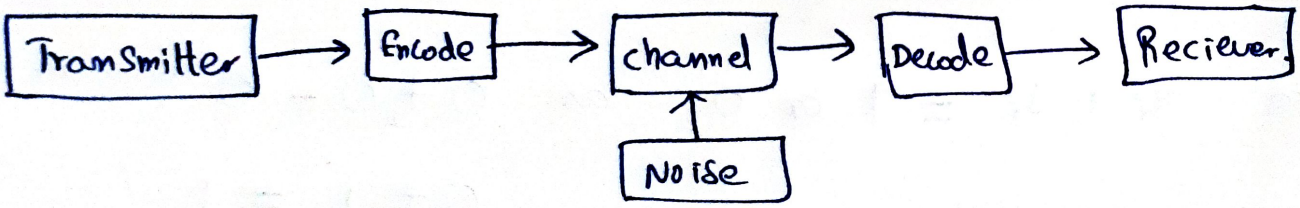
Coding Theory

The Process of Communication involves transmitting some information Carrying Signal (message) that is Conveyed by a Sender to a reciever.

Coding theory deals with minimizing the distortions of the Conveyed message due to noise and to retrieve the original message to the optimal extent Possible from the Corrupted message.

Encoder: An encoder is a device which transforms the incoming message in such a way that the presence of noise in the transformed messages is detectable.

Decoder: A decoder is a device which transforms the encoded message into their original form that can be understood by the reciever.



* We will use only a binary channel in which the encoder will transform an input message into a binary string consisting of the symbols 0 and 1.

* Decoding is only the inverse operation of encoding.

GROUP CODE:

$B = \{0, 1\}$, then $B^n = \{ x_1, x_2, \ldots, x_n \mid x_i \in B \}$ is a group under the binary operation of addition modulo 2, denoted by $\oplus$. This group $(B^n, \oplus)$ is called a group code.

Proof If $x_1, x_2, \ldots, x_n \equiv (x_1, x_2, x_3, \ldots, x_n)$

and $y_1, y_2, \ldots, y_n \equiv (y_1, y_2, y_3, \ldots, y_n) \in B^n$

Then $x_1, x_2, x_3, \ldots, x_n \oplus y_1, y_2, \ldots, y_n$

$= (x_1 +_2 y_1, x_2 +_2 y_2, \ldots, x_n +_2 y_n) \in B^n$

Since $x_i +_2 y_i = 1$ or 0 as $0 +_2 0 = 0$

$0 +_2 1 = 1$

$1 +_2 0 = 1, 1 +_2 1 = 0.$

$(0,0,0,\ldots,0)$ is the identity element of B^n.

The inverse of $x_1, x_2, \ldots, x_n$ is itself

$[\because (x_1, \ldots, x_n) + (x_1, \ldots, x_n) = (0,0,\ldots,0)]$

$\therefore$ $(B^n, \oplus)$ is a group – It is also a abelian group.

* In general, any code which is a group under the operation $\oplus$ is called a group code.

HAMMING CODE:

The codes obtained by introducing additional digits called parity digits to the digits in the original message is a binary string of length 'm', the Hamming encoded message is a string of length 'n' $(n>m)$ of the n digits, m digits are used to represent the information part of the message and the remaining $(n-m)$ digits are used for the detection and correction of errors in the recieved message.

HAMMING SINGLE ERROR DETECTING CODE:

It is a code of length 'n', the first (n-1) digits contain the information part of the message and the last digit is made either 0 or 1.

ODD Parity / even Parity

If the digit introduced in the last (Portion) position gives an even number of count 1's / odd number of count 1's in the encoded word of length n, then extra digit is called an even/odd Parity check.

Examples:

words	000	001	011
even parity	↓ 0000	↓ 0011	↓ 0110

words	000	001	011
odd parity	↓ 0001	↓ 0010	↓ 0111

Weight of a code:

The number of $1^{'s}$ in the binary string $x \in B^n$ is called the weight of code 'x'. and denoted by $|x|$

$$x = 110101101 \Rightarrow |x| = 6.$$

Hamming distance $H(x,y)$

If x and y represents the binary strings $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ the number of positions in the strings for which $x_i \neq y_i$ is called the Hamming distance between x and y denoted by $H(x,y)$

$$\therefore H(x,y) = \text{weight of } x \oplus y = \sum_{i=1}^{n} (x_i +_2 y_i)$$

Example: If $x = 11010$ $y = 10101$

$$H(x,y) = |x \oplus y| = |01111| = 4.$$

Minimum distance:

The minimum distance of a code is the minimum of the Hamming distances between all pairs of encoded words in that code.

Example

If $x = 10110$, $y = 11110$, $z = 10011$

$H(x,y) = |x \oplus y| = |01000| = 1$

$H(y,z) = |y \oplus z| = |01101| = 3$

$H(z,x) = |z \oplus x| = |00101| = 2$

So the minimum distance between the code words is $\boxed{1}$, is the Hamming distances.

THEOREM

A code (an (m,n) encoding fn) can detect at most 'k' errors if and only if the minimum distance between any two code words is at least $(2k+1)$.

Encoding.

$e : B^m \rightarrow B^n$, where $m, n \in Z^+$, $n > m$

encoding function: e is given by a matrix $(m \times n)$ G over B.

ERROR CORRECTION USING MATRICES.

The matrix G is called the generator matrix for the code

$$G = [I_m \mid A]$$

I_m is the identity matrix / (or) Unit matrix of order $m \times m$.

I_m – order $m \times m$

A – order $(m) \times (n-m)$.

If w is a message (orginal message) $\in B^n$.

Then $e(w) = wG$ [encoding]

The code $C = e(B^m) \subseteq B^n$, where w is a $(1 \times m)$ vector.

$B = \{0, 1\}$

$B^2 = \{00, 01, 10, 11\}$

$B^3 = \{000, 001, 010, 100, 011, 101, 110, 111\}$

DECODING

For Decoding we need Parity check matrix H.

$$H = [\, A^T \mid I_{n-m} \,]$$

H is a $(n-m) \times n$ matrix

G is a $m \times n$ matrix

I_m is a $m \times m$ Unit matrix

A is a $m \times (n-m)$ matrix Suitably choosen.

r is a $1 \times n$ matrix.

r – is the recieved message.

* This Unique Parity check matrix H provides a decoding scheme that Corrects a Single error in transmission as explained below.

Case (i): If 'r' is a recieved word considered as $(1 \times n)$ matrix and if

$H \cdot r^T = [0]$, then we Conclude that there is no error in transmission and that r is the code word transmitted.

Case (ii) If $H \cdot r^T =$ i^{th} column of H, then we Conclude that a Single error has occuring during transmission. and it has occured in the i^{th} Component of r, changing the i^{th} Component of r, we get the code word 'c' transmitted.

* The first 'm' Components of 'C' gives the original message.

* **Note:** $H \cdot r^T =$ the first column of H, a single error has occured in the first Component of 'r'.

Case (iii) If neither Case (i) nor Case (ii) occurs then, we conclude that more than one transmission error have occured. Though detection of error is possible in this case, correction is not possible.

Theorem : 2

A code can correct a set of at most k errors if and only if minimum distance between any two code words is atleast $(2k+1)$.

Problem: 1 Find the code words for $w \in B^2$

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

Given $w \in B^2$ $\quad B^2 = \{00, 01, 10, 11\}$

W.K.T $\quad e(w) = w \cdot G$

$$e(00) = [0\ 0] \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} = [0\ 0\ 0\ 0\ 0]$$

$$e(01) = [0\ 1] \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} = [0\ 1\ 0\ 1\ 1]$$

$$e(10) = [1\ 0] \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} = [1\ 0\ 1\ 1\ 0]$$

$$e(11) = [1\ 1] \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} = [1\ 1\ 1\ 0\ 1]$$

(2)

Find the code words generated by the encoding function $e: B^2 \to B^5$ with respect to the Parity check matrix

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

W.K.T $e: B^m \to B^n$ $\quad m=2,\ n=5,\ n-m=3.$

$$G = [I_m \mid A] = [I_2 \mid A]$$

$$H = [A^T \mid I_{n-m}] = [A^T \mid I_3]$$

Here in the problem H^T is given we have to rewrite

$$H = \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{array}\right]_{3\times 5} = [A^T \mid I_3]$$

$$A^T = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \qquad G = [I_2 \mid A]$$

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

$B^2 \equiv \{00, 01, 10, 11\}$ w.k.T $e(\omega) = \omega . G$

$$e(00) = [0\ 0] \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} = [0\ 0\ 0\ 0\ 0]$$

$$e(01) = [0\ 1] \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} = [01\ 01\ 1]$$

$$e(10) = [1\ 0] \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} = [1\ 0\ 0\ 11]$$

$$e(11) = [1\ 1] \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} = [1\ 1\ 000]$$

Hence the Code words generated by H are

00000, 01011, 10011, 11000.

Problem: 3

Find the Code words generated by the Parity check matrix $H = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ When the encoding function is $e: B^3 \to B^6$.

$e: B^m \to B^n$, $m=3$, $n=6$, $n-m=6-3=3$.

H is not in the Correct form. (Taking transpose)

$$H = \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] = [A^T \mid I_{n-m}] = [A^T \mid I_3]$$

$$G = [I_m \mid A] = [I_3 \mid A] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right]$$

$B^3 = \{000, 001, 010, 100, 011, 101, 110, 111\}$

$e(000) = [000]G = [000\ 000]$

$e(001) = [001]G = [001\ 011]$

$e(010) = [010]G = [010\ 101]$

$e(011) = [011]G = [011\ 110]$

$e(100) = [100]G = [100\ 111]$

$e(101) = [101]G = [101100]$

$e(110) = [110]G = [110010]$

$e(111) = [111]G = [111001]$

∴ The generated codes words are

000000, 001011, 010101, 100111, 011110

100100, 110010. and 111001.

Problem: 4 Given the generator matrix

$$G \equiv \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$ Corresponding to the encoding function $e: B^3 \to B^6$, find the corresponding Parity check matrix and use it to decode the following received words and hence, to find the original message. Are all words decoded uniquely?

(i) 110101 (ii) 001111 (iii) 110001

(iv) 111111

given $e : B^3 \to B^6$ $\quad m=3,\ n=6,\ n-m=3.$

$$G = \left[\begin{array}{ccc|ccc} 1&0&0&1&1&0\\ 0&1&0&0&1&1\\ 0&0&1&1&0&1 \end{array}\right] = [I_m \mid A] = [I_3 \mid A]$$

$$H = [A^T \mid I_{n-3m}] = [A^T \mid I_3] = \left[\begin{array}{ccc|ccc} 1&0&1&1&0&0\\ 1&1&0&0&1&0\\ 0&1&1&0&0&1 \end{array}\right]$$

— x — x —

we Compute the Syndrome of each of the recieved word by using $H \cdot [r]^T$

(i) $H \cdot [r]^T = \left[\begin{array}{ccc|ccc} 1&0&1&1&0&0\\ 1&1&0&0&1&0\\ 0&1&1&0&0&1 \end{array}\right] \cdot \begin{bmatrix} 1\\1\\0\\1\\0\\1 \end{bmatrix} = \begin{bmatrix} 0\\0\\0 \end{bmatrix}$

$r = [1\,1\,0\,1\,0\,1]$

$r \to$ recieved message.

Since $H \cdot [r]^T = H \cdot [e(\omega)]^T = \begin{bmatrix} 0\\0\\0 \end{bmatrix}$

The recieved word in this Case is the transmitted word itself

Hence the original message is 110

[* [Since Size of m is 3] take first 3 letters]

(ii) $r = [0\ 0\ 1\ 1\ 1\ 1]$

$$H \cdot [r]^T = \begin{matrix} & 1 & 2 & 3 & 4 & 5 & 6 \\ \end{matrix} \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Since the Syndrome $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ is Same as the 5th Column of H,

∴ The element in the fifth position of 'r' is changed.

∴ The decoded word is 0 0 1 1 0 1 and the original message is $\boxed{001}$ [Since m=3]

(iii) $r = 11000$

$$H \cdot [r]^T = \begin{matrix} & 1 & 2 & 3 & 4 & 5 & 6 \\ \end{matrix} \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Since the Syndrome $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is the Same as the 4th Column of H,

The element in the 4th position of 'r' is changed.

The decoded word 110101 and the orginal message is $\boxed{110}$

(iv) $r = [1\ 1\ 1\ 1\ 1\ 1]$

$$H \cdot [r]^T = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Since the Syndrome $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is not identical to any column of H,

The recieved word cannot be decoded uniquely.

Problem: 5 Construct the decoding table for the group coder given by the generator matrix

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

Decode the following words recieved using the decoding table obtained. Which of the words could not be decoded uniquely?

101111, 011010, 101110, 111111

Given $G = [\ ]_{3\times 6} = [I_3 \mid A]$

$m=3$, $n=6$ since $e: B^3 \to B^6$.

$B^3 = \{000, 001, 010, 100, 011, 101, 110, 111\}$

$e(\omega) = \omega G$

$e(000) = [000]G = [000000]$

$e(001) = [001]G = [001011]$

$e(010) = [010]G = [010101]$

$e(100) = [100]G = [100111]$

$e(011) = [011]G = [011110]$

$e(101) = [101]G = [101100]$

$e(110) = [110]G = [110010]$

$e(111) = [111]G = [111001]$

Code words →							
000000	001011	010101	100111	011110	101100	110010	111001
100000	101011	110101	000111	111110	001100	010010	011001
010000	011011	000101	110111	001110	111100	000010	110001
001000	001011	011101	101111	010110	100100	111010	111101
000100	001111	010001	100011	011010	101000	110110	111011
000010	001001	010111	100101	011100	101110	110000	111011
000001	001010	010100	100110	011111	101101	110011	111000
011000	010011	001101	111111	000110	110100	101010	100001

↑ Coset leaders

For 8th row we can choose '110000', But it is already in 7th row 7th column.

So we choose '011000'.

The decoding table is not unique Since for 8th row we can choose

100001
000110

Recall Theorem

A Code Can Correct a set of at the most k ~~often~~ errors if and only if the minimum distance between any two code words is at least $(2k+1)$.

Theorem

In a **group code**, the minimum distance between distinct code words is the minimum weight of the non zero code words in it.

Example: 101111, 011010, 101110, 111111

minimum weight : 3

from above theorem: $2k+1 = 3$

$2k = 2$ $\quad k = 1$

$\therefore$ At most one error Can be corrected.

[that is zero or one error can be Corrected].

DECODING OF THE RECIEVED WORDS:-

101111 , 011010, 101110, 111111

we note that minimum distance between the group codes is minimum weight of the non-zero code words. That is $\boxed{3}$ Here.

$\therefore$ At most 1 error can be Corrected

$2K+1 = 3$

$2K = 2$

$\boxed{K=1}$

---

1) 101111

101111 appears in 4^{th} row and 4^{th} column.

The coset leader of the 4^{th} row is 001 000, which contain only one 1.

The corrected (recieved) word, the code word transmitted is the top element of the 4^{th} column.

It is 100 111 and hence the original message is $\boxed{100}$

(ii) 011010

011010 appears in 5th row and 5th column

Hence the corresponding code word transmitted is **011110** and hence the original message is **011**

(iii) 101110

101110 appears in 6th row and 6th column

Hence the corresponding code word transmitted is **101100** and hence the original message is **101**

(iv) 111111

111111 appears in 8th row,

The coset leader of which contains two 1's.

The recieved word has **2** errors.

Hence, it cannot be corrected and the code word transmitted cannot be **Uniquely determined**