# Algebraic Coding Theory (Revisited)

## — From the perspective of linear algebra

Hengfeng Wei

hfwei@nju.edu.cn

April 17 $\sim$ April 20, 2017

# Algebraic Coding Theory (Revisited)

1. Block Codes
2. Linear Codes

# Block coding

flow chart here

Where is cryptography?

# Important code parameters

$$k$$

$$n > k$$

$$r = n - k$$

$$|C| \leq 2^n$$

$$0 < R \triangleq \frac{k}{n} < 1$$

# Hamming distance

$$w(c) = \#1's \text{ in } c$$

$$d(c_1, c_2) = w(c_1 + c_2)$$

$$\begin{aligned} d(C) &= \min\{d(c_1, c_2) \mid c_1 \neq c_2, c_1, c_2 \in C\} \\ &= \min\{w(c_1 + c_2) \mid c_1 \neq c_2, c_1, c_2 \in C\} \\ &\neq \min\{w(c) \mid c \neq 0, c \in C\} \end{aligned}$$

# Detecting and correcting errors

$$d(C) \geq 2t+1 \implies 2t\text{-detecting}$$

$$d(C) \geq 2t+1 \implies t\text{-correcting}$$

# Sphere-packing bound

## Theorem (Sphere-packing bound)

*A $t$-error-correcting binary code of length $n$ must satisfy*

$$|C| \sum_{i=0}^{t} \binom{n}{i} \leq 2^n$$

$$t = 1 \implies |C| \leq \frac{2^n}{n+1}$$

## Definition (Perfect code)

$$|C| \sum_{i=0}^{t} \binom{n}{i} = 2^n$$

# Algebraic Coding Theory (Revisited)

1. Block Codes
2. Linear Codes

# Definition of linear code

**Definition (Linear code)**

A linear code $C$ of length $n$ is a linear subspace of the vector space $\mathbb{F}_2^n$.

$$c_1 \in C, c_2 \in C \implies c_1 + c_2 \in C$$

$$\begin{aligned} d(C) &= \min\{w(c_1 + c_2) \mid c_1 \neq c_2, c_1, c_2 \in C\} \\ &= \min\{w(c) \mid c \neq 0, c \in C\} \end{aligned}$$

# Definition of linear code

**Problem TJ–8.18**

Let $C$ be a linear code.
Show that either the $i$-th coordinates in the codewords of $C$ are all zeros or exactly half of them are zeros.

# Definition of linear code

**Problem TJ–8.19**

Let $C$ be a linear code.
Show that either every codeword has even weight or exactly half of them have even weight.

Parity: $w(c_1) + w(c_2)$ *vs.* $w(c_1 + c_2)$

# Definition of linear code

**Definition (Linear code)**

An $(n, k)$ linear code $C$ of length $n$ and rank $k$ is a linear subspace with dimension $k$ of the vector space $\mathbb{F}_2^n$.

Basis: $c_1, c_2, \ldots, c_k$

$$c_i = \alpha_1 c_1 + \alpha_2 c_2 + \cdots + \alpha_k c_k$$

$$|C| = 2^k$$

# Generator matrix

**Definition (Generator matrix)**

A matrix $G_{n\times k}$ is a generator matrix for an $(n,k)$ linear code $C$ if

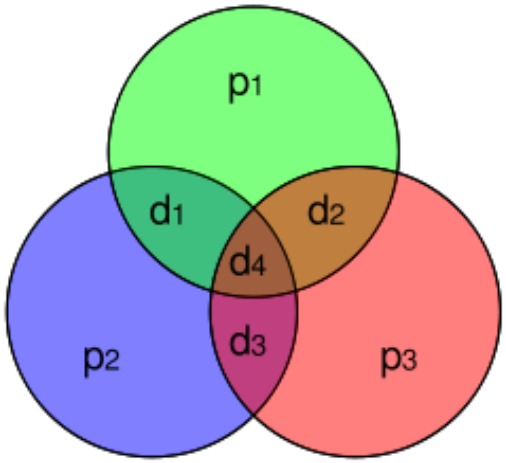
$$C = \mathsf{Col}(G)$$

$$\mathsf{rank}(G) = k$$

$$G_{(n\times k)} \cdot d_{k\times 1} = c_{n\times 1} \in C$$

$$G(c_1 + c_2) = G(c_1) + G(c_2)$$

# Generator matrix for Hamming code $(7, 4)$




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

$$G \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

# Standard generator matrix

**Problem TJ–8.7**

Generator matrices are NOT unique.

**Definition (Standard generator matrix)**

A generator matrix $G_{n\times k}$ is standard if

$$G_{n\times k} = \begin{bmatrix} I_k \\ A_{(n-k)\times k} \end{bmatrix}$$

# From generator matrix to parity-check matrix

$$G \cdot \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \cdot \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ p_1 = d_1 + d_2 + d_4 \\ p_2 = d_2 + d_3 + d_4 \\ p_3 = d_1 + d_3 + d_4 \end{pmatrix}$$

# From generator matrix to parity-check matrix

$$\begin{aligned} d_1 + d_2 \quad + d_4 + p_1 \quad &= 0 \\ d_2 + d_3 + d_4 \quad + p_2 \quad &= 0 \\ d_1 \quad + d_3 + d_4 \quad + p_3 &= 0 \end{aligned}$$

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} = 0$$

# Parity-check matrix

**Definition (Parity-check matrix)**

A matrix $H_{(n-k)\times n}$ is a parity-check matrix for an $(n, k)$ linear code $C$ if

$$C = \mathsf{Nul}(H)$$

$$H_{(n-k)\times n} \cdot c_{n\times 1} = 0_{(n-k)\times 1}$$

$$\mathsf{rank}(H) = n - k$$

# Standard parity-check matrix

**Problem TJ–8.11**

Parity-check matrices are NOT unique.

Elementary row operations.

**Definition (Standard parity-check matrix)**

A parity-check matrix $H_{(n-k)\times n}$ is standard if

$$H_{(n-k)\times n} = \left[ A_{(n-k)\times k} \mid I_{n-k} \right]$$

# Generator matrix and Parity-check matrix

$$H_{(n-k)\times n} \cdot G_{n\times k} \cdot d_{k\times 1} = 0_{(n-k)\times 1}$$

$$\begin{aligned} & H_{(n-k)\times n} \cdot G_{n\times k} \\ &= \left[A_{(n-k)\times k} \mid I_{n-k}\right] \cdot \begin{bmatrix} I_k \\ A_{(n-k)\times k} \end{bmatrix} \\ &= A_{(n-k)\times k} \cdot I_k + I_{n-k} \cdot A_{(n-k)\times k} \\ &= A_{(n-k)\times k} + A_{(n-k)\times k} \\ &= 0_{(n-k)\times k} \end{aligned}$$

# Syndrome decoding

$$r = c + e_i$$

$$r = c + (e_i + e_j + \cdots)$$

## Definition (Syndrome)

$$\begin{aligned} S(r) &= Hr \\ &= H(c + (e_i + e_j + \cdots)) \\ &= H(e_i + e_j + \cdots) \\ &= He_i + He_j + \cdots \\ &= S(e_i) + Se_j + \cdots \end{aligned}$$

# Syndrome decoding

Problem TJ–8.13

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

(d) Errors in the third and fourth bits

# Extracting $d(C)$ from $H$

**Theorem**

*If $H$ is the parity-check matrix for a linear code $C$, then $d(C)$ equals the minimum number of columns of $H$ that are linearly dependent.*

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

# Error-detecting and error-correcting capabilities

$$
\begin{aligned}
t = 1 &\Longrightarrow d(C) = 3 \\
&\Longleftrightarrow \forall\{c_i\}, \forall\{c_i, c_j\} \text{ linearly independent} \\
&\Longleftrightarrow \text{no zero column, no identical columns}
\end{aligned}
$$

# Error-detecting and error-correcting capabilities

**Problem TJ–8.21; TJ–8.23**

If we are to use an error-correcting linear code to transmit the 128 ASCII characters, what size matrix must be used? What if we require only error detection?

$$t = 1$$

$$2^r - 1 \geq 7 + r \implies r \geq 4$$

# Generalized Hamming codes

$$C_{Ham} = (7, 4)$$

$$H_{3\times 7} : r = 3, n = 2^3 - 1 = 7$$

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} = 0$$

# Generalized Hamming codes

## Definition (Generalized Hamming codes)

$H_r \triangleq H_{r\times(2^r-1)}$ : all nonzero binary vectors of length $r$

$$C_{Ham}^r = (n = 2^r - 1, k = 2^r - 1 - r)$$

$$|C_{Ham}^r| = \frac{2^n}{n+1}$$

$$R(C_{Ham}^r) = \frac{k}{n}$$


THANK YOU