Today

- Coding theory basics
- Hamming Bound
- Linear codes
- Hamming Code

CS3.318.1
Coding Theory
Lecture 2 (2022-9-2)
Instructor: Prahladh Harsha.

Coding Theory Basics

- Σ - finite alphabet

  eg: $\Sigma = \{0,1\} = \mathbb{F}_2$ (binary)

  $\Sigma = \{0,1\}^8$ (bytes)

- Σ^n: set of all n-symbol words (ambient space).

Hamming Distance: $\Delta : \Sigma^n \times \Sigma^n \to \mathbb{R}_{\geq 0}$

$x, y \in \Sigma^n$, $\Delta(x,y) = \#\{i \mid x_i \neq y_i\}$

Hamming Weight:

$x \in \Sigma^n$, $wt(x) = \#\{i \mid x_i \neq 0\}$

Observations:

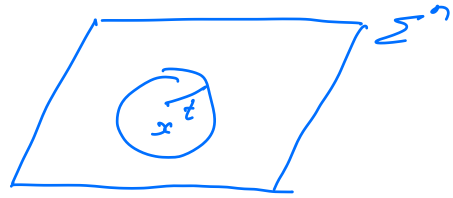
1. $\Delta(x,y) = 0 \iff x = y$
2. $\Delta(x,y) = \Delta(y,x)$
3. $\Delta(x,y) + \Delta(y,z) \geq \Delta(x,z)$

} metric distance.

Code: $\mathcal{C} \subseteq \Sigma^n$ $\quad$ $\mathcal{C}$ - code

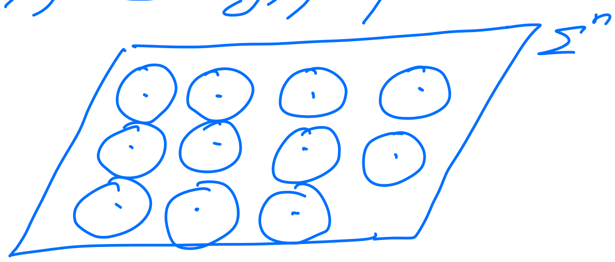
elements of $\mathcal{C}$ - codewords

$$\mathrm{Ball}(x,t) = \{y \in \Sigma^n / \Delta(x,y) \le t\}$$



<u>Defn:</u> $\mathcal{C}$ is **t-error correcting** if

$$\forall x,y \in \mathcal{C}, \quad \mathrm{Ball}(x,t) \cap \mathrm{Ball}(y,t) = \phi$$



Defn: $\mathcal{C}$ is **e-error detecting**

$$\forall x \in \mathcal{C}, \quad \mathrm{Ball}(x,e) \cap \mathcal{C} = \{x\}$$

Distance of code $\mathcal{C}$: $\Delta(\mathcal{C}) = \min\limits_{x \neq y \in \mathcal{C}} \Delta(x,y)$

fractional distance $\quad \delta(\mathcal{C}) = \dfrac{\Delta(\mathcal{C})}{n}$

Proposition [Hamming]

$\mathcal{C}$ is t-error correcting

$\Updownarrow$

$\mathcal{C}$ is $2t$-error detecting

$\Updownarrow$

$\Delta(\mathcal{C}) \geq 2t+1$

From picture, for any t-error correcting code

$$|\mathcal{C}| \cdot Vol_\Sigma(n,t) \leq |\Sigma|^n$$

$\Sigma = \{0,1\}$, $\quad Vol(n,t) = \sum_{i=0}^{t} \binom{n}{i}$ $\quad$ (For q-sized alphabets $Vol_q(n,t) = \sum_{i=0}^{t} \binom{n}{i}(q-1)^i$)

For $t=1$. & $\Sigma = \{0,1\}$

$$|\mathcal{C}| \cdot (n+1) \leq 2^n \quad \left(\text{i.e. } |\mathcal{C}| \leq \frac{2^n}{n+1}\right)$$

$$\left(n=63 \ ; \ |\mathcal{C}| \leq \frac{2^{63}}{64} = 2^{57}\right)$$

Cor: Hamming's construction can't be improved.

Packing Bound / Hamming Bound $\quad \mathcal{C} \subseteq \Sigma^n$ & $\mathcal{C}$ is t-error correcting

$$|\mathcal{C}| \leq \frac{|\Sigma|^n}{Vol_\Sigma(n,t)}$$

Linear Codes: Σ - finite field ($\mathbb{F}_q$) & $\mathcal{C} \subseteq \mathbb{F}_q^n$ linear subspace.

$\mathbb{F}_q^n \qquad \langle \ ; \ \rangle : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{F}_q$

$$\langle x, y \rangle \mapsto \sum x_i \cdot y_i$$

$$C^{\perp} = \{ y \in \mathbb{F}_q^n \mid \langle x, y \rangle = 0, \ \forall x \in C \}$$

<u>Fact:</u> ① $\dim(C) + \dim(C^{\perp}) = n$

② $(C^{\perp})^{\perp} = C.$

Representation of a linear code.

① Generator Matrix: using basis of C.

$C \subseteq \mathbb{F}_q^n$

$\dim(C) = k$

$$\underbrace{\begin{bmatrix} | & | & & | \\ c_1 & c_2 & \cdots & c_k \\ \downarrow & | & & | \end{bmatrix}}_{G \in \mathbb{F}_q^{n \times k}} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}}_{\mathbb{F}_q^k}$$

$$Enc: \mathbb{F}_q^k \to \mathbb{F}_q^n$$
$$x \mapsto Gx$$

$$C = \{ Gx \mid x \in \mathbb{F}_q^k \}$$

② Parity-check matrix - using basis of $C^{\perp}$

$$\underbrace{\begin{bmatrix} \text{---} & f_1^T & \longrightarrow \\ \text{---} & f_2^T & \longrightarrow \\ & \vdots & \\ \text{---} & f_{n-k}^T & \text{---} \end{bmatrix}}_{H \in \mathbb{F}_q^{(n-k) \times n}} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{F}_q^n$$

$$C = \{ c \in \mathbb{F}_q^n \mid Hc = 0 \}.$$

$$HG = \bar{0}_{(n-k) \times k}$$

Distance of a linear code.

$$\Delta(x,y) = \Delta(x-y, 0) = wt(x-y)$$

— min weight of a non-zero codeword

$$\Delta(C) = \min_{0 \neq c \in C} wt(c)$$

— smallest r such that there exist r dependent columns in H.

Codes Notation

C –

1. $q = |\Sigma|$; alphabet size
2. n – blocklength
3. k – dimension of C

   $k = \log_{|\Sigma|} |C|$
4. $\Delta(C) \geq d$

C – $(n,k,d)_q$ -code

$(n,k)_q$ - code

Furthermore, C is linear

$[n,k,d]_q$ -code

— <u>Hamming Codes:</u>

$$C : \{0,1\}^{57} \to \{0,1\}^{63}$$

For any positive integer r

$$H_r = \begin{array}{c} \xleftarrow{\quad 2^r - 1 \quad} \\ \begin{bmatrix} | & | & & | \\ bin(1) & bin(2) & \cdots & bin(2^r-1) \\ | & | & & | \end{bmatrix} \end{array} \Big\updownarrow r$$

$$\boxed{C_{Ham}^{(r)} = \{ c \in \mathbb{F}_2^{2^r-1} \mid H_r c = 0 \}}$$

— c $\leq$ 1-corruption

$c \longrightarrow c$ or $c + e_i$ for some $i \in [2^r - 1]$

$H_r(c + e_i) = H_r e_i = bin(i)$

— Follows, that $C_{Ham}^{(r)}$ is 1-error correcting

hence, $\Delta(C_{Ham}^{(r)}) \geq 3$, $\forall r$.

Claim: $\Delta(C_{Ham}^{(r)}) = 3$.

Pf: There exist 3 cols in H_r (say $bin(1)$, $bin(2)$ & $bin(3)$) which are dependent. ☒

Hence, $C_{Ham}^{(r)}$ is $[2^r - 1, 2^r - r - 1, 3]_2$

Hamming Bound for t-error correcting codes

$$|C| \leq 2^n / Vol(n, t) \quad \ldots (A)$$

Perfect codes are codes for which (*) is tight

Theorem [Tietavainen & van Lint] The only perfect codes are over $\{0,1\}$

- Hamming codes $\mathcal{C}_{Ham}^{(r)}$
- Trivial examples: $\mathcal{C} : |\mathcal{C}| = 1$
  $\mathcal{C} = \{0^n, 1^n\}$, n - odd.
- Golay code $G = [23, 12, 7]_2$-code.

$(c_0 \ldots, c_{22}) \in G$ (G - cyclic code)

$\Leftrightarrow$ $c_0 + c_1 X + \ldots + c_{22} X^{22}$ is a multiple of
$1 + X + X^5 + X^6 + X^7 + X^9 + X^{11}$
in $\mathbb{F}_2[X] / (X^{23} - 1)$

Families of Codes:

$[n, k, d]_q$ $\qquad$ $\{\mathcal{C}_n\}_{n=1}^{\infty}$

$$\delta(\mathcal{C}) = \frac{d}{n}$$

$$\text{Rate} = \frac{k}{n} = \frac{\log_{|\Sigma|} |\mathcal{C}|}{n}$$

Qn: Given $R, \delta \in (0,1)$, does there exist a family of codes $\{\mathcal{C}_n\}$ s.t.

$$\left.\begin{array}{l} R(C_n) \geq R \\ \delta(C_n) \geq \delta . \end{array}\right\}$$ good codes. (understood)

Qn: R, vs δ tradeoff – open.