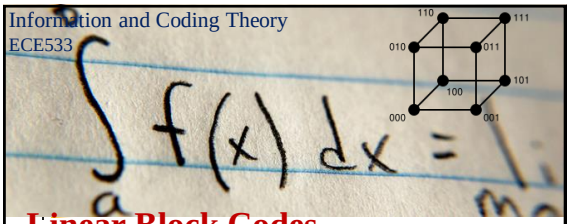
Information and Coding Theory
ECE533



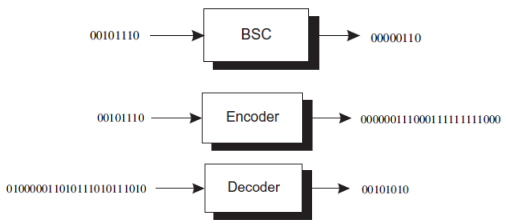
Linear Block Codes

Nikesh Bajaj
nikesh.14730@lpu.co.in
Digital Signal Processing
Lovely Professional University

Overview

- Codes?
- Digital Communication
- Signal Detection
- Channel
- Hamming Codes
- Shannon Theorem

Channel Coding



Channel Coding

- Words and Codeword
- Code(n,k) or code(n,k,d)
- Code rate k/n
- Code C = {000 111 }
- Hamming distance *d*
- Hamming weight *w*
- Error detection capability *ed*
- Error Correction Capability *ec*

Block Codes

- Words, Codeword
- Code rate
- Hamming distance/Min d*
  - Non binary
- Hamming weight/ Min w*
  - d(c1,c2)=w(c1-c2)
- Block length (n)
  - Code(n,k) $M=q^k$

Linear Block Codes

- All Zeros
- Sum of any two codewords is a codeword
- d* = w* Proof
  - Example
    - C ={00000 10100 11110 11001}
    - C={0100 1111}
    - C ={0000 1010 0101 1111}
    - C={12, 21} over GF(3)

Encoding

- C ={0000 1010 0101 1111}

Example : Repetition Code

- *Code(n,1)* *n is Odd*
  - *Code(3,1)*
  - 0 → 000
  - 1 → 111
  - Perform following for this code
    - Encode Msg = 11010100110
    - Decode Rx = 111101000010001001001111

Encoding: Matrix form G

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

Linear Codes

- Encoding

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

Generator Matrix (Properties)

- Size kxn,
- *rank k*
- Rows are linearly independent and basis vector
  - Any set of basis vector can be used to generate code span
- Generator matrix is not unique for given linear code
- Saves memory space
- Sufficiency

Find linear code

- Q: S={1100 0100 0011} over GF(2)
- Q: S={12 21} over GF(3)

Systematic Codes

- Code(5,2) over GF(3)

S.N.	Information symbols (k = 2)	Codewords (n = 5)
1.	00	00 000
2.	01	01 121
3.	02	02 220
4.	10	10 012
5.	11	11 221
6.	12	12 210
7.	20	20 020
8.	21	21 100
9.	22	22 212

Find systematic G

- .
- G=
- $\begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 2 & 2 & 1 \end{bmatrix}$

Linear Codes

- Encoding

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

- Check matrix or Parity Check Matrix H
  - G =[I | P] or G = [P | I]
  - H = [$-P^T$ | I] or H = [I | $-P^T$]

Parity check matrix H

- Size==??
- purpose
  - Limitation
- Uniqueness
- Sufficiency
- d* =min rows in H^T or column of H whose sum is zero (in linear code) and binary

Hamming Code

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

Parity check matrix H

- Q: find for H
  - . 0 1 2 1
  - G= 1 0 1 0
  - 1 2 2 1

What you can find?

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

Error Correction and Detection

Singleton Bound

Singleton Bound For code(n,k)

- $d^* \leq n - k + 1$

Hamming Sphere

Maximum Distance Separable (MDS) code

- $d^* = n-k+1$

Error detecting capability

- Code (n, k, d*)
- d*-1
- C={000 111}
- C={001 110 101}

Decoding

- Spheres of t radius, t errors
- Vector space, q-ary, n-tuples
- $d^* \geq 2t+1$
  - Worst case
- Nearest neighborhood decoding
  - C={00000 01010 10101 11111}
  - Tx=11111--- Rx =11110
  - Tx=00000 --- Rx =01000

Decoding

We call such nonintersecting spheres *decoding spheres* and such a decoder a *bounded-distance decoder.*

- Coders
  - Incomplete Coder
  - Complete Coder

Coset

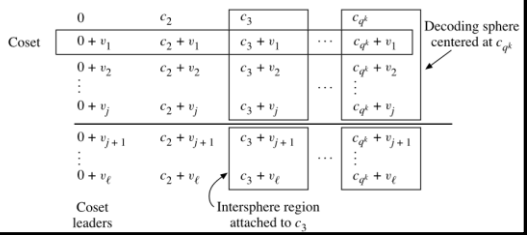
- *For C (n, k) over GF(q) and a be any vector of length n then the set*
  - Coset a + C ={c1+a, c2+a … } c1, c2 ∈C
- Theorems (Prove) *Ref: 106 in RB*
  - *Every vector b of length n is in some Coset*
  - *Each Coset contain exactly q^k vector*
  - *Two Cosets are either disjoint or Concide*
  - *If a+C is Coset of C and b ∈a+C then b+C =a+C*
  - *a and b are said to be in same Coset if (a-b) ∈C*

Coset

- Coset Leader: Vector of minimum weight
  - If two, choose any one
- Find Cosets for G=[1 0 1; 0 1 0]
- GF(q^n) = C ∪ (a1+C) ∪ (a2+C)… ∪ (at+C)
- t =q^{n-k} -1

Standard Array

- *Standard Array: for Code C (n,k), q^{n-k} x q^k array of all vector In GF(qn) …*



Standard array: Decoding

- STEPS
  - ….
  - C={0000 1011 0101 1111}

0000	1011	0101	1111
1000	0011	1101	0110
0100	1111	0001	1010
0010	1001	0111	1100

↑ coset leader

- If Rx v = 1101
- c= v - e

Standard array: Decoding

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

This code corrects one error. The standard array is

0	0	0	0	0		1	0	1	1	1		0	1	1	0	1		1	1	0	1	0
0	0	0	0	1		1	0	1	1	0		0	1	1	0	0		1	1	0	1	1
0	0	0	1	0		1	0	1	0	1		0	1	1	1	1		1	1	0	0	0
0	0	1	0	0		1	0	0	1	1		0	1	0	0	1		1	1	1	1	0
0	1	0	0	0		1	1	1	1	1		0	0	1	0	1		1	0	0	1	0
1	0	0	0	0		0	0	1	1	1		1	1	1	0	1		0	1	0	1	0
0	0	0	1	1		1	0	1	0	0		0	1	1	1	0		1	1	0	0	1
0	0	1	1	0		1	0	0	0	1		0	1	0	1	1		1	1	1	0	0

Example

For a (7, 3) code, a generator matrix is

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

0000000	0111100	1011010	1100110	1101001	1010101	**0110011**	0001111
1000000	1111100	0011010	0100110	0101001	0010101	1110011	1001111
0100000	0011100	1111010	1000110	1001001	1110101	0010011	0101111
0010000	0101100	1001010	1110110	1111001	1000101	0100011	0011111
0001000	0110100	1010010	1101110	1100001	1011101	0111011	0000111
0000100	0111000	1011110	1100010	1101101	1010001	0110111	0001011
0000010	0111110	1011000	1100100	1101011	1010111	0110001	0001101
0000001	0111101	1011011	1100111	1101000	1010100	0110010	0001110
1100000	1011100	0111010	0000110	0001001	0110101	1010011	1101111
1010000	1101100	0001010	0110110	0111001	0000101	1100011	1011111
0110000	0001100	1101010	1010110	1011001	1100101	0000011	0111111
1001000	1110100	0010010	0101110	0100001	0011101	1111011	1000111
0101000	0010100	1110010	1001110	1000001	1111101	**0011011**	0100111
0011000	0100100	1000010	1111110	1110001	1001101	0101011	0010111
1000100	1111000	0011110	0100010	0101101	0010001	1110111	1001011
1110000	1001100	0101010	0010110	0011001	0100101	1000011	1111111

Syndrome: Decoding

- Why syndrome?

Theorem 3.4.1. *All vectors in the same coset have the same syndrome that is unique to that coset.*

x and y are in same coset iff both have same syndrome

					Syndrome
Codewords →	0000	1011	0101	1111	00
	1000	0011	1101	0110	11
	0100	1111	0001	1010	01
	0010	1001	0111	1100	10
	↑				
	coset leader				

Syndrome: Decoding

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

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

$v = 10010$
$s = vH^T = 101.$
$10010 - 01000 = 11010,$
∴ dataword is 11.

The new table is

Coset Leader	Syndrome
0 0 0 0 0	0 0 0
0 0 0 0 1	0 0 1
0 0 0 1 0	0 1 0
0 0 1 0 0	1 0 0
0 1 0 0 0	1 0 1
1 0 0 0 0	1 1 1
0 0 0 1 1	0 1 1
0 0 1 1 0	1 1 0

Example: Syndrome Decoding

For a (7, 3) code, a generator matrix is

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

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

$\mathbf{r} = [0, 0, 1, 1, 0, 1, 1],$
$\mathbf{s} = \mathbf{r}H^T = [0 \quad 1 \quad 0 \quad 1]$
$\mathbf{e} = [0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0].$
$\hat{\mathbf{c}} = [0, 0, 1, 1, 0, 1, 1] + [0, 1, 0, 1, 0, 0, 0] = [0, 1, 1, 0, 0, 1, 1],$

Error	Syndrome
0000000	0000
1000000	1000
0100000	0100
0010000	0010
0001000	0001
0000100	0111
0000010	1011
0000001	1101
1100000	1100
1010000	1010
0110000	0110
1001000	1001
0101000	**0101**
0011000	0011
1000100	1111
1110000	1110

Important skills that you should have is to solve such problems

Problems

- Decode Rx sequence and compute Transmitted Tx and Original massage Mx if LBC with generator matrix G is used over GF(2)
- *Rx =1101010001010101001010001110101*0*
- Consider different G as separate case

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

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

Important skills that you should have is to solve such problems

Problems

- Decode Rx sequence and compute Transmitted Tx and Original massage Mx if LBC with generator matrix G is used over *GF(3)*
- *Rx =021201020200011110202*1*
- Consider different G as separate case

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

Bounds and Codes

- Singleton Bound
- Hamming Bound
- Perfect Code
- Hamming Code
- Optimal Code
- MDS Code

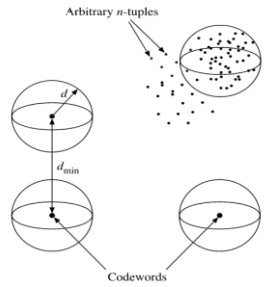
Sphere and # vector

- A sphere of radius t (0<=t<=n) contains vectors

$$V = \sum_{\ell=0}^{t} \binom{n}{\ell} (q-1)^{\ell}$$

  - Example: # of vector of distance 2 from a vector of length 4 in GF(2)
- Hamming Bound
  - MV <= q^n

Hamming sphere



Perfect Codes

- *Perfect Code is one which achieves the Hamming Bound*
  - MV = q^n

Definition 3.5.1. *A perfect code is one for which there are spheres of equal radius about the codewords that are disjoint and that completely fill the space.*

Equivalent Codes

- Two q-ary codes are called equivalent codes if one can be obtained from other by one or both
  - Permutation of components
  - Permutation of position of codewords

$$C = \begin{Bmatrix} 2 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{Bmatrix} \qquad C = \begin{Bmatrix} 2 & 2 & 2 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{Bmatrix}$$

Equivalent Linear Codes

- Two *Linear* q-ary codes are called equivalent codes if one can be obtained from other by one or both
  - Multiplication of components by any non-zero scalar
  - Permutation of position of codewords

(i) Permutation of rows,
(ii) Multiplication of a row by a non-scalar,
(iii) Addition of a scalar multiple of one row to another,
(iv) Permutation of columns,
(v) Multiplication of any column by a non-zero scalar.

Hamming Codes (binary*)

- $n = 2^m - 1$
- $k = 2^m - m - 1$
- $(n, k) = (2^m - 1, 2^m - m - 1)$

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

Optimal Linear Codes

- For optimal code (n, k, d*)
  - No (n-1, k, d*),
  - (n+1, k+1, d*) or
  - (n+1, k, d*+1) exist
  - (24, 12, 8)
  - 23, 12, 8 X 23 12 7
  - 25, 13, 8 X
  - 25, `12, 9 X

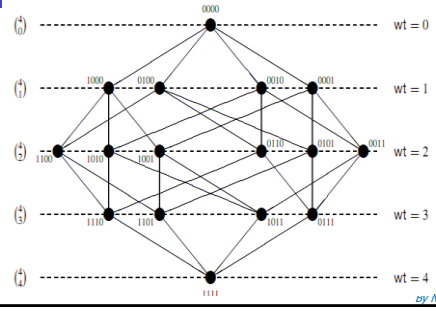
MDS Codes

- (n, n-r, d*) code satisfy d* =r+1
- (n, n-r, r+1) is MDS Codes

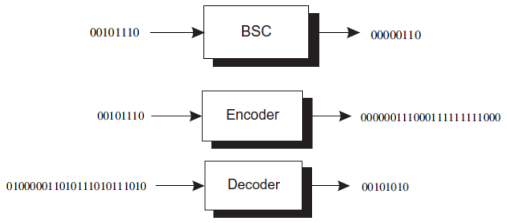
Linear Codes

- *A linear code $\mathcal{C}$ is a subspace of $GF(q)^n$.*
  - Set S,
  - Extended field GF(q^n) linear span <S>
  - linear span <S> is sub space of GF(q^n)
  - Linear code C= <S>

Visualization of Code space



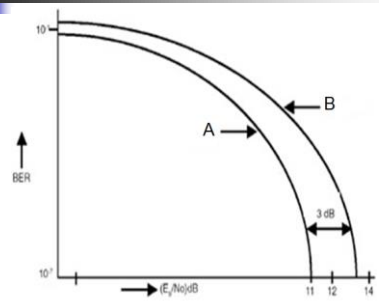
Channel Coding



Repetition Code

- 0—000
- 1—111
- Code rate =1/n
  - Example 2.4 from Ranjan boss

Coding Gain



By Nikesh Bajaj