线性分组码

目录

- 数学描述
- q进制对称信道上的译码
- 码间最小距离与纠错能力
- 重量枚举多项式及译码错误概率估计
- 常见的线性分组码

数学描述

- 定义

  GF(q)上的 (n, k) 线性分组码,

  是 n维矢量空间$V_n[GF(q)] = \{(x_1、x_2、\ldots x_n): x_i \in GF(q)\}$的一个k维子空间.

  n是码长

  k是信息位数

  n-k是校验位数

  k/n是编码速率

数学描述

- 生成矩阵:

　　对于GF(q)上的 (n, k) 线性分组码C, 若矩阵G的行空间等于C, 则 **G**称为 C的<u>**生成矩阵**</u>.

　　反之, 如果**G**是由 GF(q)上元素构成的矩阵, G的行空间称作生成的码.

设 **u**是 $(1\times k)$ 维信息矢量

G是 $(k\times n)$ 维的生成矩阵

x是编码后的 $(1\times n)$维码字

则 $\mathbf{x}=\mathbf{u}\cdot\mathbf{G}$

(7,4)Hamming Code

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

数学描述

- 系统线性分组码及其生成矩阵:

如果线性分组码 C, 码字 $\mathbf{x}\in C$的高 k个分量恰好等于信息矢量$\mathbf{u}$的 k个分量, 则码 C称作**系统码**.

系统码的生成矩阵G是 $(k\times n)$ 维的, 它可以写作 $\mathbf{G} = (\mathbf{AI})$.

其中$\mathbf{I}$是 $(k\times k)$ 维单位矩阵

$\mathbf{A}$是 $[k\times(n-k)]$ 维矩阵

(7,4)Hamming Code

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

数学描述

- 对偶码

GF(q)上(n, k) 线性码 C的校验方程是：

$$\langle \mathbf{a} \cdot \mathbf{x} \rangle = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0 \quad \forall \mathbf{x} \in C$$

对偶码$C^{\perp}$是使校验方程成立的全部矢量的集合.

$C^{\perp}$也是V_n[GF(q)]上的子空间, 是n维矢量空间中的n-k维子空间, $C^{\perp}$是GF(q)上的一个(n, n-k) 线性分组码.

数学描述

- 校验矩阵

  设C是GF(q)上的一个 (n, k) 线性码,

  若矩阵**H**满足

  $$\mathbf{H}\mathbf{x}^{T} = 0,\ \text{当且仅当}\ x \in C\ ,$$

  则**H**称作码C的<u>**校验矩阵**</u>.

  H是[(n-k)×n]维的，

  是对偶码$C^{\perp}$的生成矩阵.

(7,4)Hamming Code

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

数学描述

HG^T=?

- 系统线性分组码的校验矩阵

定理11.1: 如果(n, k)系统码C的生成矩阵

$$\mathbf{G} = (\mathbf{A}\mathbf{I})$$

则其校验矩阵$\mathbf{H} = (\mathbf{I}\mathbf{A}^T)$,

其中$\mathbf{A}^T$是$\mathbf{A}$的转置, 是一个(n-k)×k阶矩阵.

I是(n-k)×(n-k)阶单位矩阵.

即

$$\mathbf{G} = \left[\mathbf{A}_{k\times(n-k)}\mathbf{I}_k\right]$$

$$\mathbf{H} = \left[\mathbf{I}_{n-k}\mathbf{A}_{(n-k)\times k}{}^T\right]$$

(7,4)Hamming Code

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

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

数学描述

- 小结 $\langle \mathbf{a} \cdot \mathbf{x} \rangle = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0 \quad \forall \mathbf{x} \in C$

	线性码C	对偶码$C^{\perp}$
参数	（n, k)	（n, n-k)
定义	n维空间中的k维子空间	n维空间中的n-k维子空间
码字个数	2^k	2^{n-k}
生成矩阵	$\mathbf{G}_{k\times n}$ $\mathbf{x} = \mathbf{u} \cdot \mathbf{G}$	$\mathbf{H}_{(n-k)\times n}$
校验矩阵	$\mathbf{H}_{(n-k)\times n}$ $\mathbf{H}\mathbf{x}^T = 0$	$\mathbf{G}_{k\times n}$

目录

- 数学描述
- q进制对称信道上的译码
- 码间最小距离与纠错能力
- 重量枚举多项式及译码错误概率估计
- 常见的线性分组码

Q进制对称信道上的译码

- 汉明重量：

  码字**x**的汉明重量是它的非零分量的个数，记作$W_H(\mathbf{x})$.

- 汉明距离：

  码字**x**和**y**之间的汉明距离是它们不相同分量的个数，记作$d_H(\mathbf{x}, \mathbf{y})$.

➢ 汉明重量与汉明距离的关系：

$$d_H(\mathbf{x},\mathbf{y}) = W_H(\mathbf{x}-\mathbf{y})$$

码间最小距离等于其非零码字的最小重量，即

$$d_{H,\min}(\mathbf{x},\mathbf{y}) = W_{H,\min}(\mathbf{x}-\mathbf{y})$$

Q进制对称信道上的译码

- q进制对称信道:

  输入$\mathbf{x} \in V_n[GF(q)]$，输出 $\mathbf{y} \in V_n[GF(q)]$

  $$\mathbf{y} = \mathbf{x} + \mathbf{z}$$

  $\mathbf{z}$是错误图案，$z_i \neq 0$意味着第i个分量出错，

  $\mathbf{z}$中非零分量的个数记作$W_H(\mathbf{z})$.

$$p(\mathbf{y}/\mathbf{x}) = p(\mathbf{z})$$

$$\begin{cases} p(z_i) = 1-(q-1)\varepsilon & if\ z_i = 0 \\ p(z_i) = \varepsilon & if\ z_i = 1, 2, \cdots, q-1 \end{cases}$$

$$p(\mathbf{z}) = \left[1-(q-1)\varepsilon\right]^{n-W_H(\mathbf{z})} \varepsilon^{W_H(\mathbf{z})}$$

Q进制对称信道上的译码

- 译码准则:

  接收到矢量**y**,采用**最大似然概率译码准则**，译码输出的码字**x**应使下列信道转移概率最大.

$$p(\mathbf{y}/\mathbf{x})=p(\mathbf{z})=[1-(q-1)\varepsilon]^{n-W_H(\mathbf{z})}\varepsilon^{W_H(\mathbf{z})}$$

$$=[1-(q-1)\varepsilon]^n\left(\frac{\varepsilon}{1-(q-1)\varepsilon}\right)^{W_H(\mathbf{z})}$$

$$=[1-(q-1)\varepsilon]^n\left(\frac{\varepsilon}{1-(q-1)\varepsilon}\right)^{W_H(\mathbf{y}-\mathbf{x})}$$

Q进制对称信道上的译码

- 译码准则:

  选择的**x**应使$W_H(\mathbf{y}-\mathbf{x})$最小，

  即使**y**和**x**之间的汉明距离最小，

  应采用**最小汉明距离译码准则**.

Q进制对称信道上的译码

- 伴随式：$\mathbf{s} = \mathbf{H}\mathbf{y}^{\mathrm{T}} = \mathbf{H}(\mathbf{x}^{\mathrm{T}}+\mathbf{z}^{\mathrm{T}}) = \mathbf{H}\mathbf{z}^{\mathrm{T}}$

  其中$\mathbf{H}$是校验矩阵，$\mathbf{y}$是接收矢量.

- 陪集：$C + \mathbf{z}_0 = \{ \mathbf{x} + \mathbf{z}_0,\ x \in C\}$

  称作以$\mathbf{z}_0$为陪集首的线性码C的陪集.

0000000	1000000	0100000	0010000	0001000	0000100	0000010	0000001
1010001	0010001	1110001	1000001	1011001	1010101	1010011	1010000
1110010	0110010	1010010	1100010	1111010	1110110	1110000	1110011
0100011	1100011	0000011	0110011	0101011	0100111	0100001	0100010
0110100	1110100	0010100	0100100	0111100	0110000	0110110	0110101
1100101	0100101	1000101	1110101	1101101	1100001	1100111	1100100
1000110	0000110	1100110	1010110	1001110	1000010	1000100	1000111
0010111	1010111	0110111	0000111	0011111	0010011	0010101	0010110
1101000	0101000	1001000	1111000	1100000	1101100	1101010	1101001
0111001	1111001	0011001	0101001	0110001	0111101	0111011	0111000
0011010	1011010	0111010	0001010	0010010	0011110	0011000	0011011
1001011	0001011	1101011	1011011	1000011	1001111	1001001	1001010
1011100	0011100	1111100	1001100	1010100	1011000	1011110	1011101
0001101	1001101	0101101	0011101	0000101	0001001	0001111	0001100
0101110	1101110	0001110	0111110	0100110	0101010	0101100	0101111
1111111	0111111	1011111	1101111	1110111	1111011	1111101	1111110

Q进制对称信道上的译码

- 定理11.2: 同一陪集中的矢量具有相同的伴随式; 不同陪集中的矢量具有不同的伴随式; 所有q^{n-k}种可能的伴随式作为某个陪集的伴随式出现.

- 定理11.3: 如果C是二进制码，且**e**是一个陪集首, 则由**e**生成的陪集C+**e**所对应的伴随式，等于**H**中与**e**中非零分量所对应的列之和.

Q进制对称信道上的译码

- 利用伴随式译码:
  - 计算 $\mathbf{s} = \mathbf{H}\mathbf{y}^{\mathrm{T}}$
  - 确定错误图案：在给定伴随式**s**对应的赔集中， 寻找最有可能的错误图案, 即相应重量 $W_H(\mathbf{z})$最小的错误图案$\mathbf{z}_0$，使$p(\mathbf{z}_0)$最大
  - 输出 $\hat{\mathbf{x}} = \mathbf{y} - \mathbf{z}_0$

Q进制对称信道上的译码

- 利用标准阵列译码:
  - 按照陪集分解的方法构造标准阵列，标准阵列的陪集首是这一陪集中作为错误的图案最可能出现的;
  - 收到接收矢量后，在标准阵列中找到对应的元素，它所在行/列的第一个元素就是相应的译码输出.

目录

- 数学描述
- q进制对称信道上的译码
- 码间最小距离与纠错能力
- 重量枚举多项式及译码错误概率估计
- 常见的线性分组码

码间最小距离和纠错能力

- $d_{H,min}$与校验矩阵的关系:

定理11.4: 码C是GF(q)上(n, k)线性分组码, 其校验矩阵**H**中线性相关列矢量组的最小数目等于$d_{H,min}(C)$及$W_{H,min}(C)$.

推论11.1: 校验矩阵**H**中任意($d_{H,min}(C)$-1)个列矢量是线性不相关的.

(7,4)Hamming Code

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

码间最小距离和纠错能力

- **定理11.5**

　　辛格尔顿界：任意（n，k）线性分组码的码间最小距离d_{min}满足不等式$d_{min} \leq n-k+1$.

分析：

- （n，k）系统线性分组码中存在只有一个信息位非零，其它信息位全为零的码字
- 该码字的校验位中至多可以有n-k位非零
- 因此该码字的最大重量不超过n-k+1,即该码与全零码的最小距离至多是n-k+1
- 因此有$d_{min} \leq n-k+1$

码间最小距离和纠错能力

- **定理11.6**

  普洛特金界： GF(q)上（q进制）（n，k）线性分组码的最小距离d_{min}满足不等式：

$$d_{\min} \le \frac{n(q-1)q^{k-1}}{q^k - 1}$$

码间最小距离和纠错能力

- **定理11.6证明：**
  - GF(q)上（n,k）线性分组码总共有q^k个码字
  - 第一个分量不为零的码字总数是$(q-1)q^{k-1}$
  - 同样其他分量不为零的码字总数也都是$(q-1)q^{k-1}$
  - 由此可知该码的总重量是$n(q-1)q^{k-1}$
  - 只有当这$n(q-1)q^{k-1}$个非零分量平均地分配到q^k-1个非零码字时，才能得到最大的w_{min}，即得到最大的码间d_{min}
  - 因而码间最小距离满足

$$d_{\min} \le \frac{n(q-1)q^{k-1}}{q^k - 1}$$

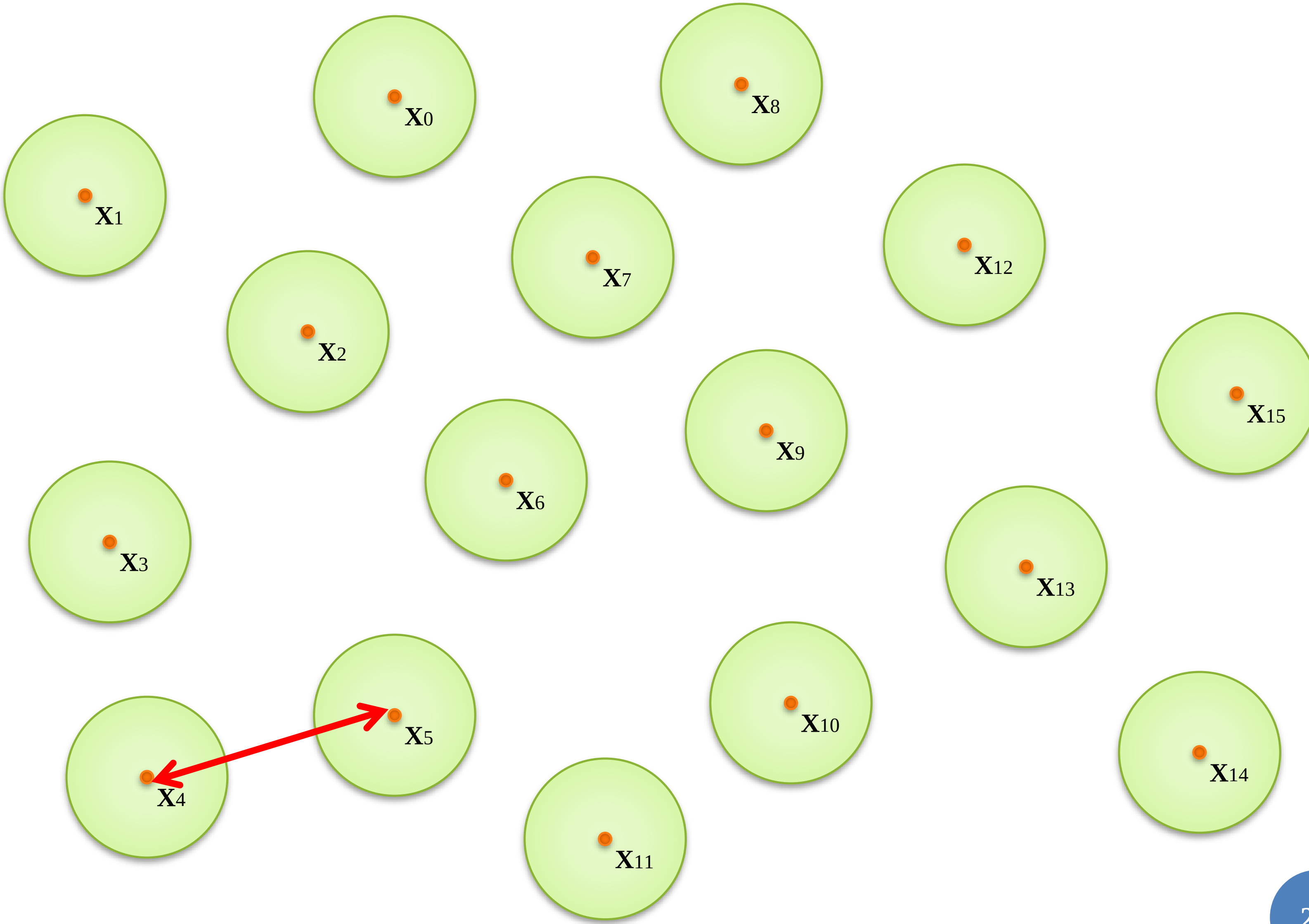
0000000
1010001
1110010
0100011
0110100
1100101
1000110
0010111
1101000
0111001
0011010
1001011
1011100
0001101
0101110
1111111

码间最小距离和纠错能力

- 纠错能力与最小汉明距离$d_{H,min}$的关系:

定理11.7: 码$C=\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M\}$ 能纠正所有非零个数小于等于t的错误图案$\mathbf{z}$, 当且仅当

$$t \le \lfloor (d_{min}-1)/2 \rfloor.$$


X0
X1
X2
X3
X4
X5
X6
X7
X8
X9
X10
X11
X12
X13
X14
X15

码间最小距离和纠错能力

- **定理11.8**

  汉明界：任何GF(q)上（q进制）能纠正t个错误的（n，k）线性分组码满足汉明不等式：

$$q^{n-k} \geq 1 + (q-1)\binom{n}{1} + (q-1)^2\binom{n}{2} + \cdots + (q-1)^t\binom{n}{t}$$

码间最小距离和纠错能力

- **定理11.8证明：**
  - 一个（n, k）码共有q^k个码字，最多有q^{n-k}个伴随式；
  - 若要纠正所有不多于t个错误的图案，则要求所有不多于t个错误的图案所对应的伴随式各不相同；
  - 没有错误的全零错误图案的个数、含一个错误的错误图案的个数、含两个错误的错误图案的个数、直到含t个错误的错误图案的个数的总和，是

$$1+(q-1)\binom{n}{1}+(q-1)^2\binom{n}{2}+\cdots+(q-1)^t\binom{n}{t}$$

  - 能纠正t个错误的（n, k）码的伴随式总数满足

$$q^{n-k}\geq 1+(q-1)\binom{n}{1}+(q-1)^2\binom{n}{2}+\cdots+(q-1)^t\binom{n}{t}$$

码间最小距离和纠错能力

- **定理11.8**

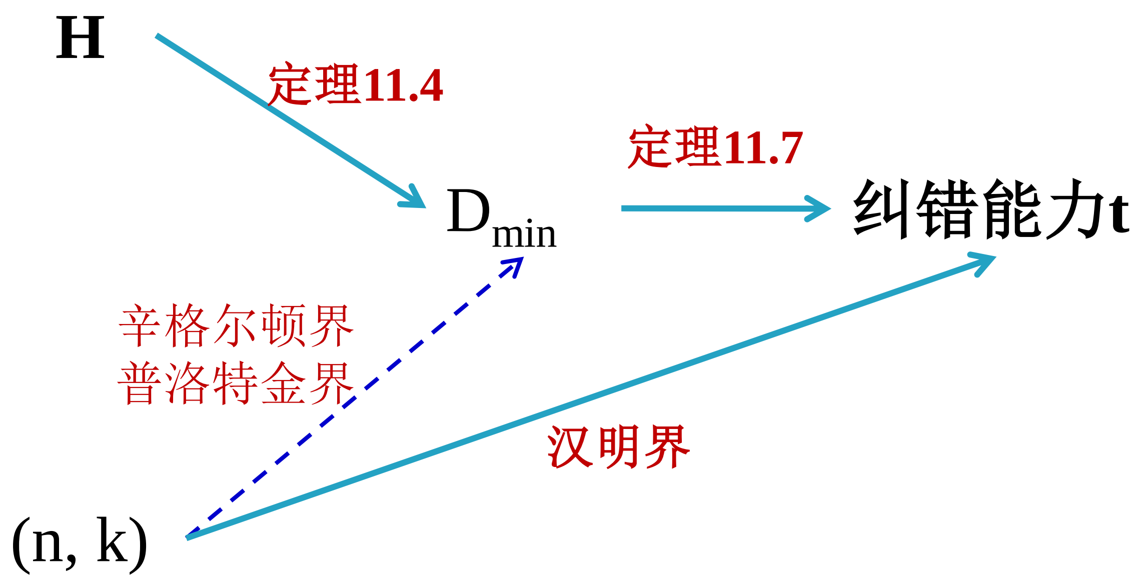
当只考虑二进制（q=2）编码时，汉明界可以用下列不等式表达：

$$2^{n-k} \geq 1+\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{t}$$

<u>思考题</u>：能否构造一个二进制（15，11）线性分组码，纠正t=2个信道传输错误？

码间最小距离和纠错能力

- **小结：**



目录

- 数学描述
- q进制对称信道上的译码
- 码间最小距离与纠错能力
- 重量枚举多项式及译码错误概率估计
- 常见的线性分组码

重量枚举多项式及译码错误概率估计

- 线性分组码的**纠错能力**主要取决于它的**码字之间的距离**
- 对于**线性码**只需考虑全零码字与其它非零码字的距离，即**非零码字的重量**
- 码字的重量分布是用**重量枚举多项式**来描述的，线性分组码的重量枚举多项式表示为：

$$A(D)=\sum_{d=0}^{n}A_d D^d$$

其中A_d表示重量为d的非零码字的个数

重量枚举多项式及译码错误概率估计

- 对于二进制对称信道，设传输的错误概率为ε，根据译码错误概率的并合界公式，传输码字的译码错误概率满足：

$$P_{block} \le \sum_{d=0}^{n} A_d z^d - 1$$

其中 $$z = 2\sqrt{\varepsilon(1-\varepsilon)}$$

重量枚举多项式及译码错误概率估计

- 例：(7,4)Hamming码

$$A(D)=1+7D^3+7D^4+D^7$$

$$P_{block}\le 7z^3+7z^4+z^7$$

$$z=2\sqrt{\varepsilon(1-\varepsilon)}$$

$$P_{block}=\sum_{k=2}^{7}\binom{7}{k}\varepsilon^k(1-\varepsilon)^{7-k}=21\varepsilon^2-70\varepsilon^3+etc.$$

重量枚举多项式及译码错误概率估计

- **线性分组码的性能评价**
  - 非零码字的最小重量应该尽可能地大
  - 对应具有最小重量的非零码字的个数应该尽可能地少

目录

- 数学描述
- q进制对称信道上的译码
- 码间最小距离与纠错能力
- 重量枚举多项式及译码错误概率估计
- 常见的线性分组码

常见的线性分组码

(7,4)Hamming Code

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

- Hamming 码

二进制Hamming 码:

码长 $n = 2^m-1$

信息位数 $k = 2^m-m-1$

校验位数 $n-k = m$

校验矩阵H是$m\times(2^m-1)$维矩阵,

其列矢量是 所有2^m-1个二进制非零列矢量的某种排列.

问题：码长最短的汉明码？**G**？**H**?

常见的线性分组码

- Hamming 码

q进制Hamming 码:

码长 $n = (q^m-1)/(q-1)$

信息位数 $k = n-m$

校验位数 $n-k = m$

校验矩阵H是$m \times n$维矩阵，

其列矢量是所有$(q^m-1)/(q-1)$个q进制两两互不相关的非零列矢量的某种排列.

以q=3，m=3为例说明！

常见的线性分组码

- Hamming 码

纠错能力

校验矩阵的任意两个列矢量线性不相关

- 非零码字的重量大于等于3
- 码间最小距离大于等于3
- 能纠正单个错误

常见的线性分组码

- Hamming 码

例11.1：（7,4）Hamming 码，其生成矩阵、校验矩阵及码字分量与信息比特的关系如下：

$$H = \begin{pmatrix} 1001011 \\ 0101110 \\ 0010111 \end{pmatrix}$$

$$G = \begin{pmatrix} 1101000 \\ 0110100 \\ 1110010 \\ 1010001 \end{pmatrix}$$

$$x_0 = x_3 + x_5 + x_6$$
$$x_1 = x_3 + x_4 + x_5$$
$$x_2 = x_4 + x_5 + x_6$$
$$x_3 = u_0$$
$$x_4 = u_1$$
$$x_5 = u_2$$
$$x_6 = u_3$$

常见的线性分组码

- Hadamard码
  - Hadamard码是将Hadamard矩阵及其补阵的行矢作为码字而构成的
  - Hadamard矩阵$\mathbf{M}_n$是一个由“0”和“1”组成的n×n阶矩阵（n为偶数）
  - 矩阵每一行恰好有n/2个元素与其它行不同
    - 其中一行全部是“0”
    - 而其余行包含 n/2个“0”和n/2个“1”

常见的线性分组码

- Hadamard码

当n=2时，Hadamard矩阵为

$$\mathbf{M}_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

而通过下面的关系，可以由$\mathbf{M}_\mathrm{n}$得到$\mathbf{M}_{2\mathrm{n}}$

其中 $\overline{\mathbf{M}}_n$是 $\mathbf{M}_\mathrm{n}$的补阵

$$\mathbf{M}_{2n} = \begin{bmatrix} \mathbf{M}_n & \mathbf{M}_n \\ \mathbf{M}_n & \overline{\mathbf{M}}_n \end{bmatrix}$$

常见的线性分组码

- Hadamard码

根据上面介绍，可得到M4和$\overline{\mathbf{M}}_4$如下

$$\mathbf{M}_4=\begin{bmatrix}0&0&0&0\\0&1&0&1\\0&0&1&1\\0&1&1&0\end{bmatrix}\qquad\overline{\mathbf{M}}_4=\begin{bmatrix}1&1&1&1\\1&0&1&0\\1&1&0&0\\1&0&0&1\end{bmatrix}$$

$\mathbf{M}_4$和 $\overline{\mathbf{M}}_4$ 的行可以构成2n = 8个长度为 n = 4的二进制线性分组码

其最小距离d_{min}=n/2=2

常见的线性分组码

- Hadamard码

一般的，Hadamard码的参数如下：

码长: $n = 2^m$

信息位数: $k = \log_2 2n = \log_2 2^{m+1} = m+1$

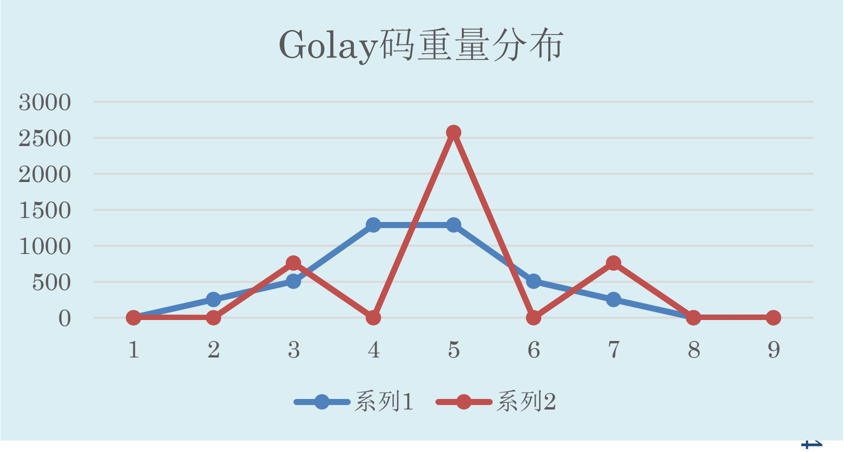
最小距离: $d_{\min} = \frac{1}{2} n = 2^{m-1}$

问题：重量枚举多项式？

常见的线性分组码

- Golay码
  - Golay码是一种最小重量$d_{min} = 7$的二进制（23，12）线性分组码
  - 扩展Golay码是增加了偶校验位的Golay码，是一种最小重量$d_{min} = 8$的二进制（24，12）线性分组码
  - Golay码和扩展Golay码的重量分布见表11-1:

常见的线性分组码



- Golay码

表11-1 （23，12）Golay码和（24，12）扩展Golay码的重量分布

重　　量	码字数目 （23，12）码	码字数目 （24，12）码
0	1	1
7	253	0
8	506	759
11	1288	0
12	1288	2576
15	506	0
16	253	759
23	1	0
24	0	1

作业

- 习题11.2
- 习题11.5(1), (2)
- 习题11.8
- 习题11.9
- 习题11.10