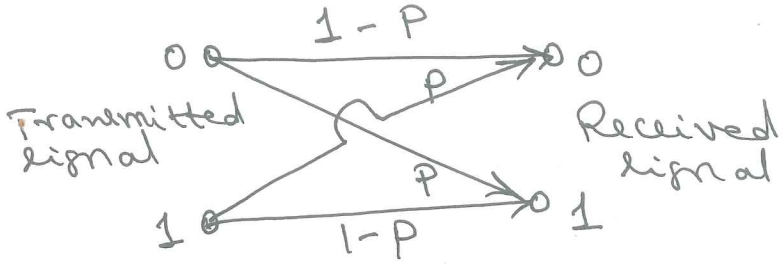
Chapter-10 : Coding Theory.

Introduction :- The algebraic coding theory was inspired by the fundamental paper of Claude Shannon (1948) along with results by Marcel Golay (1949) and Richard Hamming (950). Since that time it has become an area of great intrest where algebraic structures, probability, and combinatorics all play a role.

In this chapter, our coverage will be held to an introductory level as we seek to model the transmission of information represented by strings of the signals 0 and 1.

In digital communications, when information is transmitted in the form of strings of 0's and 1's, certain problems arise. As a result of "noise" in the channel, when a certain signal is transmitted a different signal may be received, thus calling the receiver to make wrong decision. Hence, we want to develop techniques to help us detect and perhaps even correct transmission errors. Only we can improve the chances of correct transmission there are no guarantees.

Our model uses a <u>binary symmetric channel</u>, as shown below.



The adjective binary appears because an individual signal is represented by one of the bits 0 or 1. When a transmitter sends the signal 0 or 1 in such a channel, associated with either signal is a probability p for incorrect transmission. When that probability p is same for both signals, the channel is called symmetric. Here we have probability p of sending 0 and having 1 received. The probability of sending signal 0 and having it received correctly is then $1-p$.

Ex: Consider a string $c = 10110$.
$c \in Z_2^5$ group, formed from the direct product of five copies of $(Z_2, +)$.

When sending each bit of c through the binary symmetric channel, we assume that the probability of incorrect transmission is $p = 0.05$, so that the probability of transmitting c with no errors is $(0.95)^5 = 0.77$.

NOTE: The transmission of each signal does not depend in any way on the transmissions of prior signals. Consequently, the probability of the occurrence of all of these independent events is given by the product of their individual probabilities.

What is the probability that the party receiving the five-bit message receives the string $r = 00110$?

⇒ Here the original message is $c = 10110$, therefore the error occur at first position.
The probability is 0.05 for the first bit and 0.95 for the remaining 4 bits. Therefor, $(0.05)(0.95)^4 = 0.041$ is the probability of sending $c = 10110$ and receiving $r = 00110$. By considering $e = 10000$, we can write $c + e = r$. That is, r is the sum of the original message and error pattern $e = 10000$. Here, $c + r = e$ or $r + e = c$, since, $e, c, r \in Z_2^5$ and

In transmitting $C = 10110$, the probability of receiving $r = 00100$ is $(0.05)(0.95)^2(0.05)(0.95) = 0.002$

hence, this occurance of this multiple error is less likely to happen.

What is the probability that r differs from C in exactly two places?

TWO 1's & 3 zeros

$$\binom{5}{2}(0.05)^2(0.95)^3 = 0.021$$

From this we can conclude the following.

" Let $c \in Z_2^n$, for the transmission of C through a binary symmetric channel with probability p of incorrect transmissions, (a) the probability of receiving $r = c + e$, where e is a particular error pattern consisting of k 1's and $(n-k)$ zero is $p^k(1-p)^{n-k}$ (b) the probability that k errors are made in the transmission is

$$\binom{n}{k} p^k (1-p)^{n-k}$$

The probability of making, at most one error in the transmission of $C = 10110$ is $(0.95)^5 + \binom{5}{1}(0.05)^1(0.95)^4 = 0.977$.

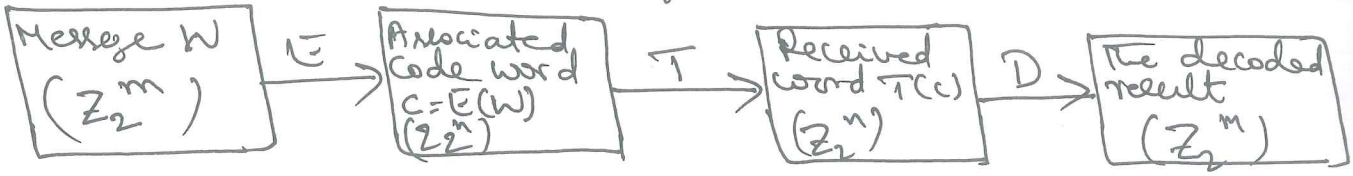
Thus the chance of multiple errors in transmission will be considered negligible. This assumption is valid when p is very small. In a binary symmetric channel is considered "good" when $p < 10^{-5}$, or $p < 1/2$

To improve the accuracy of transmission in a binary symmetric channel, certain types of coding schemes can used where extra signals are provided.

For $m, n \in \mathbb{Z}^+$, let $n > m$. Consider $\phi \neq W \subseteq Z_2^m$. The let W consists of the messages to be transmitted. For each $w \in W$, append $n-m$ extra bits to form the codeword c, where $c \in Z_2^n$. This process is called <u>encoding</u>. This is represented by a function $E: W \to Z_2^n$. Therefore $E(w) = c$ and $E(W) = C \subseteq Z_2^n$. Since the function E simply appends extra bits to the distinct messages, the encoding is <u>one-to-one</u>.

After transmission of c, is received as $T(c)$, where $T(c) \in Z_2^n$. Unfortunately, T is not ~~one-to-one~~ a function because $T(c)$ may be different at different transmission times (noise changes with times).

Upon receiving $T(c)$, we want to apply a decoding function $D: Z_2^n \to Z_2^m$ to remove the extra signals and to obtain the original message w. Ideally $D \circ T \circ E$ should be the identity function on W, with $D: C \to W$. Since this cannot be expected, we seek functions E & D such that there is a high probability of correctly decoding the received word $T(c)$ and recapturing the original message w.



The functions E & D are called encoding & decoding functions of an <u>(n, m) block code</u>.

Ex: Consider the $(m+1, m)$ block code for $m = 8$. Let $W = Z_2^8$. For each $w = w_1 w_2 \cdots w_8 \in W$, define $Z_2^8 \to Z_2^9$ by $E(w) = w_1 w_2 \cdots w_9$, where $w_9 = \sum_{i=1}^{8} w_i$ where the addition is modulo 2. For example, $E(11001101) = 110011011$, & $E(00110011) = 001100110$ For all $w \in Z_2^8$, $E(w)$ contains an even number of 1's. So for $w = 11010110$ and $E(w) = 110101101$, if we receive $T(c) = T(E(w))$ as 100101101, this has odd no. of 1's, hence we know that mistake happens in transmission. Hence, we are able to detect single errors in transmission. But, we don't know how to correct such errors.

The probability of sending the code word 110101101 and allowing at most one error in transmission is

$$(1-p)^9 + \binom{9}{1} p(1-p)^8$$

All 9 bits are transmitted Correctly

One bit changed in transmission and it is detected.

For $p = 0.001$, this gives $(0.999)^9 + \binom{9}{1}(0.001)(0.999)^8 = 0.9999964$

If we detect an error and we are able to relay a signal back to the transmitter to repeat the transmission of the codeword, and continue this process until the received word has an even number of 1's then the probability of sending the code word 110101101 and receiving the correct transmission is approximately 0.9999964.

This $(m+1, m)$ block code is called parity-check code. What is the drawback of this code?

Ex: The $(3m, m)$ block code [Triple Repetition Code] is one in which we can detect and correct single errors in transmission.

With $m=8$ and $W = Z_2^8$, we define

$E: Z_2^8 \to Z_2^{24}$ by $E(w_1 w_2 \cdots w_8) = w_1 w_2 \cdots w_8 w_1 w_2 \cdots w_8 w_1 w_2 \cdots w_8$

The decoding function $D: Z_2^{24} \to Z_2^8$ is carried out by the majority rule.

For example, if the $w = 10110111$ then

$c = E(w) = 10110111 : 10110111 : 10110111$

If the received word $T(c)' = 10100111 : 00110111 : 10110110$

In this we have 3 errors at position 4, 9 & 24.

Decode this $T(c)$ by examining $1^{st}, 9^{th}, 17^{th}$ bits and check which signal appears more times. Here it is 1, so the first bit is 1 in the decoded message. Similarly counting in this pattern for $2^{nd}, 10^{th}$ & 18^{th} positions, here all 3 are zero hence the 2^{nd} bit in decoded word is 0. By continuing in this pattern we decoded $T(c)$ back to $w = 10110111$.

What is the Draw back in this?

Even though, many errors are detected and corrected, but if many or 2-errors occur at the same position it can't be corrected / or incorrect transmission occurs.

What is the probability? $p = 0.001$.

$(0.999)^3 + \binom{3}{1}(0.001)(0.999)^2 = 0.999997$

$\therefore (0.999997)^8 = 0.999976$

Hamming Metric:

Consider a code $C \leq Z_2^4$, where $c_1 = 0111$ & $c_2 = 1111 \in C$. If both sender and receiver know the elements of C. If the transmitter sends c_1 and the person receiving the code word received $T(c_1)$ as 1111, then he or she feels that c_2 was transmitted and makes the appropriate decision (wrong) c_2 implied. Here, the two code words they are almost similar, they are close to each other, for they differ in only one component.

Definition: For each element $x = x_1 x_2 \cdots x_n \in Z_2^n$, where $n \in Z^+$, the **weight** of x, denoted as $wt(x)$, is the number of components x_i of x, for $1 \leq i \leq n$, where $x_i = 1$. If $y \in Z_2^n$, the **distance** between x & y, denoted by $d(x,y)$, is the number of components where $x_i \neq y_i$, for $1 \leq i \leq n$.

Ex: For $n=5$, let $x = 01001$ and $y = 11101$,
Then $wt(x) = 2$, $wt(y) = 4$ and $d(x,y) = 2$.

$x+y = 10100$, so $wt(x+y) = 2$, In this example by chance $d(x,y) = wt(x+y)$.

For each $1 \leq i \leq 5$, $x_i + y_i$ contributes a count of 1 to $wt(x+y) \iff x_i \neq y_i$, whereas x_i, y_i contribute a count of 1 to $d(x,y)$.

∴ This is actually true for all $n \in Z^+$, so $wt(x+y) = d(x,y)$

Lemma: For all $x, y \in Z_2^n$, $\boxed{wt(x+y) \leq wt(x) + wt(y)}$ //

$x = x_1 \cdots x_n$
$y = y_1 \cdots y_n$
$x+y = z_1 z_2 \cdots z_n$

The distance function d, defined on $Z_2^n \times Z_2^n$ satisfies the following for all $x, y, z \in Z_2^n$

a] $d(x,y) \geq 0$

b] $d(x,y) = 0$ iff $x = y$

c] $d(x,y) = d(y,x)$

d] $d(x,z) \leq d(x,y) + d(y,z)$

Any function which satisfies all the above properties is called a distance function or metric, and we call (Z_2^n, d) a metric space. Here d is often refered as Hamming metric:

Definition: {Sphere} :- For $n, k \in Z^+$ and $x \in Z_2^n$, the sphere of radius k centered at x is defined as

$$S(x,k) = \{ y \in Z_2^n \mid d(x,y) \leq k \}$$

Ex: For $n = 3$ and $x = 110 \in Z_2^3$

$\therefore S(x,1) = \{110, 010, 100, 111\}$

$S(x,2) = \{110, 010, 100, 111, 000, 101, 011\}$

What is the error detecting and correcting capabilities of Hamming metric?

Let $E: W \to C$ be a encoding function with the set of messages $W \subseteq Z_2^m$ and the set of code words $E(W) = C \subseteq Z_2^n$, where $m < n$. For $k \in Z^+$, we can detect transmission errors of weight $\leq k$ iff the minimum distance between code words is at least $k+1$.

proof :- The set C is known to both transmitter and the receiver, so if $w \in W$ is the message and $c = E(w)$ is transmitted, let $c \neq T(c) = r$. If the minimum distance between the code words is at least $k+1$, then the transmission of c can result in at most k errors and r will not be listed in C. Hence, we can detect all errors e where $wt(e) \leq k$.

(or) If c_1 and c_2 are the codewords with $d(c_1, c_2) < k+1$. Then $c_2 = c_1 + e$ where $wt(e) \leq k$. If we send c_1 & $T(c_1) = c_2$, then we feel that c_2 is sent, thus failing to detect an error of weight $\leq k$.

With E, W, and C are as above, and $k \in Z^+$, we can construct a decoding function $D: Z_2^n \to W$ that corrects all transmission errors of weight $\leq k$ iff minimum distance between code words is at least $2k+1$.

proof: For $c \in C$, consider $S(c,k) = \{x \in Z_2^n \mid d(c,x) \leq k\}$.

Define $D: Z_2^n \to W$ as follows. If $r \in Z_2^n$ and $r \in S(c,k)$ for some code word c, then $D(r) = w$ where $E(w) = c$. If $r \notin S(c,k)$ for any $c \in C$, then we define $D(r) = w_0$, where w_0 is some arbitrary

The only problem we would face here is that D might not be a function. This will happen if there is an element r of Z_2^n with r in both $S(c_1, k)$ and $S(c_2, k)$ for distinct code words c_1, c_2.

But $r \in S(c_1, k) \Rightarrow d(r, c_1) \leq k$, and $r \in S(c_2, k)$ $\Rightarrow d(r, c_2) \leq k$, so $d(c_1, c_2) \leq d(c_1, r) + d(r, c_2) \leq k + k \leq 2k+1$. Consequently, if the minimum distance between code words is at least $2k+1$, then D is a function, and it will decode all possible received words, correcting any transmission error of weight $\leq k$. Conversely, $c_1, c_2 \in C$ & $d(c_1, c_2) \leq 2k$, then c_2 can be obtained from c_1 by making atmost $2k$ changes. Starting at code word c_1 we make approximately half of these changes. This brings us to $r = c_1 + e_1$ with $wt(e_1) \leq k$.

Continuing from r, we make the remaining changes to get to c_2 and find $r + e_2 = c_2$ with $wt(e_2) \leq k$. But then $r = c_2 + e_2$. Now with $c_1 + e_1 = r$ $= e_2 + c_2$ and $wt(e_1), wt(e_2) \leq k$, how can one decide on the code word from which r arises? This ambiguity arises in a possible error of weight $\leq k$ that cannot be corrected.

Ex: With $W = Z_2^2$ let $E: W \to Z_2^6$ given by

$E(00) = 000000$, $E(10) = 101010$, $E(01) = 010101$, $E(11) = 111111$.

$\Rightarrow$ The minimum distance between the code words is 3.

Hence, we can detect double errors & correct single ones.

$$S(000000, 1) = \{x \in Z_2^6 \mid d(000000, x) \leq 1\}$$

$$= \{000000, 100000, 010000, 001000, 000100, 000010, 000001\}.$$

The decoding function $D: Z_2^6 \to W$ given

$D(x) = 00$ for all $x \in S(000000, 1)$

Now

$$S(010101, 1) = \{x \in Z_2^6 \mid d(010101, x) \leq 1\}$$

$$= \{010101, 110101, 011101, 010001, 010111, 010100\}$$

or 101

here $D(x) = 01$ for each $x \in S(010101, 1)$

At this point our definition of D accounts for 14 of the elements in Z_2^6. Continuing to define D for the 14 elements in $S(101010, 1)$ & $S(111111, 1)$, there remain 36 other elements to account for, We define $D(w) = 00$ for these 36 elements and have a decoding fn that will correct single errors.

If $c = 010101$ & $T(c) = r = 1\,1\,1\,1\,01$, we can detect this double error because r is not a code word. But if $T(c) = r_1 = 111111$, 3 errors, we think that $c = 111111$, & it decodes to 11, (incorrectly) instead it is sent as 01.

Two bit errors can't be corrected.

Parity check and Generator matrix :-

Here the encoding and decoding functions are given by matrices over Z_2. One of these matrices will help us to locate the nearest code word for a received word. This is helpful when the set C grows larger.

Ex: Let $G_g = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{array}\right]$ be a 3×6 matrix over Z_2.

$G_g = [I_3 \mid A]$

The first 3 columns of G_g constitute Identity matrix. The remaining 3 columns represents matrix A.

This partitioned matrix G_g is called generator matrix. Using this generator matrix G_g, we can define the encoding function $E: Z_2^3 \to Z_2^6$ as follows.

For $w \in Z_2^3$, $E(w) = wG_g$.

NOTE: Assume that all of Z_2^3 is encoded and that the transmitter and the receiver will both know the real messages of importance and their corresponding code words.

For example,

$$E(110) = 110\,G_g = [1\ 1\ 0]\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix} = [1\ 1\ 0\ 1\ 0\ 1]$$

& $E(010) = 010\,G_g = [0\ 1\ 0]\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix} = [0\ 1\ 0\ 0\ 1\ 1]$

Note that $E(110)$ is obtained by adding first two rows of G_g. $E(010)$ is simply the second row of G_g.

$W = \{000, 100, 010, 001, 110, 101, 011, 111\} \subseteq Z_2^3$ and their code words are

$C = \{000000, 100110, 010011, 001101, 110101, 101011, 011110, 111000\} \subseteq Z_2^6$

After receiving a codeword, one can recover the message by simply dropping the last 3 bits. The mapping

For all $w = w_1 w_2 w_3 \in Z_2^3$,

$E(w) = w_1 w_2 w_3 w_4 w_5 w_6 \in Z_2^6$.

$$E(w) = [w_1 w_2 w_3] \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

$$= [w_1 w_2 w_3 (w_1+w_3)(w_1+w_2)(w_2+w_3)]$$

so we have $w_4 = w_1 + w_3$, $w_5 = w_1 + w_2$, $w_6 = w_2 + w_3$ and these equations are called parity-check equations.

Since $w_i \in Z_2$ for each $1 \le i \le 6$, it follows that $w_i = -w_i$, and so these parity check equations are rewritten as

$$w_1 + w_3 + w_4 = 0$$
$$w_1 + w_2 + w_5 = 0$$
$$w_2 + w_3 + w_6 = 0$$

Thus we find that

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \end{bmatrix} = H \cdot (E(w))^T = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

where $(E(w))^T$ denotes the transpose of $E(w)$.

If the received word is $r = r_1 r_2 \cdots r_6 \in Z_2^6$, we can identify r as a code word iff

$$H \cdot r^T = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$H = [B \mid I_3]$ Note that $B = A^T$

In this example, since the minimum distance is 3, we are able to define a decoding function which corrects a single errors.

Suppose we receive $r = 110110$. We want to find the code word c which is nearest to r. Just compare r with each code allowable code words. If $|C|$ is very high, just examine $H \cdot r^T$, this is called as <u>syndrome</u> of r. (18)

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

so r is not a codeword, hence the error is detected.

Now compute the distance between r & each codewords $c \in C$, and consider the one which has less distance.

$\therefore d(100110, r) = 1$ & all other $c \in C$, $d(r, c) \geq 2$.

Writing $r = c + e = (100110 + 010000)$, we find that the transmission error of (weight 1) occurs in the second component of r. Hence, changing the second bit of r, we will get c, the message is comprises the first three components of c.

Let $r = c + e$, where c is a code word and e is an error pattern of weight 1. Suppose that 1 is the ith component of e, where $1 \leq i \leq 6$. Then

$$H \cdot r^T = H(c+e)^T = H \cdot (c^T + e^T) = H \cdot c^T + H \cdot e^T$$

With c the code word, it follows that $H \cdot c^T = 0$, so

$H \cdot r^T = H \cdot e^T = i^{th}$ column of the matrix H. Hence c & r differ only in the ith component, and we can determine c by simply changing the ith component of r.

Suppose, $r = 000111$, its syndrome,

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

If $c = \overset{0}{\cancel{1}}00\,1\cancel{1}\overset{1}{0}$: 000111 = r