Elements of Algebraic Structures DATE: 20-th April,2024

Topic: Application of Coset in Error Correction

Lecture: Sujata Ghosh *Presented by: Soumik Guha Roy*

Contents

1 Introduction

1.1 Definition

- **Message:** The basic unit of information called a **message** which is finite sequence of characters from a finite alphabet.Here ,our alphabet set is $B = 0, 1$.

- **Word:** Basic unit of information,called **word**, is a sequence of m 0's or 1's.

- **Weight:** It is defined as the number of 1's in a code word. the weight of the code word 001 is $|001| = 2$.

- **Distance:** It is defined as the number of differing positions among two same length code word w_1, w_2.It is denoted by $\delta(w_1, w_2) = |w_1 \oplus w_2|$.Consider two words 1010 and 1011. The distance is $\delta(1010, 1011) = |1010 \oplus 1011| = |0001| = 1$

- The set $B = \{0, 1\}$ is forming a **group** under the operation '+' defined as:

+	0	1
0	0	1
1	1	0

- $B^m = B \times B \times B \cdots \times B$ is group under the operation $\oplus$ defined as $(x_1, x_2, \ldots, x_m) \oplus (y_1, y_2, \ldots, y_m) = (x_1 + y_1, x_2 + y_2, \ldots, x_m + y_m)$ and the **identity element** of $B^m = (0, 0, \ldots, 0) = \bar{0}$

- Elements of B^m will be written as $b_1, b_1, \ldots, b_m$

- Sender sends $x \in B^m$ and the receiver receives $x_t \in B^m$. Due to **Noise** $x \neq x_t$.Hence,if noise exists then x_t can be any element in B^m.

2 Group Code

2.1 Definition

- $(B^n, \oplus)$ is a group

- An (m,n) Encoding function $e : B^m \mapsto B^n$ is called **Group Code** if $e(b^m) = \{e(b) | b \in B^m\} = Range(e)$ is a subgroup of $(B^n, \oplus)$

Example of Group code

- Given an encoding function $e : B^2 \mapsto B^3$ where the *odd parity of zero* encoding is used.The function is shown below:

B^2	B^3
00	000
01	011
10	101
11	110

Hence, elements of $B^3 = N = \{000, 011, 101, 110\}$.Lets check the composition table of $(N, \oplus)$.

$\oplus$	000	011	101	110
000	000	011	101	110
011	011	000	110	101
101	101	110	110	011
110	110	101	011	000

Here,$(N, \oplus)$ forming a group and $(N, \oplus)$ is subgroup of $(B^3, \oplus)$.So,$(N, \oplus)$ is a Group code.

- Consider the following encoding function:

B^2	B^3
00	001
01	010
10	101
11	111

Hence, elements of $B^3 = N = \{001, 010, 101, 111\}$ and $(B^3, \oplus)$ not forming a group.Hence $(B^3, \oplus)$ not a group code.

2.2 Generating Group Code

- **Minimum distance of a group code**
  Theorem 1: Given $e : B^m \mapsto B^n$ is a group code. The ***minimum distance of e*** is the minimum weight of the nonzero code word.

- **Theorem 2:** Let $\mathbf{D}$ and $\mathbf{E}$ be $m \times p$ Boolean matrices, and let $\mathbf{F}$ be a $p \times n$ Boolean matrix. Then $(\mathbf{D} \oplus \mathbf{E}) \star \mathbf{F} = (\mathbf{D} \star \mathbf{F}) \oplus (\mathbf{E} \star \mathbf{F})$

- **Theorem** Let $m, n \in \mathbb{N}$ with $m < n$ and $r = n - m$ and let $\mathbf{H}$ be an $n \times r$ Boolean matrix. Then the function $f_H : B^n \mapsto B^r$ defined by $f_H(x) = x \star \mathbf{H}$ where $x \in B^n$ is a homomorphism from the group B^n to B^r.

Proof: Let $x, y \in B^n$. Then

$$\begin{aligned} f_H(x \oplus y) &= (x \oplus y) \star H \\ &= (x \star H) \oplus (y \star H)[Using Theorem 2] \\ &= f_H(x) \oplus f_H(y) \end{aligned}$$

Hence, f_H is a homomorphism from B^m to B^n.

- **Corollary 3.1:** $N = \{x | x \in B^n, x \star \mathbf{H} = 0\}$ is a normal subgroup of $(B^n, \oplus)$.
  Proof: N is the kernel of the homomorphism f_H,so it is a normal subgroup of $(B^n, \oplus)$.

3 Parity Check Matrix and Encoding function

- **Parity check matrix** An $n \times r$ Boolean matrix defined as $\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1r} \\ h_{21} & h_{22} & \dots & h_{2r} \\ \\ \dots \\ h_{m1} & h_{m2} & \dots & h_{mr} \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}$
  where the last r rows is a identity matrix I_r.

- **Encoding function using parity check matrix:** $e_H : B^m \mapsto B^n$.Let $b = b_1 b_2 \dots b_m$ and $x = e_H(b) = b_1 b_2 \dots b_m x_1 x_2 \dots x_r$ where

$$\begin{aligned} x_1 &= b_1.h_{11} + b_2.h_{21} + \dots + b_m.h_{m1} \\ x_2 &= b_1.h_{12} + b_2.h_{22} + \dots + b_m.h_{m2} \\ \dots \\ x_r &= b_1.h_{1r} + b_2.h_{2r} + \dots + b_m.h_{mr} \end{aligned} \tag{1}$$

- **Theorem** Let $x = y_1y_2 \dots y_m x_1 x_2 \dots x_r \in B^n$. Then $x \star H = 0 \leftrightarrow x = e_H(b), b \in B^m$
  - **Corollary** $e_H(B^m) = \{e_H(b) | b \in B^m\}$ is a sub group of B^n

Example of Encoding

- Given the group code $e_H : B^2 \mapsto B^5$ then $m = 2, n = 5$ and the parity check matrix $\mathbf{H} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- $B^2 = \{00, 01, 10, 11\}$.
- Then $e(00) = 00x_1x_2x_3$ where $b_1 = b_2 = 0$ and x_1, x_2, x_3 can be obtained from the **H** matrix.$x_1 = x_2 = x_3 = 0$.Hence, $e(00) = 00000$
- Similarly,$e(10) = 10x_1x_2x_3$ where $b_1 = 1, b_2 = 0$ and $x_1 = x_2 = 1, x_3 = 0$.Hence,$e(01) = 10110$
- In the same way, $e(01) = 01x_1x_2x_3$ where $b_1 = 0, b_2 = 1$ and $x_1 = 0, x_2 = x_3 = 1$.Hence,$e(01) = 01011$
- Finally, $e(11) = 11x_1x_2x_3$ where $b_1 = 1, b_2 = 1$ and $x_1 = 1, x_2 = 0, x_3 = 1$.Hence,$e(01) = 11101$
- Minimum distance of this group code (2,5) is 3 (**Why this distance?**)

4 Decoding

4.1 Maximum likelihood decoding function:

Given $e_H : B^m \mapsto B^n$.Let us list the code words in a fixed order: $x^{(1)}, x^{(2)}, \dots, x^{(2^m)}$.Let x_t be the received word and compute $\delta(x^{(i)}, x_t) \forall i = 1$ to 2^m and choose the first code word , $x^{(s)}$ such that min $\delta(x^{(i)}, x_t) = \delta(x^{(s)}, x_t)$ $\forall i = 1$ to 2^m. Hence,$x^{(s)}$ is the closest code to x_t and the first in the list $x^{(1)}, x^{(2)}, \dots, x^{(2^m)}$.Let $x^{(s)} = e(b)$.Then **maximum likelihood decoding function d** associated with e by $d(x_t) = b$ where x_t is the received word.The maximum likelihood decoding function d depends on the order

$x^{(1)}, x^{(2)}, \dots, x^{(2^m)}$.

Theorem: Given that e is an (m, n) encoding function and d is the maximum likelihood decoding function associated with e. Then (e,d) can correct k or fewer errors if and only if the minimum distance of e is at least $2k + 1$.

4.2 Coset leader

Let $e : B^m \mapsto B^n$ be an (m, n) encoding function.N is set of code words in B^n such that $N = \{x^{(1)}, x^{(2)}, \dots, x^{(2^m)}\}$ Let $x = e(b)$ where $b \in B^n$ is transmitted and received as $x_t \in B^n$. Left coset of N is $x_t + N = \{x_t + x^{(1)}, x_t + x^{(2)}, \dots, x_t + x^{(2^{(m)})}\} = \{\epsilon_1, \epsilon_2, \dots, \epsilon_{2^{(m)}}\}$ where $\epsilon_{2^{(i)}} = x_t \oplus x^{(i)}$. Distance between the received code word x_t and $x^{(i)}$ is $|\epsilon_i|$ ϵ_j is a coset member with smallest weight, then $x^{(j)}$ must be the code word that is closest to x_t.Here, $x^j = \bar{0} \oplus x^j = x_t \oplus x_t \oplus x^j = x_t \oplus \epsilon_j$

Coset Leader An element ϵ_j having the smallest weight,called the *Coset leader.Coset leader may not be unique*

Example of coset leader

Given the encoding function $e : B^2 \mapsto B^3$, which is odd parity of zero function.Hence, $B^2 = \{00, 01, 10, 11\}$ and $N = B^3 = \{000, 011, 101, 110\}$ and $B^3 - N = \{001, 010, 100, 111\}$.

$000 = \epsilon_1$	$011 = \epsilon_1 \oplus x^{(2)}$	$101 = \epsilon_1 \oplus x^{(3)}$	$110 = \epsilon_1 \oplus x^{(4)}$
$001 = \epsilon_1$	$010 = \epsilon_2 \oplus x^{(2)}$	$100 = \epsilon_2 \oplus x^{(3)}$	$111 = \epsilon_2 \oplus x^{(4)}$

Here, in the $row1$,$\epsilon_1 = 000$ is the coset leader.In the $row2$, the coset leader is $\epsilon_2 = 001$.Multiple possible coset leader exists in the 2nd row.But only one coset leader exists in the row 1.Coset leaders can also be used to generate each row.

4.3 Construction of decoding table using coset leader

Let $N = \{x^{(1)}, x^{(2)}, \dots, x^{(2^m)}\}$, where $x^{(1)}$ is $\bar{0}$, the identity element of B^n.Now if $x^{(1)} \in N$, then $x^{(1)} + N = \{x^{(1)}, x^{(2)}, \dots, x^{(2^m)}\}$.

Now take $y \in B^n - N$, then $y \oplus N = \{y \oplus x^{(1)}, y \oplus x^{(2)}, \dots, y \oplus x^{(2^m)}, \}$.In the left coset $y \oplus N$ pick an element of least weight,a coset leader,which can be denoted by ϵ_2.In case of tie,pick any of the element of least weight.Now as $\epsilon_2 \in y \oplus N$, we can say that $\epsilon_2 \oplus N = y \oplus N$.This implies that every word in the coset of $y \oplus N$ can also be written as $\epsilon_2 \oplus v$ where $v \in N$.We can express the coset $y \oplus N =$

Table 1: Decoding Table

$\bar{0} = \epsilon_1$	$x^{(2)}$	$x^{(3)}$	$\ldots$	$x^{(2^m-1)}$
ϵ_2	$\epsilon_2 \oplus x^{(2)}$	$\epsilon_2 \oplus x^{(3)}$	$\ldots$	$\epsilon_2 \oplus x^{(2^m-1)}$
ϵ_3	$\epsilon_3 \oplus x^{(2)}$	$\epsilon_3 \oplus x^{(3)}$	$\ldots$	$\epsilon_3 \oplus x^{(2^m-1)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\epsilon_{2^{n-m}}$	$\epsilon_{2^{n-m}} \oplus x^{(2^{n-m})}$	$\epsilon_{2^{n-m}} \oplus x^{(3)}$	$\ldots$	$\epsilon_{2^{n-m}} \oplus x^{(2^m-1)}$

$\{\epsilon_2 \oplus x^{(1)}, \epsilon_2 \oplus x^{(2)}, \epsilon_2 \oplus x^{(3)}, \ldots, \epsilon_2 \oplus x^{(2^m)}\} = \{\epsilon_2, \epsilon_2 \oplus x^{(2)}, \epsilon_2 \oplus x^{(3)}, \ldots, \epsilon_2 \oplus x^{(2^m)}\}$.
Let $z \in B^n - N$ and $z \neq y$.The left coset of N $z \oplus N = \{z \oplus x^{(1)}, z \oplus x^{(2)}, \ldots, z \oplus x^{(2^m)}, \} = \{\epsilon_3, \epsilon_3 \oplus x^{(2)}, \epsilon_3 \oplus x^{(3)}, \ldots, \epsilon_3 \oplus x^{(2^m)}\}$, where ϵ_3 is the element in $z \oplus N$ with smallest weight.Hence, the coset leader for $z \oplus N$ is ϵ_3.
Continue this process until all elements of B^n have been listed. The result is listed in the following table,called the **decoding table**.The decoding table contains $2^{n-m} = 2^r$ rows one for each coset of N and 2^m columns, that is total 2^n elements.
If we receive the word x_t, we locate it in the decoding table. If $x \in N$ and x is at the top of the table of the row of x_t, then x is the code word which closest to x_t.Thus if $x = e(b)$, we let $d(x_t) = b$.

Example of decoding table construction

Given the encoding function $e : B^2 \mapsto B^3$, which is odd parity of zero function.Hence, $B^2 = \{00, 01, 10, 11\}$ and $N = B^3 = \{000, 011, 101, 110\}$ and $B^3 - N = \{001, 010, 100, 111\}$.

000=ϵ_1	011=$\epsilon_1 \oplus x^{(2)}$	101=$\epsilon_1 \oplus x^{(3)}$	110=$\epsilon_1 \oplus x^{(4)}$
001=ϵ_1	010=$\epsilon_2 \oplus x^{(2)}$	100=$\epsilon_2 \oplus x^{(3)}$	111=$\epsilon_2 \oplus x^{(4)}$

5 Basic Decoding function

- Given $e : B^m \mapsto B^n$ is a group code and sender sends the data b encoded as $x = e(b)$ to the receiver.
- **Step 1:** Determine all the left cosets of $N = e(B^m)$
- **Step 2:** For each coset, find the coset header(a word with smallest weight)
- **Step 3:** Determine in which coset of N, x_t belongs.[As N is normal subgroup of B^n ,due to partition of N, x_t will be in exactly one coset among 2^{n-m}]

- **Step 4:** Let ϵ be the coset leader as determined in Step 3.Compute $x = x_t \oplus \epsilon$. If $e(b) = x$, then $d(x_t) = b$.Hence, receiver decodes x_t as b.

- The **main problem of this algorithm** is the calculation of the entire table containing all the coset elements.

Example of decoding using the decoding table

Given the encoding function $e : B^2 \mapsto B^3$, which is odd parity of zero function.Hence, $B^2 = \{00, 01, 10, 11\}$ and $N = B^3 = \{000, 011, 101, 110\}$ and $B^3 - N = \{001, 010, 100, 111\}$. The encoding function is

b	00	01	10	11
$e(b)$	000	011	101	110

000=ϵ_1	011=$\epsilon_1 \oplus x^{(2)}$	101=$\epsilon_1 \oplus x^{(3)}$	110=$\epsilon_1 \oplus x^{(4)}$
001=ϵ_1	010=$\epsilon_2 \oplus x^{(2)}$	100=$\epsilon_2 \oplus x^{(3)}$	111=$\epsilon_2 \oplus x^{(4)}$

- Consider the received word is $x_t = 011$.The receiver will search the decoder table for the code word 011 and find the code word in the 1st row.As it is in the first row,the receiver conclude that the original sent code word was $x = 011$ and $d(011) = 01 = b$.

- Consider the received word is $x_t = 111$.The receiver will search the decoder table for the code word 111 and find the code word in the last row and last column.Then 1st element of the last column will be the actual data $x = 110$ and $d(110) = 11 = b$.Here,1-bit error is corrected by the receiver.

- Suppose the sender sends the data 00 encoded as 000 and the receiver receives the data 110.The the receiver will conclude that the original code word was 11 which is not the actual one.Hence, if the bit error in this example is 2, the the method fails.

6 Syndrome of a code word

- **Theorem:** Given $m, n, r = n - m$ and $f_H : B^n \mapsto B^m$ and defined as $f_H(x) = x \star H$, then f_H is onto function.
  Proof: Let $b = b_1 b_2 \ldots b_r \in B^r$. Letting $x = 00\ldots0 b_1 b_2 \ldots b_r$, we obtain $x \star H = b$. Thus $f_H(x) = b$,so f_H is onto.

- **Syndrome of x:** B^r and B^n/N are isomorphic where $N = ker(f_H) = e_H(B^m)$ under the homomorphism $g : B^n/N \mapsto B^r$ defined by $g(xN) = f_H(x) = x \star H$.Here , the element $x \star H$ called the *Syndrome of x*.

- **Theorem:** Let $x, y \in B^n$.Then x and y are same left coset of N in B^n if and only if $f_H(x) = f_H(y)$, that is if and only if x and y have the same syndrome.
  Proof: We know that given H is normal sub group of G if $a \star H = b \star H \implies a^{-1} \star b \in H$.Hence ,x,y lies in same left coset of N, if and only if $x \oplus y = (-x) \oplus y \in N$. Since, $N = ker(f_H), x \oplus y \in N$ if and only if

$$f_H(x \oplus y) = \bar{0}_{B^r}$$
$$f_H(x) \oplus f_H(y) = \bar{0}_{B^r}$$
$$f_H(x) = f_H(x)$$

7 Modified Decoding function

- Given $e : B^m \mapsto B^n$ is a group code and sender sends the data b encoded as $x = e(b)$ to the receiver.

- **Step 1:** Determine all the left cosets of $N = e_H(B^m)$ in B^n.

- **Step 2:** For each coset find the coset leader and find the syndrome of all coset leaders.

- **Step 3:** If x_t is the received, computer the syndrome of x and find the coset leader ϵ having the same syndrome. Then $x_t \oplus \epsilon = x$ is a code word $e_H(b)$ and $d(x_t) = b$.

- Here, we donot need to keep the entire table of cosets.

8 Example

Given the $(3,6)$ group $e_H : B^3 \mapsto B^6$ and consider the parity matrix $\mathbf{H} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

The encoding function

$e(x)$	$e(000)$	$e(001)$	$e(010)$	$e(011)$	$e(100)$	$e(101)$	$e(110)$	$e(111)$
$x \in N$	000000	001011	010101	011110	100110	101101	110011	111000

The Syndrome Coset Leader table

Syndrome of Coset Leader ($x \star \mathbf{H}$)	000	001	010	011	100	101	110	111
Coset Leader (ϵ_i)	000000	000001	000010	001000	00100	010000	100000	001100

8.1 Without noise

- Sender send the data $b = 011 \in B^3$ encoded as $x = e(011) = 011110$
- Receiver receives the data $x_t = 011110$
  - Step 1: Calculate the syndrome of x_t as $f_H(x_t) = x_t \star H = 011110 \star H = 101$
  - Step 2: Using the Coset Leader table, the coset header is $\epsilon = 010000$
  - Step 3: Finally ,compute $x = x_t \oplus \epsilon = 001110 \oplus 010000 = 011110$ and the data is $b = e^{-1}(011110) = 011$

8.2 With noise

- Sender send the data $b = 001 \in B^3$ encoded as $x = e(001) = 001011$
- Receiver receives the data $x_t = 011011$
  - Step 1: Calculate the syndrome of x_t as $f_H(x_t) = x_t \star H = 011011 \star H = 101$
  - Step 2: Using the Coset Leader table, the coset header is $\epsilon = 010000$
  - Step 3: Finally ,compute $x = x_t \oplus \epsilon = 011011 \oplus 010000 = 001011$ and the data is $b = e^{-1}(001011) = 001$

1 bit error message corrected by Receiver

8.3 With noise but erroneous acceptance

- Sender send the data $b = 010 \in B^3$ encoded as $x = e(010) = 010101$
- Receiver receives the data $x_t = 011111 \in B^6$
  - Step 1: Calculate the syndrome of x_t as $f_H(x_t) = x_t \star H = 011111 \star H = 001$
  - Step 2: Using the Coset Leader table, the coset header is $\epsilon = 000001$
  - Step 3: Finally ,compute $x = x_t \oplus \epsilon = 001110 \oplus 010000 = 011110$ and the data is $b = e^{-1}(011110) = 011$.

Wrong data accepted by Receiver.But why?

Here. the minimun distance of the (m,n) encoding function e is atleast $2k + 1 = 3$ and d is the maximum likelihood decoding function associated with e $\implies$ the (e,d) can correct atmost $k = 1$ errors.Hence , 2 errors cannot be detected by the receiver.

References

[1] Bernard Kolman,Robert C.Busby,Sharon Ross *Discrete Mathematical Structures*, Prentics Hall of India.

[2] S.K.Mapa *Higher Algebra*, Levant Books,India, 4th ed.