Applications of Linear Algebra to Coding Theory

Presented by:-
Prof.Surjeet kaur
Dept of Mathematics
SIES College.Sion(W)

Outline

- Introduction
- AIM
- Coding Theory Vs Cryptography
- Coding Theory
- Binary Symmetric Channel
- Hamming Code
- Generator and Parity check matrices
- Applications

Coding Theory

- It is concerned with reliability of communication over noisy channels.
- Number of applications in digital communication such as E-mail, internet and Intranet.
- Also used in store scanners(Bar code)
- International Standard Book Number (ISBN)

Coding Theory Vs Cryptography

- Coding theory deals with communication in a hostile channel
- Concerned with encoding and decoding messages
- Need for clearing the information sent

- Cryptography is about disguising messages so only certain people can see through the disguise
- Concerned with encrypting and decrypting
- Hidden communication

General Idea

- The main method used to recover messages that might be distorted during transmission over a noisy channel is to employ redundancy.
- Error detecting codes:-

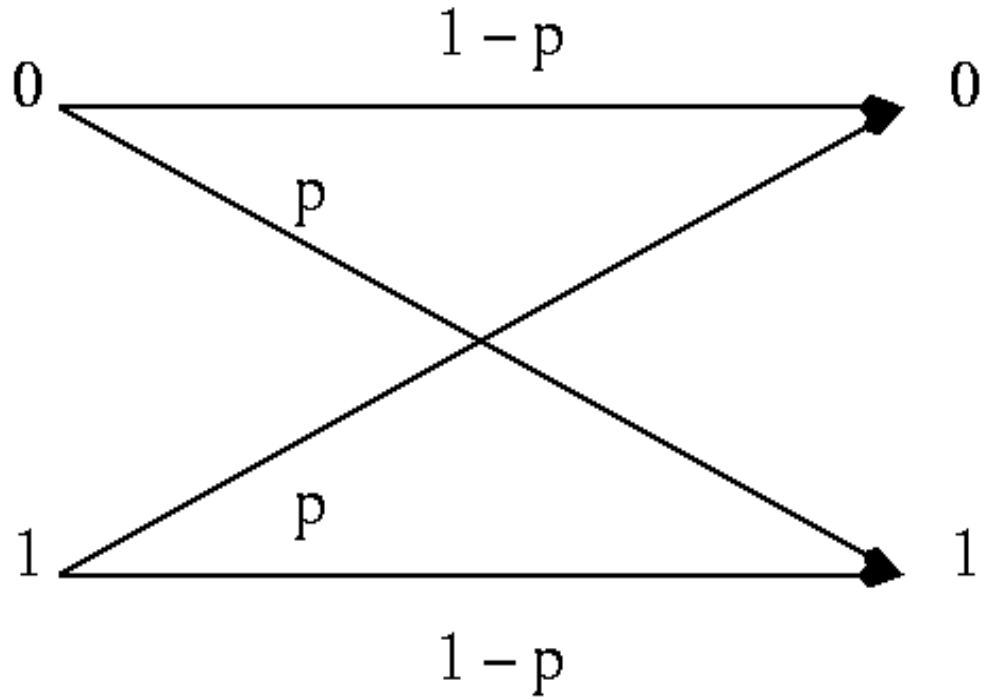
  Detect when an error occurs in transmission
- Error Correcting codes:-

  Detect and correct the errors in transmission

Simple Repetition Code

- Mathematical use of redundancy.
- Binary Symmetric Channel (BSC):-

  In this channel, every bit of a transmitted message has the same probability p of being changed to the other bit.
- 1-p is the reliability of the channel.
- In block coding theory, original data is broken into blocks of a fixed length and certain amount of redundancy is added to the data.



Binary Symmetric channel (BSC) is idealised model used for noisy channel.

- binary (0,1)
- symmetric p(0→1) =p(1→0)

Hamming Codes

- 3 bits of redundancy are added to information bits.

If the original data bits be denoted as $x_1x_2x_3x_4$ then the corresponding codeword is $x_1x_2x_3x_4x_5x_6x_7$,obtained by adding 3 redundancy bits according to the equations:-$x_5=x_1+x_2+x_4$;$x_6=x_1+x_3+x_4$;$x_7=x_2+x_3+x_4$,

where all computations are done modulo 2.

History on Hamming codes

- Middle of 20th century by Richard Hamming, Marcel Golay
- Bell Labs
- Early computers were detecting errors and halting, hence wasting a lot of computations.
- Single error-correcting codes in mid 1940s

Vector space And codes

- A nonempty set of elements called vectors on which two operations, namely addition and scalar multiplication have been defined such that V is closed with respect to these operations and satisfies certain axioms.
- In Coding theory field B={0,1} of scalars with operations of addition and multiplication defined as:-

  0+0=0; 0+1=1; 1+0=1; 1+1=0

  0.0=0; 0.1=0; 1.0=0; 1.1=1

Defn:- A binary linear code of length n is a vector subspace of B_n .

Hamming C(7,4)

- Consider V_7 vector space of 7 tuples of 0's and 1's over the field of scalars{0,1} where addition and multiplication are defined in the usual component wise manner
- For eg:-(1,0,0,1,1,0,1)+(0,1,1,1,0,0,1)=(1,1,1,0,1,0,0)
- 0(1,0,0,1,1,0,1)=(0,0,0,0,0,0,0) and 1(1,0,0,1,1,0,1)=(1,0,0,1,1,0,1)
- Since each vector in V_7 has seven components, and each of these components can be either 0 or 1,there are 2^7 vectors in this space.
- The four dimensional subspace of V_7 having basis

  B={(1,0,0,0,0,1,1),(0,1,0,0,1,0,1),(0,0,1,0,1,1,0),(0,0,0,1,1,1,1)} is called a Hamming Code and is denoted as $C_{7,4}$
- The vectors in $C_{7,4}$ can be used to send messages.
- Each vector in $C_{7,4}$ can be written as

  vi =a1(1,0,0,0,0,1,1)+a2(0,1,0,0,1,0,1)+a3(0,0,1,0,1,1,0)+a4(0,0,0,1,1,1,1)

  There are 2^4 = 16 vectors in $C_{7,4}$.. The Hamming code $C_{7,4}$ can thus be used to send 16 different messages v1,v2,v3……..v16.

Hamming Codes and Error Correction

- When an error occurs in one location of a transmitted message the resulting incorrect vector lies in V_7,outside the subspace $C_{7,4.}$
- It can be proved that there is exactly one vector in $C_{7,4}$ that differs from this incorrect vector in one location. Thus the error can be detected and corrected.
- In practice, electrical circuits called gates are used to test whether the received message is in $C_{7,4}$ or not.

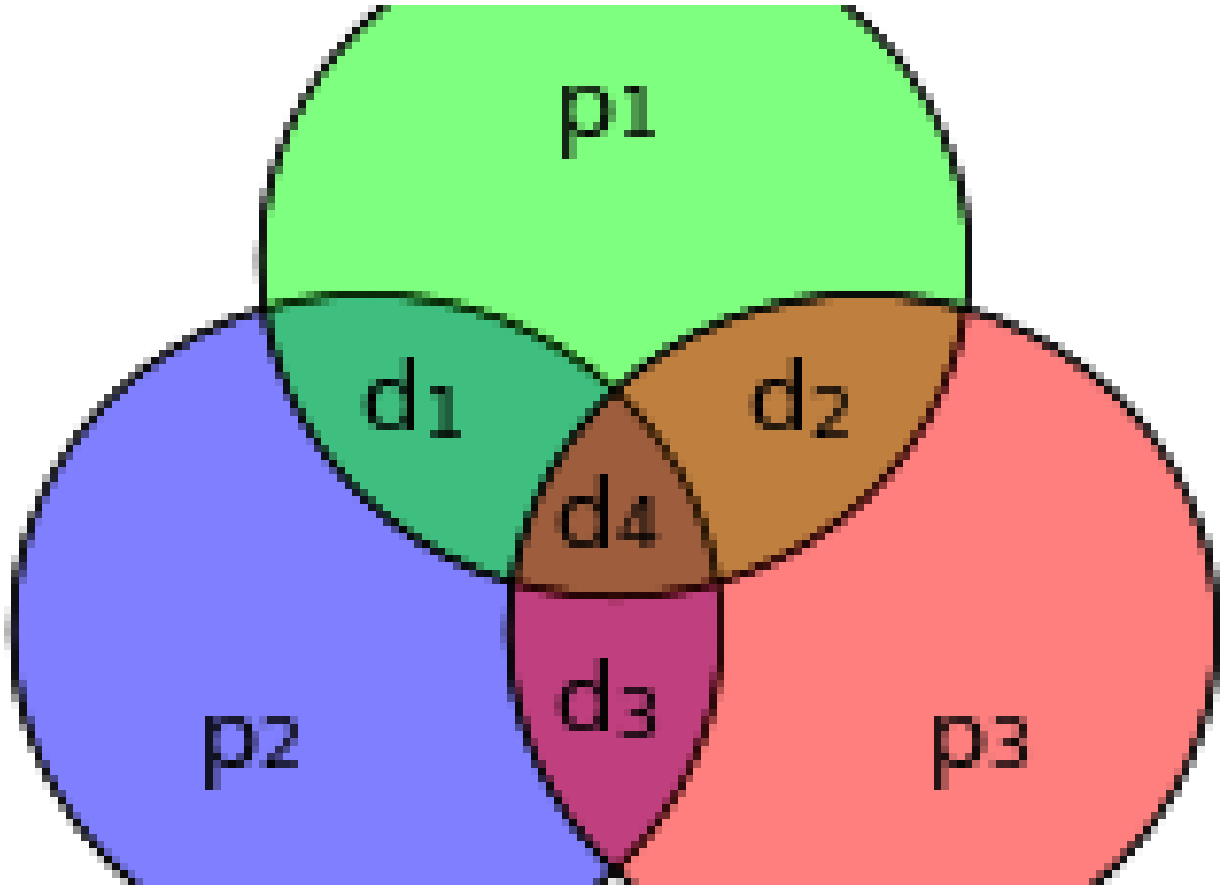
Generator and Parity check matrices

- A generator matrix of a linear code C is a matrix G whose rows span the code.
- A parity check matrix H of a linear code is a matrix whose null space is C.
- Augment binary messages with an extra bit to make an even no of 1's.
- If you receive a message with an odd no of bits you know there has been an error in transmission.

- Therefore, the code (of dimension k) can be defined as either C = { u*G : u in B_k } or C = { u in B_n : H*u = 0 vector}. The rank of G or the nullity of H give the dimension of C.
- Recall from linear algebra that a k by n matrix over B defines a linear transformation from B_k to B_n . So, the vector-matrix multiplication u*G corresponds to encoding: The information vector u of length k is transformed into a codeword v = u*G of length n.
- The redundancy is added through the vector-matrix multiplication. The parity check matrix is useful for checking for errors. Suppose the code has a parity check matrix H. If a vector w is received, we compute the product H*w, called the syndrome of w. If the syndrome is the zero vector, we assume that there was no error. If not, we know that there is an error.

Hamming Codes (Contd)

- The goal of Hamming codes is to create a set of parity bits that overlap such that a single-bit error (the bit is logically flipped in value) in a data bit *or* a parity bit can be detected *and* corrected.



Graphical representation of data bits d1,d2,d3,d4(corresponding to $x_1x_2x_3x_4$) and parity bits p1,p2,p3(corresponding to $x_5x_6x_7$)

Hamming Matrices

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

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

- The 4 data bits — assembled as a vector p is pre-multiplied by G(i.e.Gp) and taken modulo 2 to yield the encoded value that is transmitted. The original 4 data bits are converted to 7 bits (hence the name "Hamming(7,4)") with 3 parity bits added to ensure even parity.

$$\mathbf{x} = \mathbf{G}\mathbf{p} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

This means that 0110011 would be transmitted instead of transmitting 1011

Parity Check

- If no error occurs during transmission, then the received codeword r is identical to the transmitted codeword :x

$$\mathbf{z} = \mathbf{H}\mathbf{r} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Error Correction

- Suppose a single bit error has occurred

  R=x+ e_i modulo 2 where e_i=zero vector with a 1 in the i th place.
- If we multiply this vector by H, Hr=H(x+e_i)
- Since x is the transmitted data, it is without error, and as a result, Hx=0. Thus Hr=Hx+He_i=0+ He_i

- For example:

$$\mathbf{r} = (\mathbf{x} + \mathbf{e}_5) = \left(\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\mathbf{z} = \mathbf{H}\mathbf{r} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

which corresponds to the fifth column of H.

$$\mathbf{r}_{corrected} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ \overline{1} \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

This corrected received value indeed, now, matches the transmitted value

Hamming Codes and the Hat Puzzle

- At a mathematical show with 7 players each player receives a hat either red or blue
- The color of each hat is determined by a coin toss.
- Each player can see the other person's hat but not his own.
- When the host signals, all players must simultaneously guess the color of their own hats or pass.
- The group shares a $1million prize if at least one player guesses correctly and no player guesses incorrectly.
- No communication of any sort between the players is allowed
- What should they do to maximize their chance of winning?

International Standard Book Number (ISBN)

- Ten-digit number (codeword) assigned by publisher:

$$x_1x_2x_3x_4x_5x_6x_7x_8x_9x_{10}$$

- x_1: language
- x_2x_3: publisher
- $x_4x_5 \cdots x_9$: book (assigned by publisher)
- x_{10}: assigned so that $x_{10} = \sum_{i=1\ldots.9} i\, xi = 0 \pmod{11}$

Possible to

- Detect and correct error in one digit.
- Detect transposition of two digits.

Applications(contd)

- Hamming codes over integers modulo p
- Hamming codes over an arbitrary finite field
- Widely used in computer memory(ECC)
- Storage devices (CD, DVD, DRAM), mobile communication (cellular telephones, wireless, microwave links), digital television, and high-speed modems (ADSL, xDSL).

References

- Raymond Hill (1986).A First Course in Coding Theory. Oxford University Press, Oxford.
- Richard Hill (1996).Elementary Linear Algebra with Applications.

  Harcourt, Orlando, 3rd edition.
- D. E. Shasha, *Puzzling Adventures*, W. W. Norton, New York, 2005