LINEAR CODES

$C \subseteq \{0,1\}^n$

number of codewords

in general inefficient to calculate

I. C has $(n, M, \boxed{d})$ parameters

length of codewords

minimum distance

II. ENCODING

$m_0 \to c_0$
$\vdots$
$m_M \to c_M$

this table needs to be stored to encode with general C

III DECODING

for each recieved w we need to find $c \in C$ that is closest to w in Hamming distance.

need to compare to each $c \in C$

LINEAR CODES - All tasks I.-III. are much more efficient.

DEFINITION

Code C is linear if two conditions hold:

Code C is linear if two conditions hold:

mod 2 (for bit codes)

(1.) $\forall x, y \in C \Rightarrow x \oplus y \in C$

$$x = (x_0, \dots, x_{n-1})$$
$$y = (y_0, \dots, y_{n-1})$$
$$x + y = (x_0 + y_0, x_1 + y_1, \dots, x_{n-1} + y_{n-1})$$

(2.) $\forall k \in \underbrace{\{0, \dots, q-1\}}_{\text{for } q\text{-ary codes}}$

$$c = (c_0, \dots, c_{n-1})$$

$$\forall c \in C : k \cdot c \in C$$

$$k \cdot c = (k \cdot c_0, \dots, k \cdot c_{n-1})$$

$\boxed{\text{Ex. 2.1}}$

$C = \{00000, 00110, 10010, 10001, 10100, 00011, 10111, 00101\}$

10100   00011

Code is linear!

How to calculate the minimum distance d of a linear code C.

$$H(00110, 10010) = 2$$

$$H(00110 + 00110, 10010 + 00110) = 2$$

$$\parallel$$

$$H(00000, 10100)$$

$$H(w, w')$$
$$\parallel$$
$$H(w + w, w' + w)$$
$$\parallel$$
$$H(0, w' + w)$$

$$\overset{\shortparallel}{H(00000,\ 10100)} \qquad\qquad H(0,\ \overset{\shortparallel}{w'+w}) \quad w'+w \in C$$

$$\Downarrow$$

To calculate d we need to find the codeword $c \in C$ with the smallest weight (number of non-zero positions).

ENCODING

n

$M = q^k$ (for binary 2^k)

$[n, k]$ — n: length, k: size of subspace

For each linear code C you can find a basis of size k. That means there are k codewords, such that all codewords can be written as a linear combination of the k codewords.

$$C = \{00000, 00110, 10010, 10001, \underbrace{10100}_{1+2}, \underbrace{00011}_{2+3}, \underbrace{10111}_{1+3}, \underbrace{00101}_{1+2+3}\}$$

$$\left.\begin{array}{ll} 1 & 00110 \\ 2 & 10010 \\ 3 & 10001 \end{array}\right\} \text{linearly independent} \Rightarrow \text{they form a basis of } C$$

$$C = \{000, 111, 222\}$$

$(+, \cdot)$ are from $\mathbb{F}_q$ (mod q) for q prime

$2 \cdot (000) = (000) \in C$

I. ✓

II. $k \in \{0, 1, 2\}$ ✓ ✓

for q prime

$2 \cdot (000) = (000) \in C$

$2 \cdot (111) = (222) \in C$

$2 \cdot (222) = (444) = (111) \in C$

$3^k \Rightarrow k = 1$ basis is $\{111\}$

$[n, k]$ size of subspace

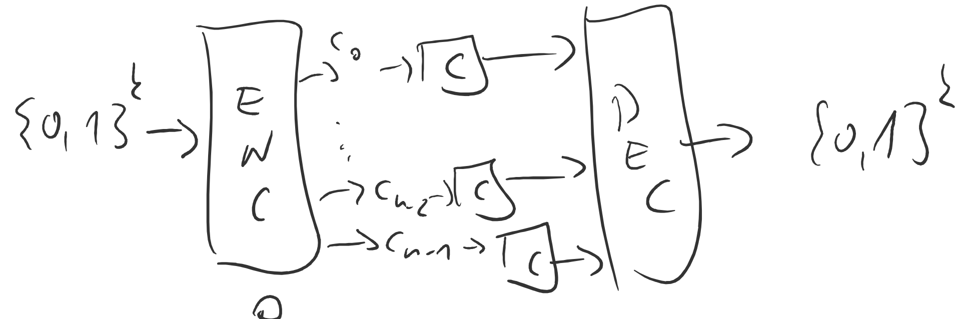
size of code word

$M = 2^k$

$C \in \{0,1,2\}^n$, $|C| = 9$ basis has $\underline{2}$ vectors $\{b_1, b_2\}$

and each $c = k_1 \cdot b_1 + k_2 \cdot b_2$

$k_i \in \{0,1,2\}$



Matrix $G = \begin{bmatrix} b_0 \\ \vdots \\ b_{k-1} \end{bmatrix}$ $n \times k$ matrix is a generating matrix of a linear code with basis $= \{b_0, \dots, b_{k-1}\}$

To encode message $\underline{m \in \{0,1\}^k}$.

calculate

$$m \cdot G = (m_0, \dots, m_{k-1}) \cdot G = (c_0, \dots, c_{n-1})$$

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

$$m_1 \cdot G = (0\,0\,1) \cdot \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$m_1 = 001$

$m_2 = 101$

$$= (0\cdot 0 + 0 \cdot 1 + 1 \cdot 1)$$

$$\parallel$$

$$(1\,0\,0\,0\,1) = c_1$$

$$m_2 \cdot G = (101) \cdot G = (1\,0\,1\,1\,1) = c_2$$

$$m_3 = 011 = b_2 + b_3 = (0\,0\,0\,1\,1)$$

For encoding we do not need a lookup table. We only need to store the generating matrix G. EFFICIENT.

Normal form of a code G

$$\boxed{G = (I_k \mid A)}$$

A: $k \times (n-k)$ matrix (checksum matrix)

I_k: $k \times k$ identity matrix

How to obtain a normal form?

Start from an arbitrary G of code C.

Use the following operations:

a.) permutation of rows

b.) multiplication of a row by a non-zero scalar
c.) addition of rows

$\Big\} \to$ change of basis

$\to$ d.) multiplication of columns by a non-zero scalar
e.) permutation of columns

$\Big\} \to$ different but equivalent code

$$G := \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \ldots \text{ operations a–e such that } G' = (I_3 \mid A)$$

I_3: (3×3), A: (3×2)

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} \sim \left(\begin{array}{ccccc|} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right) \xrightarrow{+r_2} \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right)$$

G' is of canonical form and has the same $[n, k, d]$ parameters as G

Codes with G in canonical form are called systemic.

$$(a\,b\,c) \cdot \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} = (\underbrace{a, b, c}_{\text{data part (the message)}}, \underbrace{0, (a+b+c)}_{\text{check sum}})$$

Decoding

Standard array

$$\{c+u \mid c \in C\} = \text{coset } \underline{u} \in \{0,1\}^n$$

$\forall u,v$ coset u and coset v are either identical or disjoint

Coset leader is an element of a coset with smallest weight.

Ex. 2.7

Coset leader

00000	10110	01011	11101
00001	10111	01010	11100
00010	10100	01001	11111
00100	10010	01111	$\boxed{11001}$
01000	11110	00011	10101
10000	00110	11011	01101
~~00110~~	~~10000~~	~~01101~~	~~11011~~
11000	01110	10011	00101
01100	11010	00111	10001

Code is a coset (00000)

$= \{c+U\}$

$U \in \{00100, 10010, 01111, 11001\}$

4 ways to recieve 11001

$\{(00000+11001)(10110+01111)\}$

→ NOT A NEW COSET

$\{01011 + 10010, 11101 + 00100\}$

After constructing the standard array, decoding is as follows.

I.) Recieve w,

II.) Find w in the standard array.

III.) [illegible]

II.) Find w in the standard array.

III.) Find coset leader l_w of w.

IV.) Coset leader contains positions of errors
ie. decode as $w + l_w$.

Syndrome decoding

DUAL CODE $C^\perp$ of C

Scalar product

$$\vec{x} \cdot \vec{y} = (x_1 \cdot y_1 + x_2 y_2 + \dots + x_n \cdot y_n) \bmod 2$$

more generally $(\bmod\ q)$ q prime

$\vec{x}$ and $\vec{y}$ are perpendicular if $\vec{x} \cdot \vec{y} = 0$

$$C^\perp = \{ w \in \{0,1\}^n \mid w \cdot c = 0, \forall c \in C \}$$

if C is of dimension k then $C^\perp$ is of dimension $n-k$

$$G = I_k \mid A$$

$$H = (-A^T \mid I_{n-k})$$ is a generator matrix for $C^\perp$

A

$$G = \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right) \qquad \forall c \in C$$

$$c \cdot H^T = \overbrace{0\,0\,0 \ldots 0}^{n-k}$$

$$H = \left(\begin{array}{ccc|cc} 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{array}\right)$$

scalar product of $c \in C$ and $c' \in C^{\perp}$

$$\overset{\in G}{(0\,1\,0\,0\,1)} \underset{A^T}{\begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}} = \begin{pmatrix} \underset{\displaystyle 0}{\underset{\|}{0\cdot 0 + 1\cdot 0 + 0\cdot 0 + 0\cdot 1 + 1\cdot 0}} & \underset{\displaystyle 0}{1} \end{pmatrix}$$

---

for an error vector e.

$$(c+e) \cdot H^T = (c \cdot H^T + e \cdot H^T) = 0 + \underbrace{eH^T}_{\text{Syndrome}}$$

there is a one to one correspondence between syndromes and cosets of C.

Ex 2.8

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix} = \left(\begin{array}{cccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \end{array}\right)$$

Find G

$$= \left(\begin{array}{cccc|ccc} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right) = (A^T \mid I)$$

$$G = (I \mid A) = \left(\begin{array}{cccc|ccc} 1&0&0&0&0&1&1\\ 0&1&0&0&1&1&1\\ 0&0&1&0&1&1&0\\ 0&0&0&1&1&0&1 \end{array}\right) \qquad d=3$$

e	$e \cdot H^T$
0000000	000
0000001	100
0000010	001
0000100	011
⋮	⋮
1000000	101

→ all cosets!

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

$(0000001) \cdot$

Each coset contains $16 = 2^4$ words

8 cosets $\cdot$ 16 $= 2^3 \cdot 2^4 = 2^7 =$ all words

receive (1111110)

$1111110 \cdot H^T = (100)$

decoded as 1111111.

Hamming codes

$(7,4,3)$-Hamming code

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

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

0000000	000
1000000	001 → 1
⋮	

$$\begin{vmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \leftarrow \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix} \qquad \begin{vmatrix} \vdots \\ 1000100 \end{vmatrix} \; 101 \to 5$$