Coding and Information Theory Chapter 7: Linear Codes - D

Xuejun Liang

Chapter 7: Linear Codes

1. Matrix Description of Linear Codes
2. Equivalence of Linear Codes
3. Minimum Distance of Linear Codes
4. The Hamming Codes
5. The Golay Codes
6. The Standard Array
7. Syndrome Decoding

Quick Review of Last Lecture

- Equivalence of Linear Codes
  - The definition of equivalence of linear codes C_1 and C_2
  - Generator matrix in systematic form $G = (I_k|P)$
  - Parity-check matrix in systematic form $H = (-P^T|I_{n-k})$
  - The Singleton Bound $d \leq 1 + n - k$
  - Examples
- Minimum Distance of Linear Codes
  - The d is the minimum number of linearly dependent columns of parity-check matrix H.
  - Meaning of linearly dependent of columns of H
  - Examples

Minimum Distance of Linear Codes

- Corollary 7.31

  There is a t-error-correcting linear $[n, k]$-code over F if and only if there is an $(n - k) \times n$ matrix H over F, of rank n - k, with every set of $2t$ columns linearly independent.

- Proof:

($\Rightarrow$)

Given such a code C, let H be a parity-check matrix for C,

So H has n columns and $n - k$ independent rows.

By Theorem 6.10, C has minimum distance $d \geq 2t + 1$.

By Theorem 7.27, every set of at most $d - 1$ columns are linearly independent

So, every set of $2t$ columns are linearly independent

- Corollary 7.31

  There is a t-error-correcting linear $[n, k]$-code over F if and only if there is an $(n-k) \times n$ matrix H over F, of rank n - k, with every set of $2t$ columns linearly independent.

- Proof:

(⇐) Given such a matrix H

let $\mathcal{V} = F^n$ and let $\mathcal{C} = \{\mathbf{v} \in \mathcal{V} \mid \mathbf{v}H^{\mathrm{T}} = \mathbf{0}\}$

Since H has rank $n-k$, its $n-k$ rows are linearly independent

So C has dimension k

By hypothesis, every set of linearly dependent columns of H contains at least $2t+1$ columns

So Theorem 7.27 implies that C has minimum distance $d \geq 2t+1$

Hence C corrects t errors by Theorem 6.10.

7.4 The Hamming Codes

$$\sum_{i=0}^{t} \binom{n}{i} (q-1)^i \leq q^{n-k}$$

- For a 1-error-correcting binary linear code, put t = 1 and q = 2 in the sphere-packing bound (Corollary 6.17), so the condition for a perfect code becomes

$$2^{n-k} = 1 + \binom{n}{1} = 1 + n$$

- Let $c = n - k$ (the number of check digits), then this condition is equivalent to

$$n = 2^c - 1. \qquad (7.4)$$

- So

$c =$	1	2	3	4	5	...
$n =$	1	3	7	15	31	...
$k =$	0	1	4	11	26	...

The Hamming Codes (Cont.)

Construct codes with these parameters on $F_2 = \{0,1\}$

- By Corollary 7.31, need to construct a c x n matrix H over F_2 , of rank c, with every pair of columns linearly independent (non-zero and distinct).
- Columns of H must consist of all 2^c - 1 non-zero binary vectors of length c, in some order.
- This matrix H has rank of c. Use it as the parity-check matrix, we have a code C with these parameters. This code is called the **binary Hamming code $\boldsymbol{H_n}$** of length n = $2^c - 1$.

- Example 7.32
  - H_3 has the parity checking matrix $H = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$
    - c = 2, n = 3, k = 1
  - H_3 is R_3 !!!

- Note: The rate of H_n will approaches to 1.

$$R = \frac{k}{n} = \frac{2^c - 1 - c}{2^c - 1} \to 1$$

- Nearest neighbor decoding with H_n
  - The receiver computes $s = vH^T$,
    - Called the syndrome of v.
  - If s = 0, the receiver decodes v as $\Delta(v) = v$, and
  - if s = ${c_i}^T$ (the i-th column of H) then $\Delta(v) = v - e_i$.

Nearest Neighbor Decoding

- Example 7.33
  - Let us use H_7, with parity-check matrix

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

$v = (1\ \ 1\ \ 0\ \ 1\ \ 1\ \ 0\ \ 1)$

  - Suppose that u = 1101001 is sent, and v = 1101101 is received, so the error-pattern is e = e_5.
  - The syndrome is $s = vH^T$ = 100, which is the transpose ${c_5}^T$ of the fifth column of H.
  - This indicates an error in the fifth position, so changing this entry of v we get $\Delta(v)$ = 1101001 = u

- Using the parity checking matrix as below, then a non-zero syndrome is the binary representation of the position i where a single error e, has appeared

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

1 2 3 4 5 6 7

- Example 7.34
  - Verify this using example 7.33

u = 1101001

v = 1101101

$s = vH^T$ = 101

$\Delta(v)$ = 1101001 = u

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

v = (1 1 0 1 1 0 1)

- Note: need to perform a column permutation (1362547) to change between the two representations.

Construction of perfect 1-error-correcting linear codes for prime-powers q > 2

- We take the columns of H to be

$$n = \frac{q^c - 1}{q - 1} = 1 + q + q^2 + \cdots + q^{c-1}$$

$$\sum_{i=0}^{t} \binom{n}{i} (q-1)^i \leq q^{n-k}$$

pairwise linearly independent vectors of length c over F_q.

- The resulting linear code has length n, dimension k = n - c, and minimum distance d = 3, so t = 1.

- As in the binary case, $R \to 1$ as $c \to \infty$, but $\Pr_E \nrightarrow 0$.

Construction of perfect 1-error-correcting linear codes for prime-powers q > 2

- Example 7.35
  - If $q = 3$ and $c = 2$, then $n = 4$ and $k = 2$.
  - We can take

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix}$$

  - The solutions of the simultaneous linear equations

$$vH^T = 0$$

  will give a perfect 1-error-correcting linear [4, 2]-code over F_3

$$n = \frac{q^c - 1}{q - 1} = 1 + q + q^2 + \cdots + q^{c-1}$$

7.5 The Golay Codes

- Skip this section

7.6 The Standard Array

- Suppose $C \subseteq V$ is a linear code. The standard array of C is essentially a table in which the elements of V are arranged into cosets of the subspace C.
- Suppose that $C = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M\}$ is a linear code with $M = q^k$ elements. Assume $\boldsymbol{u_1} = \mathbf{0}$.
- For $i = 1, \ldots, q^{n-k} - 1$, let the i-th row consist of the elements of the coset of C.

$$\mathbf{v}_i + C = \{\mathbf{v}_i + \mathbf{u}_1 \ (= \mathbf{v}_i),\ \mathbf{v}_i + \mathbf{u}_2,\ \ldots,\ \mathbf{v}_i + \mathbf{u}_M\}$$

  where $wt(v_i) \leq wt(v_{i+1}), i = 1, \ldots, q^{n-k} - 1$ and v_i is not in the previous (< i) rows.

- A horizontal line across the array, immediately under the last row to satisfy $wt(v_i) \leq t$, where $t = \lfloor (d-1)/2 \rfloor$.

The Standard Array (Cont.)

- Example 7.39
  - Let C be the binary repetition code R_4 of length n = 4, so q = 2, k = 1 and the code-words are $\boldsymbol{u_1}$ = **0** = 0000 and $\boldsymbol{u_2}$ = **1** = 1111
  - There are $q^{n-k} = 8$ cosets of C in V, each with two vectors
  - So, standard array has 8 rows:

    $$v_1 + C, v_2 + C, \dots, v_8 + C$$

    $v_1 = has\ weight\ 0$
    $v_2\ to\ v_5\ has\ weight\ 1$
    $v_6, v_7, v_8\ has\ weight\ 2$

	$\boldsymbol{u_1}$	$\boldsymbol{u_2}$
	0000	1111
$v_1 + C$	0000	1111
$v_2 + C$	1000	0111
$v_3 + C$	0100	1011
$v_4 + C$	0010	1101
$v_5 + C$	0001	1110
$v_6 + C$	1100	0011
$v_7 + C$	1010	0101
$v_8 + C$	1001	0110

The Standard Array (Cont.)

- Lemma 7.40
  a) If v is in the j-th column of the standard array (that is, $v = v_i + u_j$ for some i), then u_j is a nearest code-word to v.
  b) If, in addition, v is above the line in the standard array (that is, $wt(v_i) \leq t$), then u_j is the unique nearest code-word to v.

	$\boldsymbol{u_1}$	$\boldsymbol{u_2}$	
	0000	1111	
$v_1 + C$	0000	1111	
$v_2 + C$	1000	0111	
$v_3 + C$	0100	1011	
$v_4 + C$	0010	1101	$v = v_4 + u_2$
$v_5 + C$	0001	1110	
$v_6 + C$	1100	0011	
$v_7 + C$	1010	0101	
$v_8 + C$	1001	0110	

The Standard Array (Cont.)

- The sphere $S_t(u_j)$ of radius t about u_j is the part of the j-th column above the line.
- Thus C is perfect if and only if the entire standard array is above the line

The Standard Array (Cont.)

- Decoding rule
  - Suppose that a code-word $u \in C$ is transmitted, and $v = u + e \in V$ is received, where e is the error-pattern.
  - The receiver finds $v = v_i + u_j$ in the standard array, and decides that $\Delta(v) = u_j$ (u_j is header of a column)
- Note this rule gives correct decoding if and only if the error-pattern is a coset leader ($e = v_i$).

	$\boldsymbol{u_1}$	$\boldsymbol{u_2}$	
	0000	1111	
$v_1 + C$	0000	1111	
$v_2 + C$	1000	0111	
$v_3 + C$	0100	1011	
$v_4 + C$	0010	1101	
$v_5 + C$	0001	1110	$v = v_5 + u_2$
$v_6 + C$	1100	0011	
$v_7 + C$	1010	0101	
$v_8 + C$	1001	0110	

Example 7.41

- Let $C = R_4$. Suppose that $\boldsymbol{u}$ = 1111 is sent, and the error-pattern is e = 0100, $v = ?$ and $u_j = ?$

- How about when e = 0110?

- How about when e = 1100?

	$\boldsymbol{u_1}$	$\boldsymbol{u_2}$	
	0000	1111	u
$v_1 + C$	0000	1111	
$v_2 + C$	1000	0111	
$v_3 + C$	0100	1011	
$v_4 + C$	0010	1101	
$v_5 + C$	0001	1110	
$v_6 + C$	1100	0011	
$v_7 + C$	1010	0101	
$v_8 + C$	1001	0110	

7.7 Syndrome Decoding

- If H is a parity-check matrix for a linear code $C \subseteq V$ then the syndrome of a vector $v \in V$ is the vector

$$\mathbf{s} = \mathbf{v}H^{\mathrm{T}} \in F^{n-k} \qquad (7.8)$$

- Lemma 7.42
  - Let C be a linear code, with parity-check matrix H, and let $v, v' \in V$ have syndromes s, s'. Then v and v' lie in the same coset of C if and only if $s = s'$.
- Proof of Lemma 7.42

$$\begin{aligned}
\mathbf{v} + \mathcal{C} = \mathbf{v}' + \mathcal{C} &\iff \mathbf{v} - \mathbf{v}' \in \mathcal{C} \\
&\iff (\mathbf{v} - \mathbf{v}')H^{\mathrm{T}} = \mathbf{0} \qquad \text{(by Lemma 7.10)} \\
&\iff \mathbf{v}H^{\mathrm{T}} = \mathbf{v}'H^{\mathrm{T}} \\
&\iff \mathbf{s} = \mathbf{s}'.
\end{aligned}$$