Tutorial II

16 October 2020 09:22

LINEAR CODES

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

I. Important parameters (n, M, d)

- n: length of codewords
- M: number of codewords
- d: minimum distance

to calculate d, you need to calculate the Hamming distance of every pair of codewords $O(|C|^2)$.

II ENCODING

$m_0 \longrightarrow C_0$

$\vdots$

$m_M \longrightarrow C_M$

III DECODING

Upon receiving w, you need to find $c \in C$ with minimum Hamming distance $\min_{c \in C} H(c, w)$

LINEAR CODES: DEFINITION

A code C over alphabet q (q is a power of prime) is a linear code if two conditions hold:

I. $\forall x, y \in C \quad x+y \in C$

$x = (x_0, \dots, x_{n-1})$

$y = (y_0, \dots, y_{n-1})$

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

II. $\forall k \in \{0, \dots, q-1\}, \forall c \quad k \cdot c = (k \cdot c_0, k c_1, \dots, k c_{n-1}) \in C$

$(+, \cdot)$ are both operations $\mathbb{F}_q$ (mod q, if q is a prime)

Ex 2.1

$C = \{00000, 00110, 10010, 11001, 00101, 10100, 00011, 10111\}$

10100

10100

00110

II. trivial for binary codes

I.

3 dimensional subspace $(\{0,1\}^5, +, \cdot)$

Linear code C is a subspace $(\{0,1\}^n, +, \cdot) \rightsquigarrow$ n-dimensional space of 2^n vectors

What is the dimension of C?

k-dimensional subspace has 2^k vectors (q^k for alphabet of size q)

For linear codes instead of (n, M, d) we often write (n, k, d)

k: size of code space $\Leftrightarrow 2^k$ codewords

C can be characterized by k linearly independent codewords, called a basis $= \{b_0, \dots, b_{k-1}\} \subseteq C$

$$\forall c \in C \quad c = a_0 b_0 + a_1 b_1 + \dots + a_{k-1} \cdot b_{k-1}$$

$$a_i \in \{0, \dots, q-1\}$$

$$C = \{\underline{00000}, 00111, 10010, 10001, 00101, 10100, 00011, 10111\}$$

$b_0 = 00110$
$b_1 = 10010$
$b_2 = 10001$

10100 (from $b_0 + b_1$), 00011 (from $b_1 + b_2$), 10111 (from $b_0 + b_2$), 00101 (from $b_0 + b_1 + b_2$)

$0 \cdot b_0 + 0 \cdot b_1 + 0 \cdot b_2 = (00000)$ is always a codeword in a linear code

---

ADVANTAGE I How to calculate the minimum distance d of a linear code?

$$H(c_1, c_2) = H(c_1 + w, c_2 + w)$$

$$\| $$

$$H(c_1 + c_1, c_2 + c_1)$$

$$\| $$

$$H(0, c_2 + c_1)$$

$$\| $$

$$H(0, c) \quad c \in C$$

$\Rightarrow$ To find d, it suffices to find the codeword of the smallest weight (number of non-zero entries)

$$\Uparrow$$

$$O(|C|)$$

ENCODING

$k \times n$ matrix

Generating matrix $G = \begin{bmatrix} b_0 \\ \vdots \\ b_k \end{bmatrix}$ of code C

Since $M = 2^k$ we can associate each message with $m_i \in \{0,1\}^k$

To encode message m calculate

$$c = m \cdot G$$

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

d

$$G = \begin{pmatrix} 00110 \\ 10010 \\ 10001 \end{pmatrix}$$

$$c_1 = (001) \cdot G$$

$m_1 = 001$

$m_5 = 101$

$= (0\cdot 0 + 0\cdot 1 + 1\cdot 1 \mid 0001)$

$= (10001)$

$c_5 = (101) \cdot G = (10111)$

WE DO NOT NEED ENCODING TABLE. STORING G is enough.

NORMAL FORM OF G

$G = (\underline{\Pi}_k \mid A)$

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

$\underline{\Pi}_k$: $k \times k$ identity matrix

Algorithm to find normal form of G:

1.) Start with arbitrary G

2.) Do following operations until normal form is found

a.) permutation of rows

b.) multiplication of rows by non-zero scalar → change the basis

c.) addition of rows

d.) multiplication of columns by non-zero scalar

c.) permutation of columns

$\Big\}$ ↦ change linear code to an equivalent linear code

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

$$\simeq \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right) = \left(\begin{array}{c|c} \mathbb{I}_3 & A \end{array}\right) = G'$$

Codes with G in normal form are called **systemic**.

$$(abc) \cdot G' = \underbrace{abc}_{\text{output (message)}} \underbrace{0\ (a+b+c)}_{\text{checksum}}$$

DECODING

Standard array decoding

Coset $u = \{u + c \mid c \in C\}$

$\forall u, v \in \{0,1\}^n$ coset u and coset v are either identical or disjoint

Coset leader is a vector of a coset with the smallest weight

Ex 2.7 $C = \{00000, 10110, 01011, 11101\} = (5, 2, 3)$ code

u	C+u				
00000	00000	10110	01011	11101	← the code
00001	00001	10111	01010	11100	
00010	00010	10100	01001	11111	
00100	00100	10010	01111	11001	24 words
01000	01000	11110	00011	10101	
10000	10000	00116	11011	01101	
00011	00011	10101	01000	11110	
11000	11000	01110	10011	00101	8 words
01100	01100	11013	00111	10001	

Decoding procedure:

1.) receive w

2.) find w in standard array

3.) identify coset leader (l_w) of w

4.) decode as $w + l_w$

Example:

$w = 11110 \Rightarrow$ coset leader $l_w = 01000$

$\Rightarrow$ decode as 10110

SYNDROME DECODING

Dual Code $C^{\perp}$ of C

Scalar product (dot product)

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

$\vec{x}$ and $\vec{y}$ are perpendicular

if $\vec{x} \cdot \vec{y} = 0$

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

if dimension of C is k

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

$G = (I_k \mid A)$ → generator matrix of C

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

EXAMPLE $(I_3 \mid A)$

$$G = \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right) \quad A^T = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

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

$(-A^T \mid I_2)$

$\forall c \in C$

$c \cdot H^T = \overbrace{000}^{n-k}$

$(01001) \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \in C \quad \downarrow \in C^\perp & \in C \quad \downarrow \in C^\perp \\ \left[(01001)\cdot(00010)\right] & \left[(01001)\cdot(11101)\right] \\ 0 & 0 \end{pmatrix}$

error in the channel is characterized by an error vector $e \in \{0,1\}^n$

$w = (c + e)$

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

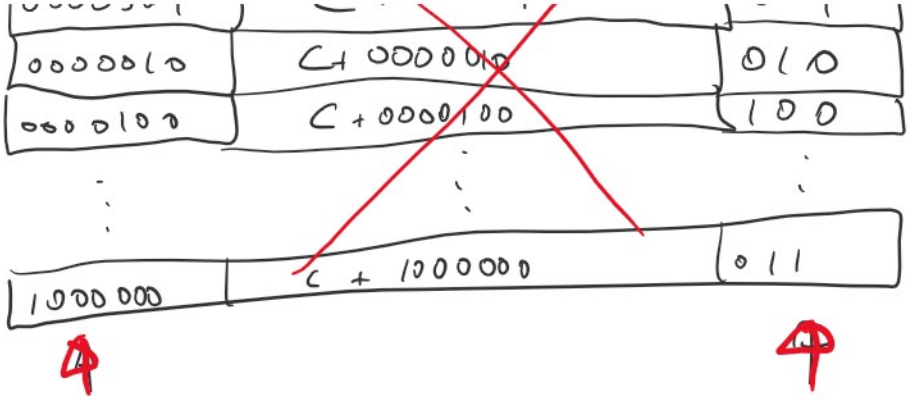
THERE is a one to one correspondence between syndromes and cosets [cosets are $\{c+e\}$]

EX. 28

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

e	$c+e$	$e \cdot H^T$
0000000	c	000
0000001	$c + 0000001$	001
0000010	$c + 0000010$	010

$H_1^T = \begin{pmatrix} 011 \leftarrow \\ 101 \\ 110 \\ 111 \\ 100 \leftarrow \\ 010 \leftarrow \\ 001 \leftarrow \end{pmatrix}$



$|\{0,1\}^7| = 2^7$

$|K| = 2^4$

Standard array contains $8 \cdot 2^4 = 2^7$

1.) receive w

2.) Calculate the syndrome $w \cdot H^T = e \cdot H^T$

3.) find $e \cdot H^T$ in (reduced) standard array

4.) find corresponding error e

5.) decode as $w+e$

Hamming codes

$$H^T = \begin{pmatrix} 001 \\ 010 \\ 011 \\ 100 \\ 101 \\ 110 \\ 111 \end{pmatrix}$$

$e = 0010000$

$$e \cdot H^T = 011$$

'3' in binary

Hamming codes → Syndrome is a binary of the error position.