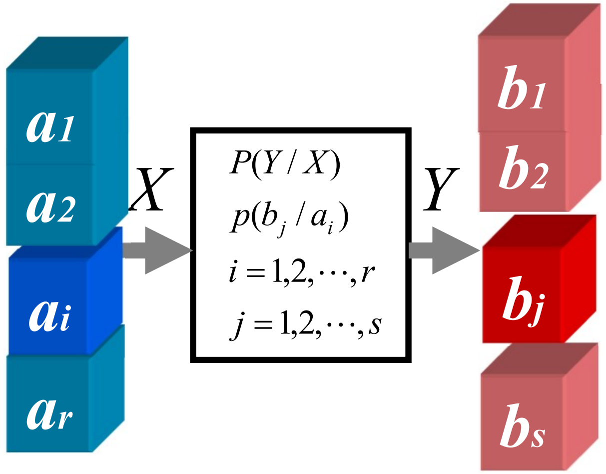
平均错误译码概率
P_e

当信源给定、信道给定时，按照什么准则来选择合适的译码规则，使平均错误译码概率达到最小，是提高通信系统可靠性的关键问题。

5-2 译码规则和错误概率

5.2.3 译码规则的选择

$$\mathbf{P}=\begin{bmatrix} p(b_1/a_1) & p(b_2/a_1) & \cdots & p(b_s/a_1) \\ p(b_1/a_2) & p(b_2/a_2) & \cdots & p(b_s/a_2) \\ \vdots & \vdots & \vdots & \vdots \\ p(b_1/a_r) & p(b_2/a_r) & \cdots & p(b_s/a_r) \end{bmatrix}$$



$$p(b_j)=\sum_X p(b_j/a_i)p(a_i)$$

$$p(a_i/b_j)=\frac{p(b_j/a_i)p(a_i)}{p(b_j)}=\frac{p(b_j/a_i)p(a_i)}{\sum_X p(b_j/a_i)p(a_i)}$$

$$\mathbf{P}'=\begin{bmatrix} p(a_1/b_1) & p(a_1/b_2) & \cdots & p(a_1/b_s) \\ p(a_2/b_1) & p(a_2/b_2) & \cdots & p(a_2/b_s) \\ \vdots & \vdots & \vdots & \vdots \\ p(a_r/b_1) & p(a_r/b_2) & \cdots & p(a_r/b_s) \end{bmatrix}$$

后验概率矩阵

5-2 译码规则和错误概率

定义： **最大后验概率译码准则**：选用译码函数$F(b_j)=a_i$，它对应于将每一个输出符号均译成具有最大后验概率的那个输入符号，那么信道的平均错误译码概率就能达到最小$P_{e\min}$，即选择译码函数：

$F(b_j)=a^*, a^*\in X, b_j\in Y$,

且$p(a^*/b_j)\geq p(a_i/b_j), a_i\neq a^*$

由定义得到：

由于$p(a^*/b_j)\geq p(a_i/b_j)$

简化式：

$$p(a^*/b_j)=\frac{p(b_j/a^*)p(a^*)}{p(b_j)}\geq\frac{p(b_j/a_i)p(a_i)}{p(b_j)}=p(a_i/b_j)$$

得到$p(b_j/a^*)p(a^*)\geq p(b_j/a_i)p(a_i)$

5-2 译码规则和错误概率

在特定条件下简化最大后验概率译码准则

因此，考虑在信道的输入符号先验等概的条件下，最大后验概率译码准则转变为：

定义：对于信道的每一个输出符号b_j，在输入先验等概的条件下，如果$p(b_j/a^*) \geq p(b_j/a_i), a_i \neq a^*$

那么选择译码函数$F(b_j)=a^*$，即将符号b_j译成a^*

将以上译码规则称为最大似然译码准则。它是在输入等概的条件下，最大后验译码准则的特例。

5-2 译码规则和错误概率

注意:

最大似然译码准则本身，并不依赖于先验概率$p(a_i)$，但是只有当先验概率为等概分布时，才可以使平均错误译码概率最小。如果先验概率不等或者不知道时，虽然可以使用这个译码准则，但是不能使平均错误译码概率最小。

5-2 译码规则和错误概率

根据**最大后验概率译码准则**和**最大似然译码准则**，研究P_e

$$P_{e\min}=\sum_{j=1}^{s}p(b_j)[1-p(F(b_j)=a^*/Y=b_j)]$$

$$=\sum_{j=1}^{s}p(b_j)-\sum_{j=1}^{s}p(b_j)p(a^*/b_j)=1-\sum_{j=1}^{s}p(b_j)p(a^*/b_j)$$

$$=\sum_{i=1}^{r}\sum_{j=1}^{s}p(a_ib_j)-\sum_{j=1}^{s}p(a^*b_j)$$

$$=\sum_{j=1}^{s}\sum_{i\neq *}p(a_ib_j)=\sum_{j=1}^{s}\sum_{i\neq *}p(a_i)p(b_j/a_i)$$

在$p(a_i)=1/r$条件下（等概）

$$P_{e\min}=\frac{1}{r}\sum_{j=1}^{s}\sum_{i\neq *}p(b_j/a_i),i=(1,2,\cdots r),j=(1,2,\cdots s)$$

5-2 译码规则和错误概率

[例题]

某信道矩阵为$P=\begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.3 & 0.5 \\ 0.3 & 0.3 & 0.4 \end{bmatrix}$,如果

(1)信道输入符号a_i先验等概，试选择译码规则，使平均错误译码概率P_e达到最小P_{emin}，并计算P_{emin}

(2)信道输入符号a_i先验概率不相等时，

$p(a_1)=1/4$，$p(a_2)=1/2$，$p(a_3)=1/4$时，

试选择译码规则，使平均错误译码概率P_e达到最小P_{emin}，并计算P_{emin}

5-2 译码规则和错误概率

解：（1）. 因信道输入符号先验等概，即$p(a_1)=p(a_2)=p(a_3)=1/3$，故采用最大似然准则选择译码规则，即若

$$p(b_j/a^*)\ge p(b_j/a_i) \qquad (i,j=1,2,3)$$

则选择译码规则：

$$F(b_j)=a^* \qquad (j=1,2,3;a^*\in\{a_1,a_2,a_3\})$$

(i)对于信道输出符号b_1而言，信道矩阵P中第一列元素：

$$p(b_1/a_1)=0.5;\quad p(b_1/a_2)=0.2;\quad p(b_1/a_3)=0.3$$

其中最大者为$p(b_1/a^*)\ge p(b_1/a_1)=0.5$，则选择译码规则$F(b_1)=a^*=a_1$

5-2 译码规则和错误概率

(ii)对于信道输出符号b_2而言，信道矩阵P中第二列元素：

$$p(b_2/a_1)=0.3;\quad p(b_2/a_2)=0.3;\quad p(b_2/a_3)=0.3$$

由于三个值相等，故选择译码规则$F(b_2)=a_1$；

$F(b_2)=a_2$；$F(b_2)=a_3$中任何一种均可.

(iii)对于信道输出符号b_3而言，信道矩阵P中第三列元素：

$$p(b_3/a_1)=0.2;\quad p(b_3/a_2)=0.5;\quad p(b_3/a_3)=0.4$$

其中最大者为$p(b_3/a^*)\geq p(b_3/a_2)=0.5$，则选择译码规则$F(b_3)=a^*=a_2$

5-2 译码规则和错误概率

为了能使信道输出符号$b_j(j=1,2,3)$与信道输入符号a_i $(i=1,2,3)$一一对应，本题按最大似然准则得到如下的译码规则：

$$\begin{cases} F(b_1)=a_1 \\ F(b_2)=a_3 \\ F(b_3)=a_2 \end{cases}$$

于是，最小平均错误译码概率为：

$$P_e=\frac{1}{r}\sum_{j=1}^{3}\sum_{i\neq *}p(b_j/a_i)=0.5667$$

5-2 译码规则和错误概率

(2). 当信道输入符号先验不等概，即 $p(a_1)=1/4; p(a_2)=1/2; p(a_3)=1/4$，要使 P_e 达到最小，则选择最大后验概率准则确定译码规则，即若

$$p(a^* / b_j) \geq p(a_i / b_j) \qquad (i, j = 1,2,3)$$

则选择译码规则：

$$F(b_j) = a^* \qquad (j = 1,2,3; a^* \in \{a_1, a_2, a_3\})$$

(i)对于信道输出符号 b_1：

$p(b_1 / a_1)p(a_1) = 1/8;\quad p(b_1 / a_2)p(a_2) = 1/10;$

$p(b_1 / a_3)p(a_3) = 3/40$

其中最大者为 $p(b_1 / a^*)p(a^*) = p(b_1 / a_1)p(a_1) = 1/8$，

则选择译码规则 $F(b_1) = a^* = a_1$

5-2 译码规则和错误概率

(ii)对于信道输出符号b_2:

$$p(b_2/a_1)p(a_1)=3/40; \quad p(b_2/a_2)p(a_2)=3/20;$$

$$p(b_2/a_3)p(a_3)=3/40$$

其中最大者为$p(b_2/a^*)p(a^*)=p(b_2/a_2)p(a_2)=3/20$,

则选择译码规则$F(b_2)=a^*=a_2$

(iii)对于信道输出符号b_3:

$$p(b_3/a_1)p(a_1)=1/20; \quad p(b_3/a_2)p(a_2)=1/4;$$

$$p(b_3/a_3)p(a_3)=1/10$$

其中最大者为$p(b_3/a^*)p(a^*)=p(b_3/a_2)p(a_2)=1/4$,

则选择译码规则$F(b_3)=a^*=a_2$

5-2 译码规则和错误概率

这样, 由按最大后验概率译码准则得到如下的译码规则:

$$\begin{cases} F(b_1) = a_1 \\ F(b_2) = a_2 \\ F(b_3) = a_2 \end{cases}$$

于是，最小平均错误译码概率为:

$$P_e = \sum_{j=1}^{3} \sum_{i \neq *} p(a_i) p(b_j / a_i) = 19/40$$

5-2 译码规则和错误概率

结论：

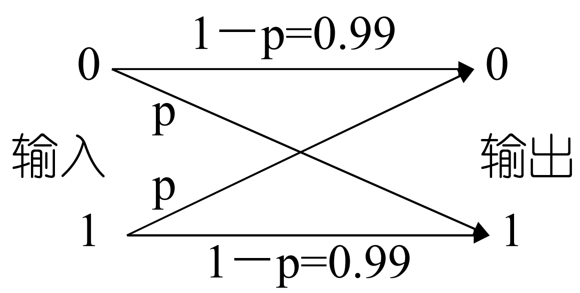
（1）对于给定信源和信道而言，只有采用最大后验概率译码准则，来选择译码规则，才能充分利用和挖掘译码规则，从而降低平均错误译码概率P_e使其达到最小值$P_{e\min}$

（2）分析P_e公式，发现它取决于信源和信道的统计特性，当信源给定，要使已经达到最小的P_e继续下降，就只有在坚持采用最大后验概率译码准则，选择译码规则的前提下，设法改变信道的统计特性

5-3 信道编码原则

[例题]已知二元对称信道如图：

如果输入符号等概$p(0)=p(1)=0.5$，试选择译码规则，并计算平均错误译码概率



解：由于输入等概，根据最大似然译码准则

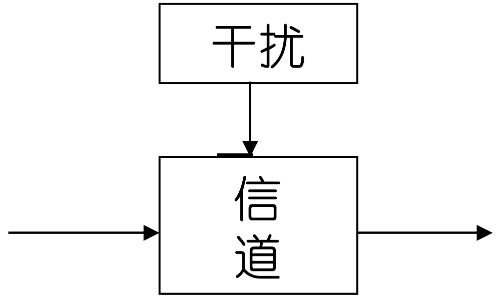
$$\text{信道矩阵}P=\begin{bmatrix}0.99 & 0.01\\ 0.01 & 0.99\end{bmatrix},$$

$$\text{得到译码规则:}\begin{cases}F(b_1)=a_1\\ F(b_2)=a_2\end{cases}\Rightarrow\begin{cases}F(0)=0\\ F(1)=1\end{cases}$$

$$\text{平均错误译码概率}:P_{e\min}=\frac{1}{r}\sum_{j=1}^{2}\sum_{i\neq *}p(b_j/a_i)=\frac{1}{2}(0.01+0.01)=0.01$$

5-3 信道编码原则

如果扩展信源符号，进行编码如下：

消息	未用码字	使用码字	信道	接收端输出序列		译码
α_1	—	000		000	β_1	
α_2	001	–		001	β_2	α_1
α_3	010	–		010	β_3	
α_4	011	–		100	β_5	
α_5	100	–		011	β_4	
α_6	101	–		101	β_6	α_8
α_7	110	–		110	β_7	
α_8	—	111		111	β_8	

$$
\text{信道矩阵} P = \begin{array}{c} \\ \alpha_1 \\ \alpha_8 \end{array}
\begin{array}{c}
\begin{array}{cccccccc} \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_6 & \beta_7 & \beta_8 \end{array} \\
\begin{bmatrix}
\bar{p}^3 & \bar{p}^2 p & \bar{p}^2 p & \bar{p} p^2 & \bar{p}^2 p & \bar{p} p^2 & \bar{p} p^2 & p^3 \\
p^3 & \bar{p} p^2 & \bar{p} p^2 & \bar{p}^2 p & \bar{p} p^2 & \bar{p}^2 p & \bar{p}^2 p & \bar{p}^3
\end{bmatrix}
\end{array}
$$

5-3 信道编码原则

对于离散无记忆信道，如果$p(0)=p(1)=0.5$,那么 $p(000)=p(111)=0.5$;

采用最大似然译码准则，得到译码规则：

$$F(\beta_1)=\alpha_1 \qquad F(\beta_5)=\alpha_1$$
$$F(\beta_2)=\alpha_1 \qquad F(\beta_6)=\alpha_8$$
$$F(\beta_3)=\alpha_1 \qquad F(\beta_7)=\alpha_8$$
$$F(\beta_4)=\alpha_8 \qquad F(\beta_8)=\alpha_8$$

5-3 信道编码原则

$$
\text{信道矩阵}P=\begin{array}{c c}
 & \begin{array}{cccccccc}\beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_6 & \beta_7 & \beta_8\end{array} \\
\begin{array}{c}\alpha_1 \\ \alpha_8\end{array} & \begin{bmatrix}
\bar{p}^3 & \bar{p}^2p & \bar{p}^2p & \boxed{\bar{p}p^2} & \bar{p}^2p & \boxed{\bar{p}p^2} & \boxed{\bar{p}p^2} & \boxed{p^3} \\
\boxed{p^3} & \boxed{\bar{p}p^2} & \boxed{\bar{p}p^2} & \bar{p}^2p & \boxed{\bar{p}p^2} & \bar{p}^2p & \bar{p}^2p & \bar{p}^3
\end{bmatrix}
\end{array}
$$

平均错误译码概率：

$$
P_e=\frac{1}{r}\sum_{j=1}^{s}\sum_{i\neq *}p(b_j/a_i)
$$

$$
=\frac{1}{2}[p^3+\bar{p}p^2+\bar{p}p^2+\bar{p}p^2+\bar{p}p^2+\bar{p}p^2+\bar{p}p^2+p^3]
$$

$$
=p^3+3\bar{p}p^2=2.98\times10^{-4}\approx3\times10^{-4}
$$

5-3 信道编码原则

问题1： 为什么通过重复编码，使得平均错误译码概率P_e降低？

根据编译码规则，使得两个输入消息分别与4个输出符号相对应：

$$\alpha_1 \to \{\beta_1\beta_2\beta_3\beta_5\};\alpha_8 \to \{\beta_4\beta_6\beta_7\beta_8\}$$

当发送消息中有一位错误时（一个码元），译码器可以正确译出所传的码字，有一位纠错能力

问题2： 虽然通过增加重复编码次数，可以降低平均错误译码概率，可是信息传输率是否变化呢?

5-3 信道编码原则

定义： 编码以后，消息的信息传输率：

$$R = \frac{\log M}{n} \text{(比特/码符号)}$$

M：信源的消息数即α个数；

n：编码以后码字的长度（码元个数）

如果传输每个码符号需要t秒钟，编码后

每秒传递的信息量$R_t = \frac{\log M}{nt} = \frac{R}{t}$(比特/秒)

$\log M$表示消息集在等概的条件下，每个消息携带的平均信息量

5-3 信道编码原则

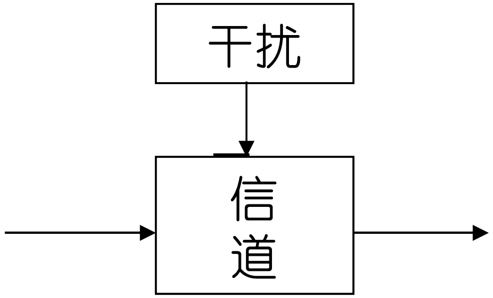
消息数M	码字长度n	平均错误译码概率P_e	信息传输率R (比特/码符号)	每秒信息传输量Rt (bit/sec) (每个码元传递时间t＝1)
2	n=1	0.01	log2/1=1	log2/1=1
2	n=3	3×10^{-4}	log2/3=1/3	log2/3=1/3
2	n=5	1×10^{-5}	log2/5=1/5	log2/5=1/5
2	n=7	4×10^{-7}	log2/7=1/7	log2/7=1/7
2	n=9	1×10^{-8}	log2/9=1/9	log2/9=1/9
2	n=11	5×10^{-10}	log2/11=1/11	log2/11=1/11

产生问题:

在以上随机编码中，当消息数M和码字长度n保持不变的条件下(也就是传输率不变)，应该遵循什么原则挑选码字，才可以得到尽量小的最小平均错误译码概率$P_{e\min}$?

5-3 信道编码原则

重新进行编码如下：

消息	未用码字	使用码字	信道	接收端输出序列		译码
α_1	—	000		000	β_1	α_1
α_2	—	001		001	β_2	α_2
α_3	010	—		010	β_3	α_1
α_4	011	—		011	β_4	α_2
α_5	100	—		100	β_5	α_1
α_6	101	—		101	β_6	α_2
α_7	110	—		110	β_7	α_1
α_8	111	—		111	β_8	α_2

$$
\text{信道矩阵}P = \begin{array}{c} \\ \alpha_1 \\ \alpha_2 \end{array}
\begin{array}{c}
\begin{array}{cccccccc} \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_6 & \beta_7 & \beta_8 \end{array} \\
\begin{bmatrix}
\bar{p}^3 & \bar{p}^2 p & \bar{p}^2 p & \bar{p} p^2 & \bar{p}^2 p & \bar{p} p^2 & \bar{p} p^2 & p^3 \\
\bar{p}^2 p & \bar{p}^3 & \bar{p} p^2 & \bar{p}^2 p & \bar{p} p^2 & \bar{p}^2 p & p^3 & \bar{p} p^2
\end{bmatrix}
\end{array}
$$

5-3 信道编码原则

对于离散无记忆信道，如果$p(0)=p(1)=0.5$,那么 $p(000)=p(001)=0.5$；

采用最大似然译码准则，得到译码规则：

$$F(\beta_1)=\alpha_1 \quad F(\beta_2)=\alpha_2$$
$$F(\beta_3)=\alpha_1 \quad F(\beta_4)=\alpha_2$$
$$F(\beta_5)=\alpha_1 \quad F(\beta_6)=\alpha_2$$
$$F(\beta_7)=\alpha_1 \quad F(\beta_8)=\alpha_2$$

5-3 信道编码原则

$$
信道矩阵P=\begin{array}{c c}
 & \begin{array}{cccccccc} \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_6 & \beta_7 & \beta_8 \end{array} \\
\begin{array}{c} \alpha_1 \\ \alpha_2 \end{array} & \begin{bmatrix} \bar{p}^3 & \boxed{\bar{p}^2 p} & \bar{p}^2 p & \boxed{\bar{p}p^2} & \bar{p}^2 p & \boxed{\bar{p}p^2} & \bar{p}p^2 & \boxed{p^3} \\ \boxed{\bar{p}^2 p} & \bar{p}^3 & \boxed{\bar{p}p^2} & \bar{p}^2 p & \boxed{\bar{p}p^2} & \bar{p}^2 p & \boxed{p^3} & \bar{p}p^2 \end{bmatrix}
\end{array}
$$

平均错误译码概率：

$$P_e = \frac{1}{r}\sum_{j=1}^{s}\sum_{i\neq *} p(b_j / a_i)$$

$$= \frac{1}{2}[\bar{p}^2 p + \bar{p}^2 p + \bar{p}p^2 + \bar{p}p^2 + \bar{p}p^2 + \bar{p}p^2 + p^3 + p^3]$$

$$= p^3 + \bar{p}^2 p + 2\bar{p}p^2 \approx 0.01$$

5-3 信道编码原则

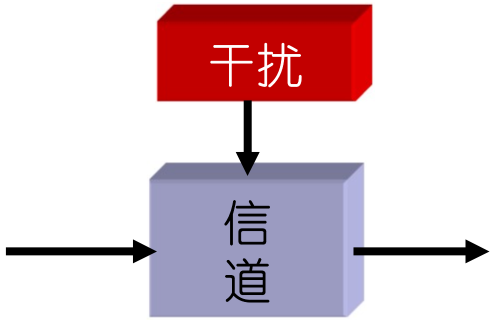
定义： 设α_i和β_j是两个由码符号集X中的码符号组成的长度为n的码符号序列，$\alpha_i=\{a_{i1},a_{i2},\cdots,a_{in}\}$，$\beta_j=\{b_{j1},b_{j2},\cdots,b_{jn}\}$，其中$a_{i1},a_{i2},\cdots,a_{in}\in X$，$b_{j1},b_{j2},\cdots,b_{jn}\in X$，在$\alpha_i$和$\beta_j$之间相同位置上的不同码符号个数，称为$\alpha_i$和$\beta_j$之间的汉明码距(*Hamming Distance*)记做$D(\alpha_i,\beta_j)$

对于二元信道，汉明距离通常可以描述成为：

$D(\alpha_i,\beta_j)=\sum_{k=1}^{n} a_{ik}\oplus b_{jk}$，其中$\oplus$为模二和运算

5-3 信道编码原则

对于一个信源和一个二元信道如果考虑输入端消息数为M个，用长度为n的码字进行信道随机编码

消息	编码	信道	接收端输出序列	
α_1	随机选择M个长度为n的码符号序列，分别代表M条消息 01010……	 对于信道传递概率而言$M\times 2^n$个传递概率为 $p(\beta_j/\alpha_i)$ $i=1,2,\ldots,M;\ \ j=1,2,\ldots,2^n$	2^n个长度为n的输出编码符号序列 01010……	β_1
α_2				β_2
α_3				β_3
α_4				β_4
∘ ∘ ∘				∘ ∘ ∘
α_M				β_{2^n}

5-3 信道编码原则

利用汉明距离来分析信道传递概率：

$$
\begin{aligned}
p(\beta_j/\alpha_i) &= p(b_{j1}b_{j2}\cdots b_{jn}/a_{i1}a_{i2}\cdots a_{in}) \\
&= p(b_{j1}/a_{i1})p(b_{j2}/a_{i2})\cdots p(b_{jn}/a_{in}) \\
&= \prod_{k=1}^{n} p(b_{jk}/a_{ik}) \\
&= \underbrace{ppp\cdots p}_{D(\alpha_i,\beta_j)\text{个}}\times \underbrace{\overline{ppp}\cdots\overline{p}}_{n-D(\alpha_i,\beta_j)} \\
&= p^{D(\alpha_i,\beta_j)}\,\overline{p}^{[n-D(\alpha_i,\beta_j)]}
\end{aligned}
$$

5-3 信道编码原则

将汉明距离与最大似然译码准则联系起来：

$$p(\beta_j/\alpha^*) \geq p(\beta_j/\alpha_i), \alpha_i \neq \alpha^*$$

$$p(\beta_j/\alpha^*) = p^{D(\alpha^*,\beta_j)}\overline{p}^{[n-D(\alpha^*,\beta_j)]} \geq p^{D(\alpha_i,\beta_j)}\overline{p}^{[n-D(\alpha_i,\beta_j)]} = p(\beta_j/\alpha_i)$$

其中 $i=1,2,\cdots,M, j=1,2,\cdots 2^n$

因此可以选择译码函数：$F(\beta_j)=\alpha^*$,其中$j=1,2,\cdots 2^n$

通常$p<1/2, p<<\overline{p}$,可见

$D(\alpha_i,\beta_j)$越大，$p^{D(\alpha_i,\beta_j)}$越小；$D(\alpha_i,\beta_j)$越小，$p^{D(\alpha_i,\beta_j)}$越大

可以将以上准则改为下列形式：

$D(\alpha^*,\beta_j) \leq D(\alpha_i,\beta_j)$，其中 $i=1,2,\cdots,M; j=1,2,\cdots 2^n$

选择译码函数：$F(\beta_j)=a^*$

其中$j=1,2,\cdots 2^n$； $a^* \neq a_i$

5-3 信道编码原则

利用汉明距离来描述最大似然译码准则：

在信道M个输入消息先验等概的条件下，离散无记忆信道的某一输出序列$\beta_j(j=1,2,\cdots 2^n)$，翻译成与$\beta_j$的汉明距离$D(\alpha_i,\beta_j)$中最小的一个$D_{\min}(\alpha_i,\beta_j)=D(\alpha^*,\beta_j)$所对应的输入消息$a^*$，则平均错误译码概率$P_e$达到最小$P_{e\min}$

5-3 信道编码原则

还可以得到，利用汉明距离表示的

最小平均错误译码概率$P_{e\min}$：

$$P_{e\min}=\frac{1}{M}\sum_{j=1}^{2^n}\sum_{i\neq *}p(\beta_j/\alpha_i)=\frac{1}{M}\sum_{j=1}^{2^n}\sum_{i\neq *}\{p^{D(\alpha_i,\beta_j)}\overline{p}^{[n-D(\alpha_i,\beta_j)]}\}$$

$$P_{e\min}=1-\frac{1}{M}\sum_{j=1}^{2^n}p(\beta_j/\alpha^*)=1-\frac{1}{M}\sum_{j=1}^{2^n}\{p^{D(\alpha^*,\beta_j)}\overline{p}^{[n-D(\alpha^*,\beta_j)]}\}$$

其中$i=1,2,\cdots,M,\ j=1,2,\cdots,2^n$

分析上面两个公式，发现信道编码的任务：

就是在保持R不变（M, n不变）的前提下，采用正确的选择M个码字的方法，使$P_{e\min}$最小

5-3 信道编码原则

进一步分析BSC信道，由于正确传递率$1-p$>>错误传递率p,得到结论：

(1)$D(\alpha_i,\beta_j)$越大，使得$P_{e\min}$越小，说明β_j与除了译码规则规定的相应码字α^*以外的$M-1$个码字α_i的汉明距离越大其平均错误译码概率就越小

(2)$D(\alpha^*,\beta_j)$越小，使得$P_{e\min}$越小，说明β_j与译码规则规定的相应码字α^*的汉明距离越小，其平均错误译码概率就越小

5-3 信道编码原则

定义: $D_{\min}(\alpha_i,\alpha_k), i \neq k$,表示$M$个码字中任何两个不同码字间的最小汉明距离，即

$$D_{\min}(\alpha_i,\alpha_k)=\min_{i\neq k}\{D(\alpha_i,\alpha_k)\}=d_{\min}$$

随机编码应该遵循的原则:

如果消息数M,编码长度n，保持不变，信息传输率$R=\dfrac{\log M}{n}$保持在相同的水平上，要使P_e最小达到$P_{e\min}$，实现方法就是在r^n个

5-3 信道编码原则

长度为n的代表消息的码符序列中，选择M个码符序列作为码字，分别代表M个消息。选择的标准是使这M个码字中的任何两个不同码字间的最小汉明距离$D_{\min}(\alpha_i,\alpha_k)$越大，也就是选出的$M$个码字间越不相似越好

问题：

(1)信道中的信息传输率R最大能到达什么样的水平？

(2)$P_{e\min}$又能降低到什么样的程度？

5-4 信道编码定理

噪声信道编码定理（香农第二极限定理）：

设某信道有r个输入符号，s个输出符号，信道的信道容量为C，当信道的信息传输率$R < C$时，只要码长n足够长，总可以在输入集合中(含有r^n个长度为n的码符号序列中)找到$M\ (M = 2^{nR})$个码字，分别代表M个等可能性的消息，组成一个信道编码，选择相应的译码规则，使信道输出端的译码过程的最小平均错误译码概率$P_{e\min}$达到任意小$(P_{e\min} \to 0)$

5-4 信道编码定理

噪声信道编码逆定理：

设某信道有r个输入符号，s个输出符号，信道的信道容量为C，当信道的信息传输率$R > C$时，无论码长n多长，都找不到一种编码，使信道输出端的译码过程的最小平均错误译码概率$P_{e\min}$达到任意小

噪声信道编码定理是一个存在性定理。它从理论上证明了平均错误译码概率趋向于零，同时信道的信息传输率R无限接近信道容量C的信道编码是存在的

5-5 差错控制系统和纠错码分类

5.5.1 差错控制

途径一（利用公式分析）

(1) 增大信道容量C (扩频，降噪)

(2) 减小信息传输率R (降低有效性换取可靠性提高)

(3) 增加码长n (用设备的复杂度换取可靠性提高)

途径二（通过概念分析）

(1) 利用冗余度

(2) 噪声均化

5-5 差错控制系统和纠错码分类

5.5.1 差错控制

信道中干扰和噪声类型：

（1）随机噪声　　（2）脉冲干扰和信道衰落

信道分类：

（1）随机信道：码元是否发生错误是随机的，相互独立，不会成串发生错误；

（2）突发信道：由于脉冲干扰和信道衰落导致产生的错误成片串现；

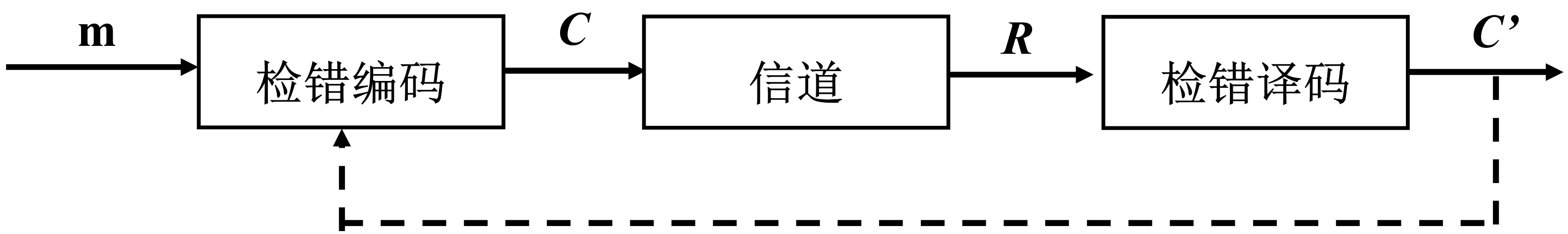
（3）混合信道：既有随机错误又有突发错误的信道为混合信道；

5-5 差错控制系统和纠错码分类

5. 5. 1 差错控制

通信系统中的纠、检错工作方式：

（1）反馈重传纠错：



（2）前向纠错：

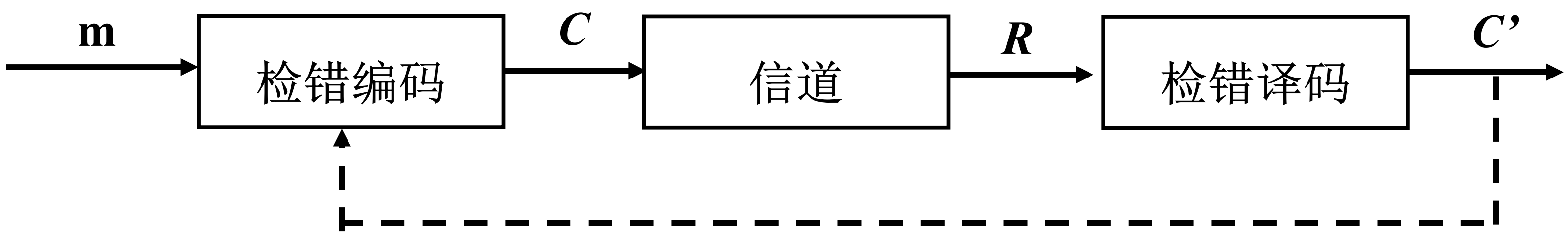
（3）混合纠错：

5-5 差错控制系统和纠错码分类

5.5.1 差错控制

通信系统中的纠、检错工作方式:

（1）反馈重传纠错ARQ (Automatic Repeat-reQuest):



发送端发出的是能够**发现错误**的**检错码**，接收端对接收到的码字进行译码，发现有错时，通过反馈系统向发送端请求重传已发送的全部或部分码字，直到接收端认为没有错误为止

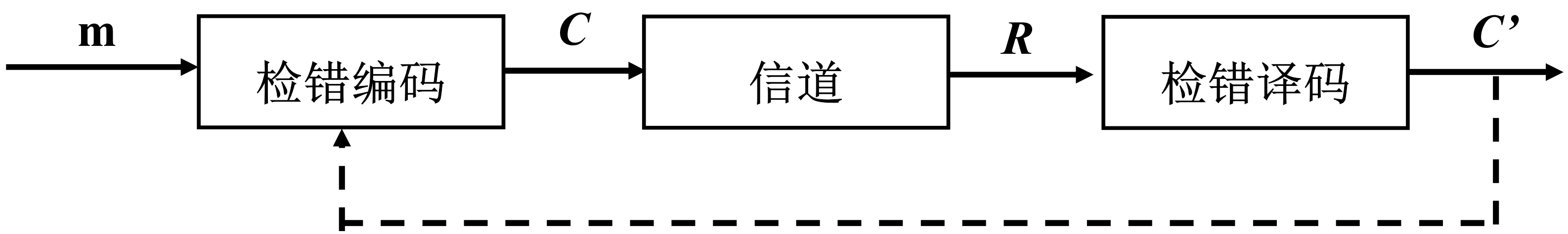
- 我国的电报系统就是一种反馈重传纠错系统

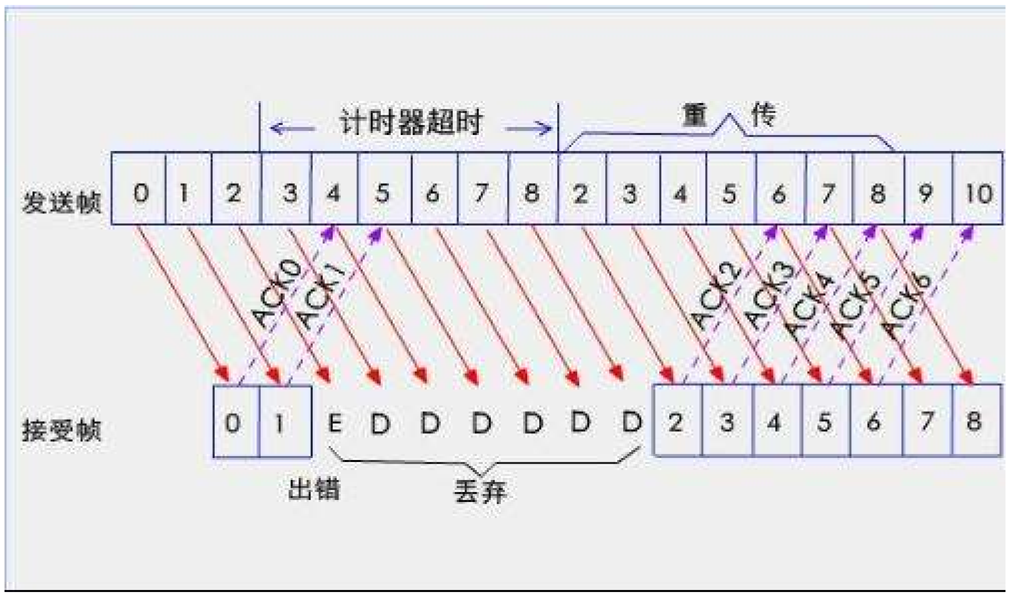
5-5 差错控制系统和纠错码分类

5.5.1 差错控制

通信系统中的纠、检错工作方式：

（1）反馈重传纠错ARQ (Automatic Repeat-reQuest)：





5-5 差错控制系统和纠错码分类

（2）前向纠错 FEC (Forward Error Correction)：

前向纠错也称为自动纠错。发送端发出的是具有**纠错能力**的**纠错码**，接收端根据编码规则进行解码。当误码个数在码的纠错能力范围之内时，译码器可以自动纠正错误。

（3）混合纠错：

对发送端进行适当的编码，当错误不严重时，在码的纠错能力之内，采用自动纠错，当产生的差错超出码的纠错能力时，则通过反馈系统向发送端要求重发，这种同时具有**反馈重传纠错**和**自动纠错**工作方式的纠错称为混合纠错。

5-5 差错控制系统和纠错码分类

5.5.2 纠错码分类

(1)按照对信息序列的处理方法分成**分组码(Block Code)**和**卷积码(Convolutional Code)**；其中，分组码又可以分为循环码和非循环码

分组码：将信息序列以每k个码元分组，然后将每组k个信息元按照一定规律产生r个多余的校验元，输出序列每组长度为$n=k+r$，则每个码字的r个校验元只与本码字的k个信息元有关，记为（n, k）分组码

5-5 差错控制系统和纠错码分类

(2)按照校验位与信息位的关系，可以分成线性码(Linear Code)和非线性码(Nonlinear Code)

线 性 码：信息位与校验位之间呈线性关系

非线性码：信息位与校验位之间不存在线性关系

(3)按照适用的差错类型，分为纠随机差错码(Random Error Correcting Code)和纠突发差错码(Burst Error Correcting Code)等

5-5 差错控制系统和纠错码分类

信道编码定理

- 1948年,信息论的奠基人C.E.Shannon在他的开创性论文“通信的数学理论”中,提出了著名的有噪信道编码定理。
- 他指出：对任何信道,只要信息传输速率R不大于信道容量C, 就一定存在这样的编码方法：在采用最大似然译码时,其误码率可以任意小。
- 该定理在理论上给出了对给定信道通过编码所能达到的编码增益的上限,并指出了为达到理论极限应采用的译码方法。
- 在信道编码定理中,香农提出了实现最佳编码的三个基本条件：(1)采用随机编译码方式; (2)编码长度 L →∞, 即分组的码组长度无限; (3)译码采用最佳的最大似然译码算法。

5-5 差错控制系统和纠错码分类

信道编码定理

➢在满足这三个条件的前提下,香农认为在有噪信道中可以实现无差错传输。在这一编码定理的理论引导下, 人们开始了对设计出信道好码的探索与研究。信道编码定理为人们探索信道的最佳编码方案提供了理论依据,但并没有指明如何获得好码。

➢后来出现了多种信道编码方案,如RS码、卷积码、级联码等。每一编码方案的提出,性能虽有所提高,但距香农极限还有很大距离。1993年出现的Turbo码,由于其很好地应用了编码定理中的随机性编译码条件和最佳译码算法, 而获得了几乎接近香农理论极限的译码性能, 立即在通信界引起了研究Turbo码的热潮。

5-5 差错控制系统和纠错码分类

信道编码历史

- 50-60年代初，主要研究各种有效的编码、译码方法，奠定了线性分组码的理论基础；提出了BCH编码、译码方法以及卷积码的序列译码；给出了纠错码的基本码限
- 60-70年代初，纠错码发展活跃。提出了如门限译码、迭代译码、软判决译码和卷积码的Viterbi译码等有效的编译码方法；同时注意到了纠错码实用化的问题，讨论了如码重量分布、译码错误概率和不可检错误概率的计算、信道的模型化等与实用化有关的各种问题
- 70年代以来，纠错码在实际应用中得到了更大发展。IC和微机的发展，为纠错码的实用打下了坚实的基础。70年代末、80年代初（20世纪），提出调码与调制相结合的网格编码调制（TCM, trellis-coded modulation）技术是编码理论的又一重要里程碑。1993年C. Berrou，A. Glavieux和P. Thitimajshima发现的Turbo码是又一重大突破

5-6 线性分组码

分组码基本思想：

- 信道编码的对象：是信源编码器输出的信息序列m。通常是二元符号1、0组成的序列。
- 信道编码的基本思想
  - ➢按一定规则给数字序列m增加一些多余的码元，使不具有规律性的信息序列m 变换为具有某种规律性的数码序列C；
  - ➢码序列中的**信息序列码元**与**多余码元**之间是相关的；
  - ➢信道译码器利用这种**预知的编码规则**译码。检验接收到的数字序列R 是否**符合既定的**规则，从而发现R 中是否有错，或者纠正其中的差错；
  - ➢根据相关性来**检测/发现**和**纠正**传输过程中产生的差错就是信道编码的基本思想。

5-6 线性分组码

分组码基本思想：

- 码元的组成及其它们之间的关系
  - **信息码组**：数字序列m 总是以k 个码元为一组传输，称这k 个码元的码组为信息码组。例如遥控系统中的每个指令字，计算机中的每个字节。
  - **码组/码字**：信道编码器按一定的规则对每个信息码组附加一些多余的码元，构成了n 个码元的码组。
  - 码组的n 个码元之间是相关的，附加的(n - k)个多余码元为何种符号序列与待编码的信息码组有关。
  - **监督码元/监督元**：附加的(n - k)个码元称为该码组的监督码元或监督元。

5-6 线性分组码

分组码基本思想：

- 可靠性与带宽、速度的关系
  - ➢从信息传输的角度，监督元不载有任何信息，所以是多余的。这种多余度使码字具有一定的纠错和检错能力，提高了传输的可靠性，降低了误码率；
  - ➢如果要求信息传输速度不变，在附加了监督元后必须减小码组中每个码元符号的持续时间，对二进制码，就是要减小脉冲宽；若编码前每个码脉冲的归一化宽度为1，则编码后的归一化宽度为k/n(k<n,k/n<1)，因此信道带宽必须展宽n/k倍；以带宽的多余度换取了信道传输的可靠性；
  - ➢如果保持码元持续时间不变，必须降低信息传输速率。以信息传输速度的多余度或称时间上的多余度换取了传输的可靠性。

5-6 线性分组码

5.6.1 分组码及其检错、纠错能力

定义

在2^k个长度为k的$\{0,1\}$序列中，设法按照一定的规则加入若干个$\{0,1\}$符号，将长度为k的$\{0,1\}$信息序列变成长度为$n(n>k)$的具有一定抗干扰能力的符号序列$(a_1a_2\cdots a_{k-1}a_ka_{k+1}\cdots a_{k+r})$，$a_1,a_2,\cdots,a_{k+r}\in\{0,1\}$，$k+r=n$. 因此由$2^k$种长度为$n=k+r$的$\{0,1\}$符号序列组成的集合构成了一个$(n,k)$分组码，代表长度为$k$的$2^k$个信息序列(消息)

（k为信息码元数，$r=n-k$为监督码元数，n为码长）

5-6 线性分组码

将2^k个长度为k的信息序列转变为2^k个长度为n的(n, k)分组码,即选择2^n种可能组合中的2^k种构成一个许用码集**C**，这实际上是一种函数对应变换问题。不同的$f_1, f_2, f_3, \ldots, f_n$体现了$r=(n-k)$个多余码元的不同安排方法，构成不同的$(n, k)$分组码.

$$\begin{cases} c_1 = f_1(a_1 a_2 \cdots a_k) \\ c_2 = f_2(a_1 a_2 \cdots a_k) \\ \qquad \vdots \\ c_n = f_n(a_1 a_2 \cdots a_k) \end{cases}$$

5-6 线性分组码

例题：长度为$k=3$的$2^k=2^3=8$种不同的$\{0,1\}$信息序列$(a_1a_2a_3)$按照如下变换：

$$\begin{cases} c_1 = a_1 \\ c_2 = a_2 \\ c_3 = a_3 \\ c_4 = a_1 + a_2 \\ c_5 = a_1 + a_3 \\ c_6 = a_2 + a_3 \end{cases}$$

就转变成为$n=k+r=6$的$\{0,1\}$序列$(c_1c_2\cdots c_6)$，$c_1,c_2,\cdots,c_6\in\{0,1\}$

即(6,3)分组码的8个码字

信息序列 $a_1a_2a_3$	码字 $\boldsymbol{C}$ $c_1c_2c_3\,c_4c_5c_6$
0 0 0	0 0 0 0 0 0
0 0 1	0 0 1 0 1 1
0 1 0	0 1 0 1 0 1
0 1 1	0 1 1 1 1 0
1 0 0	1 0 0 1 1 0
1 0 1	1 0 1 1 0 1
1 1 0	1 1 0 0 1 1
1 1 1	1 1 1 0 0 0

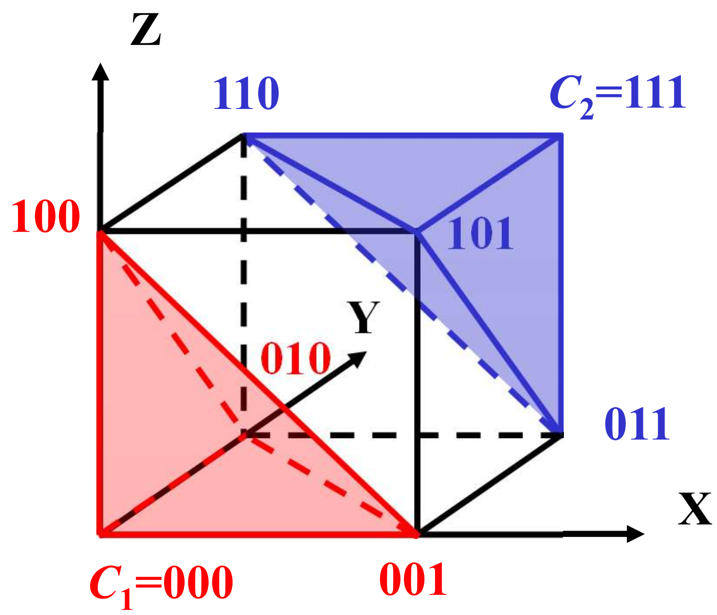
5-6 线性分组码

分组码(3, 1)的检错和纠错能力分析：

信源符号 0 ➞ 编码码字 000

信源符号 1 ➞ 编码码字 111

$$\begin{cases} c_1 = f_1(a_1) = a_1 \\ c_2 = f_2(a_1) = a_1 \qquad a_1 \in \{0,1\} \\ c_3 = f_3(a_1) = a_1 \end{cases}$$



利用三维坐标系中的一个点表示分组码的一个码字

C_1=000, X=Y=Z=0

C_2=111, X=Y=Z=1

实际上，(n, k)分组码的检错纠错能力，是直接由码字间的**汉明距离**来决定的

5-6 线性分组码

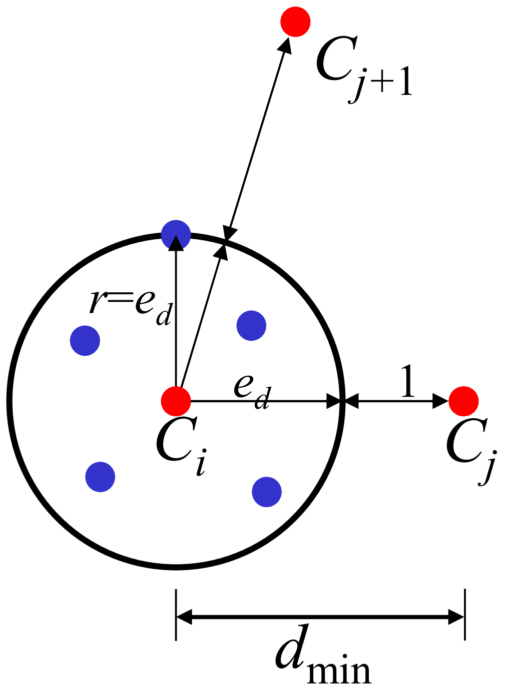
(n,k)分组码的检错纠错能力与码字间的汉明距离的联系：

(1)若(n,k)分组码能发现$\le e_d$个错误，那么(n,k)分组码中的任意两个码字C_i,C_j之间的最小汉明距离

$$D_{\min}(C_i,C_j)=d_{\min}\ge(e_d+1)$$

或者表示为

若(n,k)分组码的任意两个码字C_i,C_j之间的最小汉明距离$D_{\min}(C_i,C_j)=d_{\min}$，该$(n,k)$分组码能发现$e_d\le d_{\min}-1$个错误



5-6 线性分组码

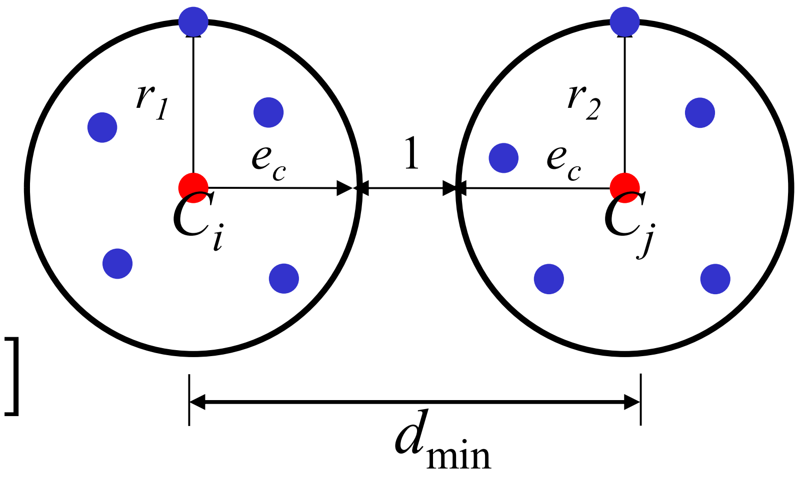
(n,k)分组码的检错纠错能力与码字间的汉明距离的联系:

(2)若(n,k)分组码能纠正$\le e_c$个错误，那么(n,k)分组码中的任意两个码字C_i,C_j之间的最小汉明距离

$$D_{\min}(C_i,C_j)=d_{\min}\ge(2e_c+1)$$

或者表示为

若(n,k)分组码的任意两个码字C_i,C_j之间的最小汉明距离$D_{\min}(C_i,C_j)=d_{\min}$，该$(n,k)$分组码能纠正$e_c\le INT\left[(d_{\min}-1)/2\right]$个错误



5-6 线性分组码

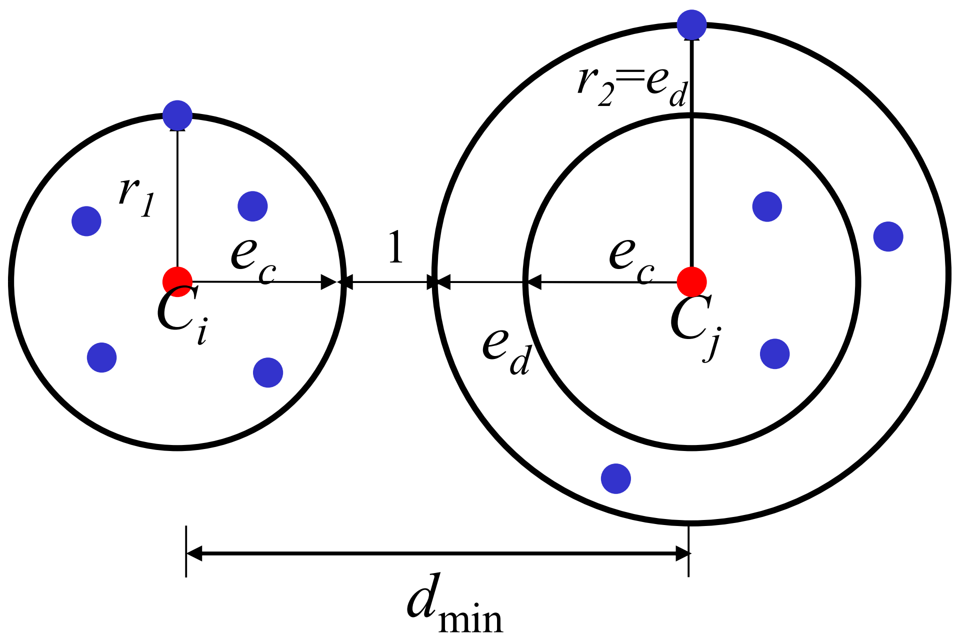
(n,k)分组码的检错纠错能力与码字间的汉明距离的联系：

(3)若(n,k)分组码能纠正$\le e_c$个错误，能发现e_d个错误$(e_c \le e_d)$，那么(n,k)分组码中的任意两个码字C_i, C_j之间的最小汉明距离$D_{\min}(C_i, C_j) = d_{\min} \ge (e_c + e_d + 1)$

或者表示为

若(n,k)分组码任意两个码字C_i, C_j之间的最小汉明距离为$D_{\min}(C_i, C_j) = d_{\min}$该分组码可以发现$e_d$，修正$e_c$个错误则必满足

$$\begin{cases} e_d + e_c \le d_{\min} - 1 \\ e_c \le e_d \end{cases}$$



5-6 线性分组码

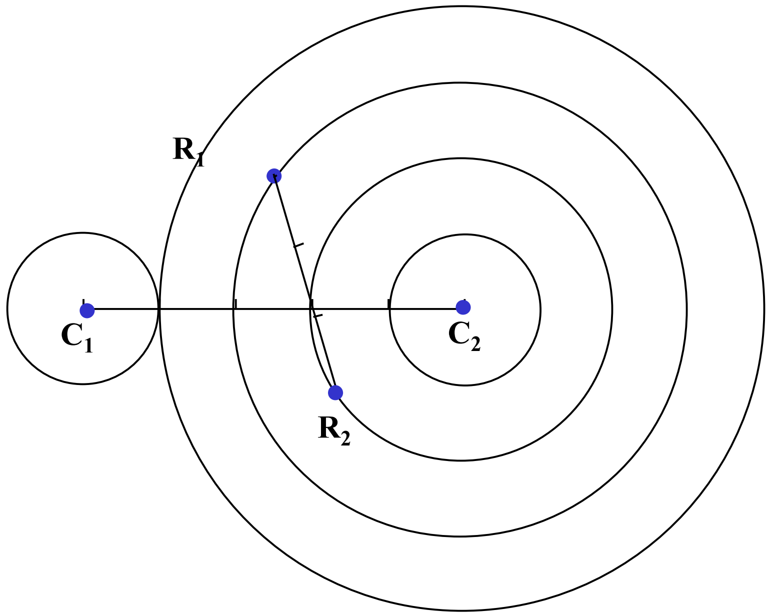
举例： 分析以下码字的检错和纠错能力

C_1=00000　C_2=11111　R_1=11000　R_2=10011

码字间的最小汉明距离：$d_{min}=D(C_1, C_2) = 5$

码字的检错能力：

$e_d \leqq d_{min} -1= 5-1 = 4$



码字的纠错能力：

$e_c \leqq (d_{min} -1) / 2 = 2$

码字的纠检错能力：

$e_c +e_d \leqq d_{min} -1= 5-1 = 4$

$e_c \leqq e_d$

则 $e_c = e_d = 2$ 或者 $e_c =1$ $e_d = 3$

5-6 线性分组码

举例： 分析以下码字的检错和纠错能力

$C_1=00000 \quad C_2=11111 \quad R_1=11000 \quad R_2=10011$

码字间的最小汉明距离：$d_{min}=D(C_1, C_2) = 5$

码字的检错能力：

$e_d \leqq d_{min}-1= 5-1 = 4$

码字的纠错能力：

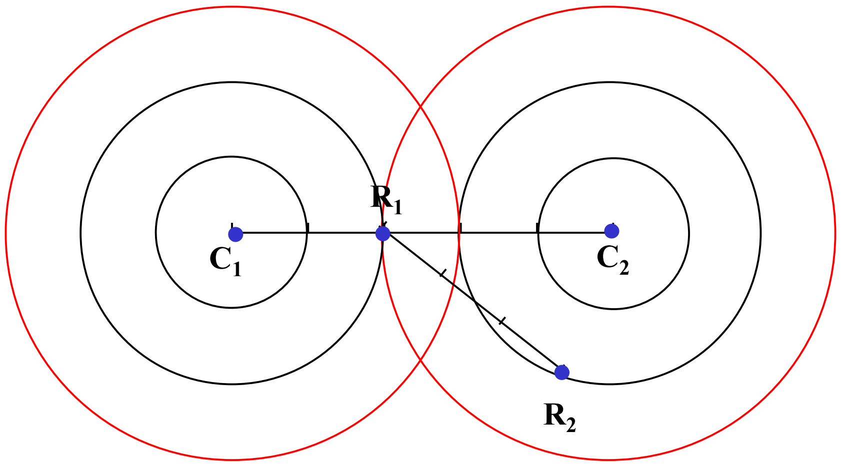
$e_c \leqq (d_{min}-1) / 2 = 2$

码字的纠检错能力：

$e_c+e_d \leqq d_{min}-1= 5-1 = 4$

$e_c \leqq e_d$

则 $e_c = e_d = 2$ 或者 $e_c =1$ $e_d = 3$



5-6 线性分组码

5.6.2 线性分组码

定义： 组成一个码字C_i的非零元素的个数，称为**码重**，表示为$W(C_i)$

定义： 如果码集中的所有码字都具有相同的码重，这种码就称为**恒重码**

5-6 线性分组码

在一个非空元素集合F，规定加法"+"和乘法"•"两种算数运算构成的域F中,算数运算必须满足以下规则:

1、集合F在加法、乘法运算下是封闭的;

即如果有$a,b\in F$,必有$a+b\in F,a\bullet b\in F$

2、满足结合律，如果$a,b,c\in F$,

则$a+(b+c)=(a+b)+c$及$a\bullet(b\bullet c)=(a\bullet b)\bullet c$

3、满足交换律，如果$a,b\in F$,

则$a+b=b+a$及$a\bullet b=b\bullet a$

4、集合F中一定包含一个零元素0，满足$a+0=a$

5、集合F中一定包含一个单位元I，满足$a\bullet I=a$

6、集合F中每个元素都有其对应的逆元素

5-6 线性分组码

定义：设C_i, C_j是某个(n,k)分组码的码字，a_1, a_2是码元符号集中的任意两个元素，那么当且仅当$a_1C_i + a_2C_j$也是码字时，称该码字为(n,k)**线性分**组码

结论：

(1) 凡是线性码必然包含**全零码**

如果$a_1, a_2=0$，$a_1C_i+a_2C_j=\underline{\mathbf{0}}$, 其中$\underline{\mathbf{0}}$表示全零码

(2) 线性码的**码字的组合**仍然**是码字**

如果$a_1, a_2=1$，$a_1C_i+a_2C_j=C_k$, 体现了封闭性

5-6 线性分组码

(3) 结合(1),(2),

$$D(C_i,C_j)=d=W(C_i+C_j)=W(C_k)= D(C_k, \underline{\mathbf{0}})$$

说明，(n, k)线性分组码码集中的两个码字的距离，必定等于码集中某一码字的码重。线性分组码的最小码距$d_{\min}$等于码集中非零码字的最小码重$W_{\min}$，可用各码字与全零码的距离或者各码字的码重代替两两码字间的码距进行研究。

5-6 线性分组码

矢量空间理论:

定义： 以n个线性无关的矢量为基底，它们的全部线性组合可以构成一个**n维n重矢量空间**S.如果从n个基底中选出一组$k(k<n)$个基底，则它们所有的线性组合也构成一个集合，这个集合是S的一个k维子集，称为**k维n重子空间**S_c

定义： 如果矢量空间中的两个矢量的点积为零，则称这两个矢量为**正交矢量**

定义： 如果一个矢量空间中的任一矢量都与另外一个矢量空间中的任一矢量正交，则称这两个**矢量空间正交**

5-6 线性分组码

矢量空间理论:

定义：以相互正交的基底张成的两个矢量空间一定正交，这两个空间称为**对偶空间**,其中一个空间是另外一个空间的零空间.

如果将n维n重矢量空间中相互正交的n个基底分成两组，一组k个基底，另一组$(n-k)$个基底，则它们分别张成k维n重和$(n-k)$维n重两个正交的对偶空间.

5-6 线性分组码

纠错编码的任务：

就是在n维n重矢量空间中的2^n种可能的组合中，选择2^k种构成一个k维n重矢量子空间(许用码集C).然后将2^k种信息组一一对应到许用码集C.

编码算法的核心问题：

(1) 寻找最佳码空间，即寻找最佳的一组(k个)基底，以张成一个码空间C.

(2) k维k重信息组空间的2^k个矢量以何种算法一一对应到k维n重码空间C.

5-6 线性分组码

5.6.3 生成矩阵和校验矩阵

(1) 生成矩阵

例题: (6,3)二进制线性分组码的输入信息组是$\{a_1a_2a_3\}$，编码输出是$C=(c_1c_2c_3c_4c_5c_6)$，已知输入输出码元的映射关系是：$c_1=a_1, c_2=a_2, c_3=a_3, c_4=a_1\oplus a_2, c_5=a_1\oplus a_3, c_6=a_2\oplus a_3$.求解：码集C以及编码时的映射算法.

5-6 线性分组码

长度为$k=3$的$2^k=2^3=8$种不同的$\{0,1\}$信息序列$(a_1a_2a_3)$按照如下变换：

$$\begin{cases} c_1=a_1 \\ c_2=a_2 \\ c_3=a_3 \\ c_4=a_1+a_2 \\ c_5=a_1+a_3 \\ c_6=a_2+a_3 \end{cases} \Longrightarrow [c_1c_2\cdots c_6]=[a_1a_2a_3]\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

$$\mathbf{C}=\mathbf{m}\cdot\mathbf{G}$$

就转变成为$n=k+r=6$的$\{0,1\}$序列$(c_1c_2\cdots c_6),c_1,c_2,\cdots,c_6\in\{0,1\}$即(6,3)分组码的8个码字

信息序列 $a_1a_2a_3$	码字C $c_1c_2c_3$ $c_4c_5c_6$
000	**000 000**
001	**001 011**
010	**010 101**
100	**100 110**
011	**011 110**
101	**101 101**
110	**110 011**
111	**111 000**

5-6 线性分组码

定义：设由$a_1, a_2, \cdots, a_k \in \{0,1\}$组成的长度为$k$的信息序列，$\mathbf{m} = [a_1 a_2 \cdots a_k]$产生$2^k$种不同的信息序列以代表信源发出的$2^k$种不同的消息，再令矩阵

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ g_{k1} & g_{k2} & \cdots & g_{kn} \end{bmatrix}$$

为该码的**生成矩阵**，其中每个元素$g_{ij}(i = 1, 2, \cdots k; j = 1, 2 \cdots n)$就是$k$个线性独立的码字$C$的$n$个分量(码符号)，并且$\mathbf{C} = \mathbf{m} \cdot \mathbf{G}$

5-6 线性分组码

注意：

(1) 任何码字均是生成矩阵**G**的k个行矢量的线性组合，只要k个行矢量线性无关，就可以作为k个基底张成一个k维n重空间，它是n维n重空间的子空间，子空间中的所有2^k个矢量构成码集**C**.

(2) 不同的生成矩阵**G**产生不同的码，生成矩阵**G**的特点决定了码的特点.

(3) 构成同一个空间的基底不是唯一的，不同的基底(或生成矩阵)有可能生成同一个码集**C** . 但是不能说是同样的编码(由于编码涉及码集和映射两方面，码集相同而映射方法不同，就不是同样的编码).

(4) 生成矩阵的k个行矢量既是k个基底，也是k个码字.

5-6 线性分组码

(2) 系统码

若(n,k)分组码的生成矩阵$\mathbf{G}$是一个$(k\times k)$的单位方阵$[I_k]$和一个k行，$r=n-k$列的$(k\times r)$阶矩阵$[P_{k,r}]$组合而成，即

$$\mathbf{G}=\left[\begin{array}{cccc:cccc} 1 & 0 & \cdots & 0 & p_{11} & p_{12} & \cdots & p_{1(n-k)} \\ 0 & 1 & \cdots & 0 & p_{21} & p_{22} & \cdots & p_{2(n-k)} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & p_{k1} & p_{2k} & \cdots & p_{k(n-k)} \end{array}\right]=\left[I_k \mid P_{k,r}\right]$$

那么,(n,k)分组码的前k位与相应信息序列完全相同。

$$\mathbf{C}=\mathbf{m}\cdot\mathbf{G}=[a_1a_2\cdots a_k]\left[\begin{array}{cccc:cccc} 1 & 0 & \cdots & 0 & p_{11} & p_{12} & \cdots & p_{1(n-k)} \\ 0 & 1 & \cdots & 0 & p_{21} & p_{22} & \cdots & p_{2(n-k)} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & p_{k1} & p_{k2} & \cdots & p_{k(n-k)} \end{array}\right]$$

$$=\left[a_1,a_2,\cdots,a_k,\left(\sum_{i=1}^{k}a_ip_{i1}\right),\left(\sum_{i=1}^{k}a_ip_{i2}\right),\cdots,\left(\sum_{i=1}^{k}a_ip_{i(n-k)}\right)\right]$$其中$\sum$为$\oplus$运算。

5-6 线性分组码

定义：所有码字的前面k位码元符号与相应的信息序列的码元符号完全相同的(n,k)线性分组码，又称为**系统码**

问题1：一个非系统线性分组码是否可以转变成为一个系统线性分组码？

例如(6,3)线性分组码的生成矩阵为**G**,相应码字为：

$$\mathbf{G}=\begin{bmatrix}0&1&1&0&0&1\\1&1&0&1&0&0\\1&0&1&0&1&0\end{bmatrix}$$

信息序列	码字	信息序列	码字
000	000 000	011	011 110
001	101 010	101	110 011
010	110 100	110	101 101
100	011 001	111	000 111

5-6 线性分组码

$$\mathbf{G}=\begin{bmatrix}0&1&1&0&0&1\\1&1&0&1&0&0\\1&0&1&0&1&0\end{bmatrix}\xrightarrow[\text{列置换}]{\substack{c6\to c1\\ c4\to c2\\ c5\to c3}}$$

$$\begin{bmatrix}1&0&0&1&1&0\\0&1&0&1&0&1\\0&0&1&0&1&1\end{bmatrix}\xrightarrow{\text{等价}}\mathbf{G}_0$$

信息序列 $m_1m_2m_3$	码字 $c_1c_2c_3$ $c_4c_5c_6$
000	000 000
001	001 011
010	010 101
100	100 110
011	011 110
101	101 101
110	110 011
111	111 000

结论： 任何一个非系统(n,k)线性分组码都和一个系统(n,k)线性分组码等价.

5-6 线性分组码

问题2：非系统线性分组码与系统线性分组码的检错和纠错能力是否相同？

结论：非系统(n,k) 线性分组码可以**等价(纠错检错能力相同)**为系统(n,k)线性分组码。由于系统码的编码过程和设备比非系统(n,k)线性分组码简单，而且可以保持与非系统(n,k)线性分组码相同的检错纠错能力，实际较多采用系统(n,k)线性分组码。

5-6 线性分组码

例题：3维重复编码是一个（3,1）线性分组码，请列写其生成矩阵

$m_1=1 \quad C_1=111 \qquad m_2=0 \quad C_2=000$

$\mathbf{G}=(1\ 1\ 1\ )$

$C_1=m_1.\mathbf{G}=1(1\ 1\ 1\ )=111$

$C_1=m_1.\mathbf{G}=0(1\ 1\ 1\ )=000$

5-6 线性分组码

例题：(4,3)偶校验码是系统线性分组码，请列写其生成矩阵

$$\mathbf{G}=\begin{bmatrix}1 & 0 & 0 & 1\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 1\end{bmatrix}$$

$$\mathbf{C}=C_1C_2C_3C_4=(m_1m_2m_3)\begin{bmatrix}1 & 0 & 0 & 1\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 1\end{bmatrix}$$

$$=\left(m_1 \quad m_2 \quad m_3 \quad m_1+m_2+m_3\right)$$

信息序列 $m_1m_2m_3$	码字 $c_1c_2c_3c_4$
0 0 0	0 0 0 0
0 0 1	0 0 1 1
0 1 0	0 1 0 1
1 0 0	1 0 0 1
0 1 1	0 1 1 0
1 0 1	1 0 1 0
1 1 0	1 1 0 0
1 1 1	1 1 1 1

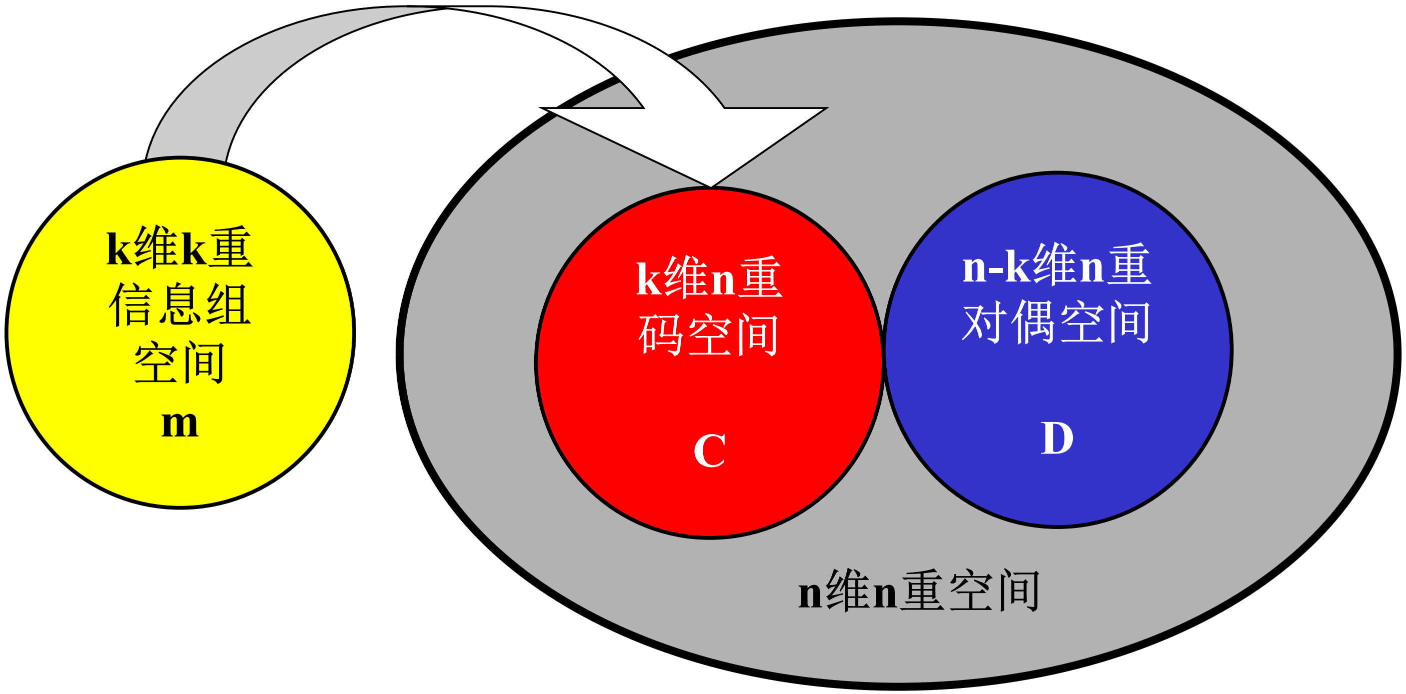
5-6 线性分组码

(3) 校验矩阵（译码）

任何一个(n,k)线性码的码空间C一定存在一个对偶空间D用$(n-k)$个基底产生一个有$2^{(n-k)}$个码矢的$(n,n-k)$线性码，称为(n,k)线性码的**对偶码**；

5-6 线性分组码

类似于码空间C的k个基底排列起来，可以将对偶空间D中的$(n-k)$个基底排列起来，构成一个$(n-k)\times n$阶矩阵，称为码空间C的校验矩阵**H**.



5-6 线性分组码

对偶空间的基底正交：码空间**C**和对偶空间**D**正交

$C_i \mathbf{H}^T = \mathbf{0}, \quad C_i \in \mathbf{C}, \quad \mathbf{0}$为$1 \times (n-k)$全零行向量

$\mathbf{G}\mathbf{H}^T = \mathbf{0}, \quad \mathbf{0}$为$k \times (n-k)$全零矩阵

考虑到(n, k)系统线性分组码的生成矩阵$\mathbf{G}_0$

$\mathbf{G}_0 \mathbf{H}^T = \mathbf{0}, \quad C_i \in \mathbf{C}, \quad \mathbf{0}$为$k \times (n-k)$全零行向量

5-6 线性分组码

$\mathbf{H}^{\mathrm{T}}$必定为如下形式：

$$\mathbf{H}^T=\begin{bmatrix} q_{11} & q_{21} & \cdots & q_{(n-k)1} \\ q_{12} & q_{22} & \cdots & q_{(n-k)2} \\ \vdots & \vdots & \cdots & \vdots \\ q_{1k} & q_{2k} & \cdots & q_{(n-k)k} \\ \hdashline 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},\text{其中}\begin{matrix} q_{11}=p_{11} & q_{21}=p_{12} & \cdots & q_{(n-k)1}=p_{1(n-k)} \\ q_{12}=p_{21} & q_{22}=p_{22} & \cdots & q_{(n-k)2}=p_{2(n-k)} \\ \vdots & \vdots & \cdots & \vdots \\ q_{1k}=p_{k1} & q_{2k}=p_{k2} & \cdots & q_{(n-k)k}=p_{k(n-k)} \end{matrix}$$

5-6 线性分组码

$$
\mathbf{G}_0\mathbf{H}^T = k\left\{\left[\begin{array}{cccc|cccc}
\multicolumn{4}{c}{\overbrace{\hphantom{1\quad 0\quad \cdots\quad 0}}^{k}} & \multicolumn{4}{c}{\overbrace{\hphantom{p_{11}\quad p_{12}\quad \cdots\quad p_{1(n-k)}}}^{n-k}} \\
1 & 0 & \cdots & 0 & p_{11} & p_{12} & \cdots & p_{1(n-k)} \\
0 & 1 & \cdots & 0 & p_{21} & p_{22} & \cdots & p_{2(n-k)} \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & 1 & p_{k1} & p_{2k} & \cdots & p_{k(n-k)}
\end{array}\right]\right.
\begin{bmatrix}
q_{11} & q_{21} & \cdots & q_{(n-k)1} \\
q_{12} & q_{22} & \cdots & q_{(n-k)2} \\
\vdots & \vdots & \cdots & \vdots \\
q_{1k} & q_{2k} & \cdots & q_{(n-k)k} \\
\hdashline
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & 1
\end{bmatrix}
$$

$$
= \underbrace{\left.\begin{bmatrix}
0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & 0
\end{bmatrix}\right\}k}_{n-k}
$$

5-6 线性分组码

$$
\mathbf{G}_0\mathbf{H}^T = k\left\{\left[\overbrace{\begin{matrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{matrix}}^{k}\;\overbrace{\boxed{\begin{matrix} p_{11} & p_{12} & \cdots & p_{1(n-k)} \\ p_{21} & p_{22} & \cdots & p_{2(n-k)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{k1} & p_{2k} & \cdots & p_{k(n-k)} \end{matrix}}}^{n-k}\right]\right.
\begin{bmatrix} \boxed{\begin{matrix} p_{11} & p_{12} & \cdots & p_{1(n-k)} \\ p_{21} & p_{22} & \cdots & p_{2(n-k)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{k1} & p_{2k} & \cdots & p_{k(n-k)} \end{matrix}} \\ \hdashline \begin{matrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{matrix} \end{bmatrix}
$$

$$
= \underbrace{\left.\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}\right\}k}_{n-k}
$$

5-6 线性分组码

$$\mathbf{G}_0\mathbf{H}^T = \begin{array}{c} \\ k\left\{\begin{matrix} \\ \\ \\ \\ \end{matrix}\right. \end{array} \begin{array}{c} \overbrace{}^{k}\;\overbrace{\phantom{p_{11}\quad p_{12}\quad \cdots\quad p_{1(n-k)}}}^{n-k} \\ \begin{bmatrix} 1 & 0 & \cdots & 0 & p_{11} & p_{12} & \cdots & p_{1(n-k)} \\ 0 & 1 & \cdots & 0 & p_{21} & p_{22} & \cdots & p_{2(n-k)} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & p_{k1} & p_{2k} & \cdots & p_{k(n-k)} \end{bmatrix} \end{array} \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1(n-k)} \\ p_{21} & p_{22} & \cdots & p_{2(n-k)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{k1} & p_{2k} & \cdots & p_{k(n-k)} \\ \hline 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} \boxed{0} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}}_{n-k} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} k$$

5-6 线性分组码

$$\mathbf{G}_0\mathbf{H}^T = k\left\{\begin{bmatrix} \overbrace{\begin{matrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{matrix}}^{k} & \overbrace{\begin{matrix} p_{11} & p_{12} & \cdots & p_{1(n-k)} \\ p_{21} & p_{22} & \cdots & p_{2(n-k)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{k1} & p_{2k} & \cdots & p_{k(n-k)} \end{matrix}}^{n-k} \end{bmatrix}\right. \begin{bmatrix} q_{11} & q_{21} & \cdots & q_{(n-k)1} \\ q_{12} & q_{22} & \cdots & q_{(n-k)2} \\ \vdots & \vdots & \cdots & \vdots \\ q_{1k} & q_{2k} & \cdots & q_{(n-k)k} \\ \hdashline 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}}_{n-k}\left.\vphantom{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}\right\}k$$

5-6 线性分组码

例题： 考虑一个(7,4)码，其生成矩阵为

$$\mathbf{G}=\begin{bmatrix}1&0&0&0&1&0&1\\0&1&0&0&1&1&1\\0&0&1&0&1&1&0\\0&0&0&1&0&1&1\end{bmatrix}$$

(1) 对于信息组m=(1 0 1 1),编出的码字是什么?

(2) 设计一个(7,4)分组码编码器原理图

(3) 若接收到一个7位码r=(1 0 0 1 1 0 0)，它是否为码字?

5-6 线性分组码

解:(1)、由 $C=mG$ 可得: $C=(1011000)$

(1)、码集 $C=(c_1c_2c_3c_4c_5c_6c_7)$, 信息序列 $m=(m_1m_2m_3m_4)$,

由 $C=mG$ 得:

$$(c_1c_2c_3c_4c_5c_6c_7)=(m_1m_2m_3m_4c_5c_6c_7)$$

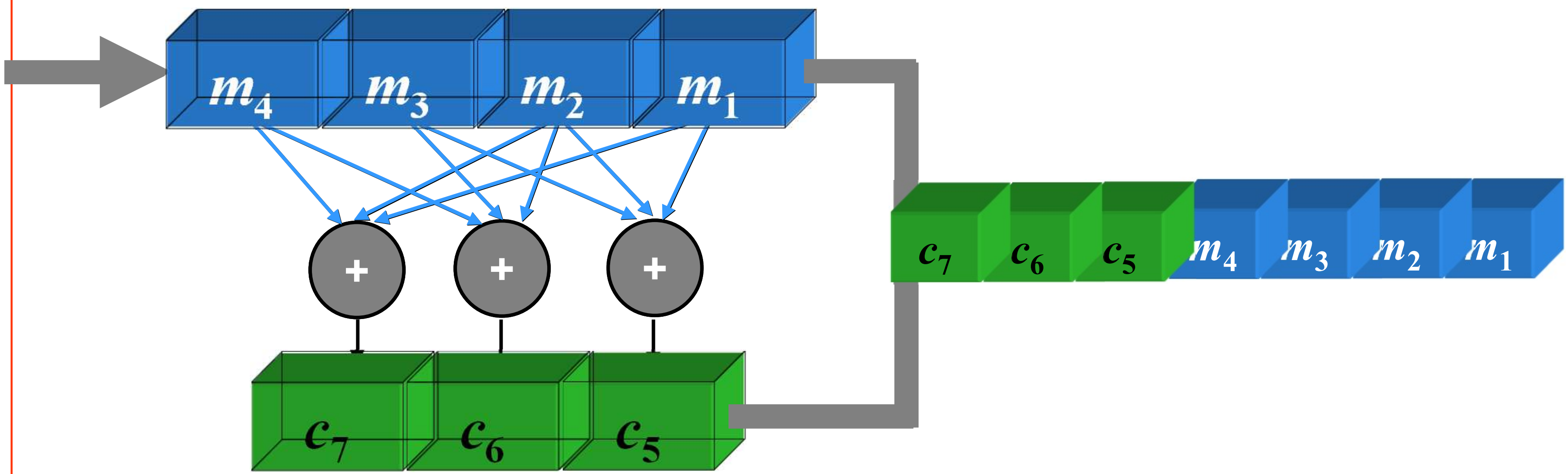
$$=m_1(1000101)+m_2(0100111)+m_3(0010110)+m_4(0001011)$$

于是，可得到线性方程组:

$$\begin{cases} c_1=m_1 \\ c_2=m_2 \\ c_3=m_3 \\ c_4=m_4 \\ c_5=m_1+m_2+m_3 \\ c_6=m_2+m_3+m_4 \\ c_7=m_1+m_2+m_4 \end{cases}$$

5-6 线性分组码

编码器的原理图如下：



5-6 线性分组码

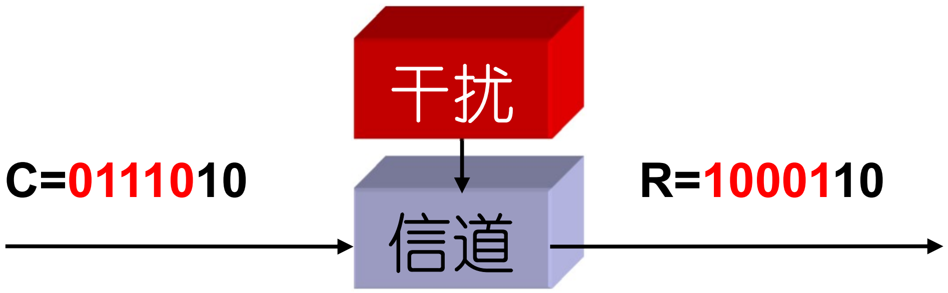
(3)、由$GH^T=0$可得校验矩阵：

$$H=\begin{bmatrix}1&1&1&0&1&0&0\\0&1&1&1&0&1&0\\1&1&0&1&0&0&1\end{bmatrix}$$

由$rH^T=\begin{bmatrix}0&1&0\end{bmatrix}\neq 0$,所以$r$不是码字.

5-6 线性分组码

5.6.4 伴随式与译码



C+R=E

C=0111010

R=1000110

E=1111100

定义：

差错图案E为

$$\mathbf{E}=(e_1,e_2,\cdots,e_n)=(c_1+r_1,c_2+r_2,\cdots,c_n+r_n)=\mathbf{C}+\mathbf{R}$$

作为信道中干扰图案的反映　（R为接收矢量）

$$\mathbf{C},\mathbf{E},\mathbf{R}\text{之间的关系}\begin{cases}\mathbf{C}=\mathbf{E}+\mathbf{R}\\ \mathbf{E}=\mathbf{C}+\mathbf{R}\\ \mathbf{R}=\mathbf{C}+\mathbf{E}\end{cases}$$

5-6 线性分组码

引入差错图案**E**以后，一致校验矩阵**H**在译码过程中的作用：

$$\mathbf{RH}^T=(\mathbf{C}+\mathbf{E})\mathbf{H}^T=\mathbf{CH}^T+\mathbf{EH}^T$$

由于 $\mathbf{CH}^T=\mathbf{0}$ 因此 $\mathbf{RH}^T=\mathbf{EH}^T$

(1)如果传输过程中没有发生错误

$\mathbf{C}=\mathbf{R}$, $\mathbf{E}=\mathbf{0}$, $\mathbf{EH}^T=\mathbf{0}$, $\mathbf{RH}^T=\mathbf{0}$

(2)如果传输过程中发生了错误， $\mathbf{E}\neq\mathbf{0}$

$\mathbf{EH}^T\neq\mathbf{0}$, $\mathbf{RH}^T\neq\mathbf{0}$

在一致校验矩阵不变的条件下， $\mathbf{RH}^T$只与**E**有关，与**C**无关

5-6 线性分组码

定义：伴随式$\mathbf{S}=(s_1,s_2,\cdots,s_{n-k})=\mathbf{RH}^T=\mathbf{EH}^T$ 作为伴随接收码**R**的一个$n-k$重矢量

要判断信道输出端的接收序列矢量**R**是否是(n,k)线性分组码的码字问题就转化成了判断接收序列矢量**R**的伴随式**S**是否是全零矢量**0**的问题，即：

$\mathbf{S}=\mathbf{RH}^T=\mathbf{EH}^T=\mathbf{0}$，那么**R**一定是$(n,k)$线性分组码码字

$\mathbf{S}=\mathbf{RH}^T=\mathbf{EH}^T\neq\mathbf{0}$，那么**R**一定不是$(n,k)$线性分组码码字

实际译码过程：

$$\mathbf{RH}^T=\mathbf{S} \quad \Rightarrow \quad \mathbf{E} \quad \Rightarrow \quad \mathbf{C}=\mathbf{E}+\mathbf{R}$$

关键在怎样由**S=>E**，由于 $\mathbf{EH}^T=\mathbf{S}$, 可将**S** **E** 联系起来

5-6 线性分组码

$$\mathbf{S}=(s_1,s_2,\cdots,s_{n-k})=\mathbf{EH}^T=(e_1,e_2,\cdots,e_n)\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ h_{(n-k)1} & h_{(n-k)2} & \cdots & h_{(n-k)n} \end{bmatrix}^T$$

$$\Rightarrow \begin{cases} s_1 = e_1h_{11} + e_2h_{12} + \cdots + e_nh_{1n} \\ s_2 = e_1h_{21} + e_2h_{22} + \cdots + e_nh_{2n} \\ \qquad\vdots \\ s_{n-k} = e_1h_{(n-k)1} + e_2h_{(n-k)2} + \cdots + e_nh_{(n-k)n} \end{cases}$$

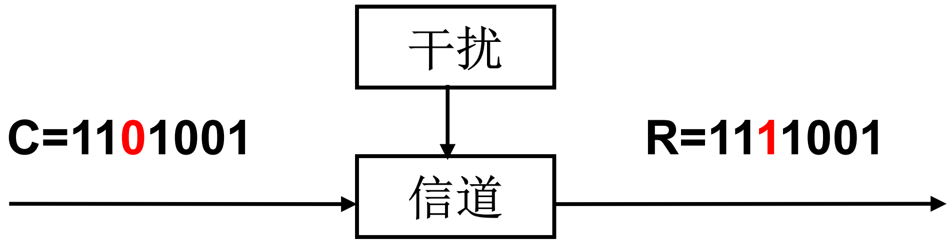
方程中有n个未知数，n-k个方程，因此每个未知数应该有2^k个解

5-6 线性分组码

在2^k个解中选择附加在接收序列**R**上的错误图案**E**的方法

概率译码方法：选择2^k个解中，汉明重量最小的那个作为附加在接收序列**R**上的错误图案**E**。

例题：考虑如图BSC信道p为错误传递概率，且$(1-p>>p)$，对于一个(7,4)系统线性分组码码字C



错误图样：

$$\mathbf{E}=\mathbf{C}+\mathbf{R}=(\mathbf{1101001})+(\mathbf{1111001})=\mathbf{0010000}$$

$$P(\mathbf{E})=P(\mathbf{R/C})=P(\mathbf{1111001/\ 1101001})$$

$$=P(1/1)\ P(1/1)\ P(1/0)\ P(1/1)\ P(0/0)\ P(0/0)\ P(1/1)$$

$$=p^1(1-p)^6=p^{W(\mathbf{E})}(1-p)^{[n-W(\mathbf{E})]}$$

其中 $W(\mathbf{E})=1,\ 6=n-1=7-1$ n是码字长度

5-6 线性分组码

标准阵列译码表

$S_1 \Rightarrow E_1$	$E_1 + C_1 = 0$	$E_1 + C_2$	$\cdots$	$E_1 + C_i$	$\cdots$	$E_1 + C_{2^k}$
$S_2 \Rightarrow E_2$	$E_2 + C_1 = E_2$	$E_2 + C_2$	$\cdots$	$E_2 + C_i$	$\cdots$	$E_2 + C_{2^k}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
$S_j \Rightarrow E_j$	$E_j + C_1 = E_j$	$E_j + C_2$	$\cdots$	$E_j + C_i$	$\cdots$	$E_j + C_{2^k}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
$S_{2^{n-k}} \Rightarrow E_{2^{n-k}}$	$E_{2^{n-k}} + C_1 = E_{2^{n-k}}$	$E_{2^{n-k}} + C_2$	$\cdots$	$E_{2^{n-k}} + C_i$	$\cdots$	$E_{2^{n-k}} + C_{2^k}$

子集中含有2^{n-k}个元素

2^k个互不相交的子集

5-6 线性分组码

例题：某一个(5,2)系统线性码的生成矩阵

$$\mathbf{G}=\begin{bmatrix}1&0&1&1&1\\0&1&1&0&1\end{bmatrix},\quad \text{设接收码}\mathbf{R}=(10101)$$

先构造该码的标准阵列译码表，再进行译码。

解：首先根据

$$\mathbf{C}=\mathbf{mG}=(m_1,m_2)\begin{bmatrix}1&0&1&1&1\\0&1&1&0&1\end{bmatrix}$$

得到信息序列与相应的码字**C**,再由**G**生成一致校验矩阵**H**

$$\mathbf{H}=\begin{bmatrix}1&1&1&0&0\\1&0&0&1&0\\1&1&0&0&1\end{bmatrix}$$

信息序列	码字
00	**00000**
01	**01101**
10	**10111**
11	**11010**

5-6 线性分组码

S_1=000				
S_2=111				
S_3=101				
S_4=100				
S_5=010				
S_6=001				
S_7=011				
S_8=110				

5-6 线性分组码

S_1=000	E_1+C_1=00000	C_2=10111	C_3=01101	C_4=11010
S_2=111				
S_3=101				
S_4=100				
S_5=010				
S_6=001				
S_7=011				
S_8=110				

5-6 线性分组码

$S_1=000$	$E_1+C_1=00000$	$C_2=10111$	$C_3=01101$	$C_4=11010$
$S_2=111$	$E_2=10000$			
$S_3=101$	$E_3=01000$			
$S_4=100$	$E_4=00100$			
$S_5=010$	$E_5=00010$			
$S_6=001$	$E_6=00001$			
$S_7=011$	$E_7=00011$			
$S_8=110$	$E_8=00110$			

5-6 线性分组码

S_1=000	E_1+C_1=00000	C_2=10111	C_3=01101	C_4=11010
S_2=111	E_2=10000	00111	11101	01010
S_3=101	E_3=01000	11111	00101	10010
S_4=100	E_4=00100	10011	01001	11110
S_5=010	E_5=00010	10101	01111	11000
S_6=001	E_6=00001	10110	01100	11011
S_7=011	E_7=00011	10100	01110	11001
S_8=110	E_8=00110	10001	01011	11100

5-6 线性分组码

S_1=000	E_1+C_1=00000	C_2=10111	C_3=01101	C_4=11010
S_2=111	E_2=10000	00111	11101	01010
S_3=101	E_3=01000	11111	00101	10010
S_4=100	E_4=00100	10011	01001	11110
S_5=010	E_5=00010	10101	01111	11000
S_6=001	E_6=00001	10110	01100	11011
S_7=011	E_7=00011	10100	01110	11001
S_8=110	E_8=00110	10001	01011	11100

对于接收码字**R**=(10101)可以选择以下3种方法进行译码：

(1)直接搜索标准阵列译码表；

(2)先求出伴随式，找到行数，再在该行中寻找；

(3)先求出伴随式，在表中查找差错图样，利用 **C=R+E** 进行译码。

5-6 线性分组码

生成标准阵列译码表的过程如下：

假设(n,k)线性分组码的2^k个不同的码字为

$$\mathbf{C}=\{\mathbf{C}_1=\mathbf{E}_1=\mathbf{0},\mathbf{C}_2,\cdots,\mathbf{C}_{2^k}\}$$

首先将这2^k个不同的码字排在表中的第一行，将$\mathbf{C}_1=\mathbf{E}_1$放在行的首位。然后取第一行($2^k$个不同的码字)中没有的，且是重量最轻的一个$n$重矢量$\mathbf{E}_2$,将$\mathbf{E}_2+\mathbf{E}_1=\mathbf{E}_2$作为第二行的行首元素，将$(\mathbf{E}_2+\mathbf{C}_i)$作为第二行中第$i$列的元素，接着再取第一第二行中没有的，且是重量最轻的一个n重矢量$\mathbf{E}_3$,将$\mathbf{E}_3+\mathbf{E}_1=\mathbf{E}_3$作为第三行的行首元素将$(\mathbf{E}_3+\mathbf{C}_i)$作为第三行中第$i$列的元素，这样一直继续下去，直到选取第$2^{(n-k)}$行前的$[2^{(n-k)}-1]$行没有的且是重量最轻的$n$重矢量$\mathbf{E}_2{}^{(n-k)}$将$\mathbf{E}_2{}^{(n-k)}+\mathbf{E}_1=\mathbf{E}_2{}^{(n-k)}$作为第$2^{(n-k)}$行的行首元素，将$(\mathbf{E}_2{}^{(n-k)}+\mathbf{C}_i)$作为第$i$列的元素。

5-6 线性分组码

S_1=000	E_1+C_1=00000	C_2=10111	C_3=01101	C_4=11010
S_2=111	E_2=10000	00111	11101	01010
S_3=101	E_3=01000	11111	00101	10010
S_4=100	E_4=00100	10011	01001	11110
S_5=010	E_5=00010	10101	01111	11000
S_6=001	E_6=00001	10110	01100	11011
S_7=011	E_7=00011	10100	01110	11001
S_8=110	E_8=00110	10001	01011	11100

译码讨论：

(1)如果收到n=5重矢量**R**=(10111)

如果码字是 **C**=(10111) 正确完成译码

此时 **E**=(00000)

5-6 线性分组码

S_1=000	E_1+C_1=00000	C_2=10111	C_3=01101	C_4=11010
S_2=111	E_2=10000	00111	11101	01010
S_3=101	E_3=01000	11111	00101	10010
S_4=100	E_4=00100	10011	01001	11110
S_5=010	E_5=00010	10101	01111	11000
S_6=001	E_6=00001	10110	01100	11011
S_7=011	E_7=00011	10100	01110	11001
S_8=110	E_8=00110	10001	01011	11100

(2)如果收到n=5重矢量**R**=(01100)

a)如果码字是 **C**=(01101),可以正确译码;

b)如果码字是 **C**=(00000),则发生错误译码;

5-6 线性分组码

$S_1=000$	$E_1+C_1=00000$	$C_2=10111$	$C_3=01101$	$C_4=11010$
$S_2=111$	$E_2=10000$	00111	11101	01010
$S_3=101$	$E_3=01000$	11111	00101	10010
$S_4=100$	$E_4=00100$	10011	01001	11110
$S_5=010$	$E_5=00010$	10101	01111	11000
$S_6=001$	$E_6=00001$	10110	01100	11011
$S_7=011$	$E_7=00011$	10100	01110	11001
$S_8=110$	$E_8=00110$	10001	01011	11100

(3)如果收到n=5重矢量**R**=(10100)

a)如果码字是 **C**=(10111),可以正确译码；

b)如果码字是 **C**=(01101),则发生错误译码；

5-6 线性分组码

结论：

这种线性分组码在保持自动纠正一位错误的纠错能力的同时不能兼有发现两位错误的能力。虽然对于特定的二位错误图样可以自动纠正错误，但是总体而言，它要么用于纠正一位错误，要么用于发现两位错误。

$$\mathbf{H}=\begin{bmatrix}1 & 1 & 1 & 0 & 0\\ 1 & 0 & 0 & 1 & 0\\ 1 & 1 & 0 & 0 & 1\end{bmatrix}$$

5-6 线性分组码

任何一个二元(n,k)线性分组码都有$2^{(n-k)}$个伴随式,如果该码字的纠错能力是e_c,那么伴随式的数目满足汉明限

$$2^{(n-k)} \geq C_n^0 + C_n^1 + C_n^2 + \cdots + C_n^{e_c} = \sum_{i=0}^{e_c} C_n^i$$

定义:

满足方程$2^{(n-k)} = \sum_{i=0}^{e_c} C_n^i$的二元$(n,k)$线性分组码称为完备码

如:(7, 4) 线性分组码

5-6 线性分组码——循环码

循环码是线性分组码的一个重要子集，是目前研究得最成熟的一类码。

1、通过严密的数学方法的研究分析，循环码有许多特殊的代数性质，这些性质有助于按所要求的纠错能力系统地构造这类码；

2、循环码的编码及译码易于用简单的具有反馈连接的移位寄存器来实现；

3、循环码的性能较好，具有较强的检错和纠错能力。

定义：对于一个(n,k)线性分组码，若某码字为$\mathbf{C}=(C_{n-1}C_{n-2}\cdots C_1C_0)$,该码字向左循环一位后为$\mathbf{C}^{(1)}=(C_{n-2}\cdots C_1C_0C_{n-1})$，向左循环移动i位后为$\mathbf{C}^{(i)}=(C_{n-i-1}C_{n-i-2}\cdots C_{n-i+1}C_{n-i})$，若$\mathbf{C}^{(i)}$，$i=1,2,3,\ldots,n-1$均为码字，则称这个$(n,k)$线性分组码为**循环码**.此处移位也可定义为向右移位.

5-6 线性分组码——循环码

（7，3）循环码举例

序号	码字		序号	码字	
	信息位 c6 c5c4	监督位 c3c2c1c0		信息位 c6 c5 c4	监督位 c3 c2 c1 c0
1	0 0 0	0 0 0 0	5	1 0 0	1 0 1 1
2	0 0 1	0 1 1 1	6	1 0 1	1 1 0 0
3	0 1 0	1 1 1 0	7	1 1 0	0 1 0 1
4	0 1 1	1 0 0 1	8	1 1 1	0 0 1 0

为了利用代数理论研究循环码，可以将码组用代数多项式来表示，这个多项式被称为**码多项式**,常常表示为:

$$\boldsymbol{C}(x)=C_{n-1}x^{n-1}+C_{n-2}x^{n-2}+\cdots+C_1x^{n-1}+C_0$$

其中**系数**C为0或1，x为**码元位置标志。**

如上表中第7个码字可以表示为:

$$\boldsymbol{C}_7(x)=1\cdot x^6+1\cdot x^5+0\cdot x^4+0\cdot x^3+1\cdot x^2+0\cdot x^1+1$$

$$=x^6+x^5+x^2+1$$

5-6 线性分组码——循环码

可以定义**码多项式的模运算法则**：若一任意多项式R(x)被一个n次多项式g(x)除，得到商式S(x)和一个次数小于n的余式e(x)，即

$$\frac{R(x)}{g(x)} = S(x) + \frac{e(x)}{g(x)} = \frac{e(x)}{g(x)} \bmod g(x)$$

通常可以表示为：$R(x) \equiv e(x)\ (模 g(x))$

例如,当$R(x)=x^4+x^2+1$ 除以$g(x)= x^3+1$

$$\frac{x^4+x^2+1}{x^3+1} = x + \frac{x^2+x+1}{x^3+1}$$

或表示为 $x^4+x^2+1 \equiv x^2+x+1\ \left(模 x^3+1\right)$

5-6 线性分组码——循环码

在循环码中，如果**码多项式**C(x)是个长度为n的许用码组，那么x^i**C**(x)在按模x^n+1运算下，也是一个许用码组。

即假如：$x^i \cdot \mathbf{C}(x) \equiv \mathbf{C}'(x)(模x^n+1)$，可以证明**C'**($x$)亦是一个许用码组，并且，正是**C**($x$)代表的码组向左循环移位$i$ 次的结果。

例如，前面表格中 的(7,3)分组码，现在给定一个 i=3

$$x^3 \cdot C_7(x) = x^3 \cdot \left(x^6 + x^5 + x^2 + 1\right) = \left(x^9 + x^8 + x^5 + x^3\right)$$

$$= \left(x^5 + x^3 + x^2 + x\right)\left(模x^7 + 1\right)$$

上述结果对应码组中第3个码字：0101110

结论：一个长度为n的循环码，它必为按模(x^n+1)运算的一个余式。

5-6 线性分组码——循环码

在循环码中，一个 (n，k) 码有 2^k个不同的许用码组。若用 **$g(x)$**表示其中 前 (k -1) 位皆为 0 的许用码组，则 $g(x), xg(x), x^2g(x), \ldots x^{k-1}g(x)$都是许用码组，而且这 k个码组是线性无关的。因此它们可以用 来构成循环码的**生成矩阵 G** 。

$g(x)$ 的确定：在循环码中除全"0"码组外，再没有连续 k 位均为"0"的码组，即连续"0"的长度最多只能有 (k - 1) 位。否则，在进行若干次循环 移位后将得到一个k位信息位全为"0"，但监督位不全为 "0" 的码组。**因此$g(x)$必须是一个常数项不为 "0"的 n - k 次多项式**，**称这唯一的 (n - k) 次多项式$g(x)$为码的生成多项式**。一旦确定了$g(x)$，则整个 (n，k)循环码就被确定了。

5-6 线性分组码——循环码

循环码的**生成多项式**用$g(x)$表示，它是在2^k个码字的码多项式中取一个次数最低(即n-k次)的多项式。它具有以下特性:

1) $g(x)$是一个常数项为1的 $r=n-k$ 次多项式;

2) $g(x)$是 x^n+1 的一个因式;

3) 该循环码其中其他码字多项式都是$g(x)$的倍式。

令 $g(x)=x^{n-k}+g_{n-k-1}x^{n-k-1}+\cdots+g_2x^2+g_1x+1$

循环码的生成矩阵为

$$G(x)=\begin{bmatrix} x^{k-1}g(x) \\ \vdots \\ xg(x) \\ g(x) \end{bmatrix}=\begin{bmatrix} 1 & g_{n-k-1} & g_{n-k-2} & \cdots & g_1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & g_{n-k-1} & g_{n-k-2} & \cdots & g_1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & g_{n-k-1} & g_{n-k-2} & \cdots & g_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & g_{n-k-1} & g_{n-k-2} & \cdots & \cdots & 1 \end{bmatrix}$$

5-6 线性分组码——循环码

例题： 分析上述(7,3)循环码，构造其生成矩阵和生成多项式

该码组n=7,k=3,r=n-k=4.

其生成多项式可以用第2个码字构造,即寻找$x^1,x^2\ldots x^{k-1}$,

而k-1=3-1=2，与$\boldsymbol{g(x)}$相乘构造生成矩阵

$$g(x) = x^4 + x^2 + x + 1$$

$$G(x) = \begin{bmatrix} x^2 g(x) \\ x g(x) \\ g(x) \end{bmatrix} = \begin{bmatrix} x^6 + x^4 + x^3 + 1 \\ x^5 + x^3 + x^2 + 1 \\ x^4 + x^2 + x + 1 \end{bmatrix} \Rightarrow G = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

构造循环码的步骤：

1) 对(x^n+1)作因式分解，找出n-k 次因式；

2) 以该n-k 次因式作为生成多项式，与不高于$(k$-1)次的信息多项式相乘得到码多项式$C(x)$=m(x)g(x),C(x)的次数不高于$(n$-1)次。

5-6 线性分组码——循环码

循环码校验矩阵的生成方法1：

例题： 已知上述(7,3)循环码，其生成矩阵如下，试构造其一致校验矩阵。

由于 $G=\begin{bmatrix}1&0&1&1&1&0&0\\0&1&0&1&1&1&0\\0&0&1&0&1&1&1\end{bmatrix}$ 并非系统矩阵，因此可以通过

矩阵变换得到系统矩阵。将矩阵第3行加到矩阵第一行得到：

$$G'=\begin{bmatrix}1&0&0&1&0&1&1\\0&1&0&1&1&1&0\\0&0&1&0&1&1&1\end{bmatrix}$$

再得到

$$H^T=\begin{bmatrix}1&0&1&1\\1&1&1&0\\0&1&1&1\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$$

5-6 线性分组码——循环码

循环码校验矩阵的生成方法2：

由于(n,k)循环码中生成多项式g(x)是x^n+1的因式，因此可知：

$$h(x)=\frac{x^n+1}{g(x)}=x^k+h_{k-1}x^{k-1}+\cdots+h_1x+1$$

将h(x)称为该循环码的**一致校验多项式**。

与前面类似，**一致校验矩阵**表示为：

$$H(x)=\begin{bmatrix} x^{n-k-1}h^*(x) \\ \vdots \\ xh^*(x) \\ h^*(x) \end{bmatrix}$$

式中h*(x)是h(x)的逆多项式：

$$h^*(x)=x^k+h_1x^{k-1}+h_2x^{k-2}+\cdots+h_{k-1}x+1$$

5-6 线性分组码——循环码

最终得到一致校验矩阵：

$$H=\begin{bmatrix} h_0 & h_1 & \cdots & h_k & 0 & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_k & 0 & \cdots & 0 \\ 0 & 0 & h_0 & h_1 & \cdots & h_k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & h_0 & h_1 & \cdots & h_k \end{bmatrix}$$

注意，这样得到的一致校验矩阵不是系统形的，但是通过矩阵变换可以将其变成系统形的一致校验矩阵。

5-6 线性分组码——循环码

例题： 上述(7,3)循环码，其生成多项式为 $g(x)=x^4+x^2+x+1$

可得 $h(x)=\dfrac{x^7+1}{g(x)}=x^3+x+1$ 即得到 $h^*(x)=x^3+x^2+1$

从而 $H(x)=\begin{bmatrix} x^{7-3-1}h^*(x) \\ x^{7-3-2}h^*(x) \\ xh^*(x) \\ h^*(x) \end{bmatrix}=\begin{bmatrix} x^6+x^5+x^3 \\ x^5+x^4+x^2 \\ x^4+x^3+x^1 \\ x^3+x^2+1 \end{bmatrix}$

$$H=\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix} \xrightarrow[\text{矩阵1行加到第3行}]{\text{矩阵1,3行加到第4行}} H=\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

5-6 线性分组码——循环码

在编码时，首先需要根据给定循环码的参数确定**生成多项式g(x)**，也就是从x^n+1的因子中选一个（n-k）次多项式作为**g(x)**；然后，利用循环码的编码特点，即所有**循环码多项式**C(x)都可以被**g(x)**整除，来定义**生成多项式g(x)**。

系统循环码编码方法：设要产生（n,k）循环码，m(x)表示信息多项式，则其次数必小于k，而$x^{n-k}\cdot m(x)$的次数必小于n，用$x^{n-k}\cdot m(x)$ 除以g(x)，可得余数r(x)，r(x)的次数必小于（n-k），将r(x)加到信息位后作监督位，就得到了系统循环码。

5-6 线性分组码——循环码

循环码的编码过程：

1)用x^{n-k} 乘m(x)。

这一运算实际上是把信息码后附加上（n-k）个“0”。例如，信息码为110，它相当于m(x) = x^2+x。当n-k = 7-3 = 4时，$x^{n-k}\cdot m(x)=x^6+x^5$，它相当于1100000。而希望到得的系统循环码多项式应当是C(x) = $x^{n-k}\cdot m(x)+r(x)$。

2)计算r(x)。

由于循环码多项式R(x)都可以被g(x)整除，也就是：

$$\frac{R(x)}{g(x)}=\frac{x^{n-k}m(x)+r(x)}{g(x)}=\frac{x^{n-k}m(x)}{g(x)}+\frac{r(x)}{g(x)}$$

此时用$x^{n-k}\cdot m(x)$除以g(x),就得到商Q(x)和余式r(x)，即

$$\frac{x^{n-k}m(x)+r(x)}{g(x)}=Q(x)+\frac{r(x)}{g(x)}$$

从而得到r(x)

5-6 线性分组码——循环码

3)编码输出系统循环码多项式R(x)。

$$R(x)=x^{n-k}m(x)+r(x)$$

例如，对于（7，3）循环码，如果选用 $g(x)=x^4+x^2+x+1$ 信息码为110，则:

$$\frac{x^{n-k}m(x)}{g(x)}=\frac{x^6+x^5}{x^4+x^2+x+1}=\left(x^2+x+1\right)+\frac{x^2+1}{x^4+x^2+x+1}$$

它相当

$$\frac{x^{n-k}m(x)}{g(x)}=\frac{1100000}{10111}=111+\frac{0101}{10111}$$

此时编码输出 1100101

循环码的译码过程：类似其他线性分组码

1）由接收到的码多项式R(x)计算校正子（伴随式）多项式S(x);

2）由校正子S(x)确定错误图样E(x);

3）将错误图样E(x)与R(x)相加，纠正错误,C(x)=R(x)+E(x)。