Information Theory & Coding

Nov 24, 2020

Yesterday : Linear Codes :

$u \in \{0,1\}^k \rightsquigarrow x \in \{0,1\}^n$ (mod 2)

$\Leftrightarrow$ $x = Gu$

$G \in \{0,1\}^{n \times k}$ generator matrix

Such x's satisfy $Hx = 0$ parity check matrix.

$\Leftarrow$

Example : Hamming Codes :

$\left(\underbrace{2^r - 1}_{n}, \underbrace{2^r - 1 - r}_{k}, 3 \right)$ d_{min}

$H = \left[\quad \vdots \quad \right] \Big\} r$

$2^r - 1$

Columns = set of $\{0,1\}^r$ vectors except all-zero vector

$H = \left[I_r \; ; \; A \right]$

$G = \begin{bmatrix} A \\ - \\ I \end{bmatrix}$

These were "single-error correcting" codes,

i.e. :

if a codeword x is modified to

$$y = x \oplus z \quad \text{with} \quad w_H(z) \le 1$$

then from y we can determine x. Reason:

Compute $s = Hy$ if $z =$ unit vector i

then $s =$ ith column of H.

rate: $R = \dfrac{2^r - 1 - r}{2^r - 1} = 1 - \dfrac{r}{2^r - 1}$

We had seen the Singleton bound:

If a code has $M \ge 2^k$ codewords, blocklength $= n$, then $d_{min} \le n - k + 1$

$\Rightarrow$ $M = 2^k > 2^{k-1} \Rightarrow d_{min} \le n - k + 1$

Consider now non-Binary codes.

$$enc(m) \in \mathcal{X}^n, \quad |\mathcal{X}| = q.$$

Suppose the number of codewords $M > q^k$.

Then, via considering the hash function

$hash(x_1 \dots x_n) = (x_1 \dots x_k)$ we see that

# of possible hash values is $q^k < M$

$\Rightarrow \exists\, m_i \neq m'$ s.t $hash(enc(m_i)) = hash(enc(m'))$.

$\Rightarrow enc(m)_i = enc(m')_i \quad \forall 1 \le i \le k$

$\Rightarrow d_H(enc(m), enc(m')) \le n-k$

$d_H(x,y) = \sum_{i=1}^{n} \mathbb{1}\{x_i \neq y_i\} \quad x, y \in \Sigma^n$.

$\Rightarrow$ d_{min} of a q-ary code of block length n

& $M \ge q^k$ codewords

is at most $n-k$.

- We will now construct a family of codes for which $M = q^k$, blocklength $= n$ &

$d_{min} = n-k+1$.

Reed-Solomon codes

Remember: algebraic fields.

An algebraic field is a set $\mathbb{F}$ of elements and two operations "$+$" and "$\cdot$" satisfying the rules of algebra:

$$\begin{cases} a+b = b+a \\ a+(b+c) = (a+b)+c \\ \exists\, 0 \in \mathbb{F} \text{ s.t } 0+a = a \ \forall a \\ \forall a \in \mathbb{F} \ \exists (-a) \text{ s.t } a+(-a) = 0. \end{cases}$$

$$\begin{cases} a \cdot b = b \cdot a \\ a \cdot (b \cdot c) = (a \cdot b) \cdot c \\ \exists\, 1 \in \mathbb{F} \text{ s.t } a \cdot 1 = a \ \forall a \in \mathbb{F} \quad 1 \neq 0 \\ \forall a \in \mathbb{F} \setminus \{0\} \ \exists (a^{-1}) \text{ s.t } a \cdot a^{-1} = 1. \end{cases}$$

$$a \cdot (b+c) = (a \cdot b) + (a \cdot c).$$

when $|\mathbb{F}| < \infty$ we call the field a finite field. ($\equiv$ Galois field)

then Galois if $(\mathbb{F}, +, \cdot)$: finite field

then $|\mathbb{F}| = p^m$ where p : prime
$m \geq 1$: integer.

Moreover if $q = p^m$ for some prime p & pos. integer m

then there is a finite field $(\mathbb{F}, +, \cdot)$ with $|\mathbb{F}| = q$.

Ex: $\mathbb{F} = \{0, 1, \dots, p-1\}$ p is prime

$+, \cdot$ = mod-p arithmetic

examples $\mathbb{F} = \{0, 1, 2\}$ with mod 3 arithmetic

$\mathbb{F} = \{0, 1, 2, 3, 4\}$ mod 5

Non examples: $\mathbb{F} = \{0, 1, 2, 3\}$ mod 4 arithmetic

we can however find a $\mathbb{F}$ with 4 elements and appropriate arithmetic:

$\mathbb{F} = \{0, 1, x, x+1\}$: polynomials with binary coefficients deg ≤ 1.

$+$: mod 2 polynomial addition

eg $\underbrace{x} + \underbrace{(x+1)} = 1$

"$\cdot$": polynomial multiplication mod 2,

and mod $\boxed{x^2 = x+1}$.

$\cdot$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	$x+1$	1
$x+1$	0	$x+1$	1	x

other options:
$$\begin{cases} x^2 = 0 \Leftarrow \text{no. because } x\cdot x = 0 \\ \underline{x^2 = 1} \leftarrow \text{"} (x+1)(x+1) = 0 \\ x^2 = x \leftarrow x(x+1) = 0 \end{cases}$$

~~$x^2 = x+1$~~

In general: if we want to have a field with

p^m elements: we let

$\mathbb{F} = \{$ polynomials with coefficients $\{0, \dots, p-1\}$

of degree $\leq m-1 \}$

$+$: mod p polynomial addition

• : mod p polynomial multiplication with

the reduction $x^m = r(x)$

↑ a special degree $\leq m-1$ polynomial.

↙

s.t $x^m - r(x)$ has no factors in $\mathbb{F}$.

Fact: such $r(x)$ always exists.

Given $\mathcal{X}$ with $|\mathcal{X}| = q = p^m$, how do we construct a code with q^k codewords, block length n & $d_{min} = n - k + 1$

Restriction: $k \leq n \leq |\mathcal{X}| = q$

① Make $\mathcal{X}$ into a field $\mathbb{F}$. (i.e., find ways of operations on $\mathcal{X}$ to obey the rules of algebra.)

② Pick distinct elements $\alpha_1, \ldots, \alpha_n$ in $\mathcal{X}$.

③ the codewords are to be indexed by $u = (u_0, \ldots, u_{k-1}) \in \mathcal{X}^k$. there are q^k such u's.

and the codeword x that corresponds to u

$\in \mathcal{X}^n$ $\in \mathcal{X}^k$

is given as follows: $\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{u(D)}$

$$x(u)_1 = \left(u_0 + u_1 D + u_2 D^2 + \cdots + u_{k-1} D^{k-1}\right)\Big|_{D=\alpha_1}$$

$$x(u)_2 = \left(\quad\right)\Big|_{D=\alpha_2}$$

$$\vdots$$

$$x(u)_n = \left(\quad\right)\Big|_{D=\alpha_n}$$

$$x(u)_i = u(\alpha_i).$$

this defines a code with blocklength $= n$

with q^k codewords.

question: what is the minimum distance d?

we know $d \le n - k + 1$ (Singleton).

we will show $d \ge n - k + 1$. $\Rightarrow$ $d = n - k + 1$.

To that end, consider $u \ne v \in \mathcal{X}^k$

and the corresponding codewords

$x(u) \ne x(v) \in \mathcal{X}^n$, and form

$x(u) - x(v)$, here see that

$$d_H(x(u), x(v)) = \sum_{i=1}^{n} \mathbb{1}\{x(u)_i - x(v)_i \ne 0\}.$$

note that

$$\big(x(u)-x(v)\big)_i = u(\alpha_i) - v(\alpha_i)$$

$$= \sum_{j=0}^{k-1} (u_j - v_j)\,\alpha_i^j$$

$$= w(\alpha_i) \quad \text{where}$$

$$w(D) = (u_0 - v_0) + (u_1 - v_1)D + \cdots + (u_{k-1} - v_{k-1})D^{k-1}$$

polynomial of degree $\le k-1$.

and

$$d_H(x(u), x(v)) = \sum_{i=1}^{n} \mathbb{1}\{w(\alpha_i) \neq 0\}$$

$$= n - \sum_{i=1}^{n} \mathbb{1}\{w(\alpha_i) = 0\}.$$

But the number of roots of $w(D) \le k-1$.

$$\sum_{i=1}^{n} \mathbb{1}\{w(\alpha_i) = 0\} \le \#\text{ of roots of } w \le k-1$$

$$\Rightarrow \quad d_H(x(u), x(v)) \ge n - (k-1) = n-k+1$$

$$\Rightarrow \quad d_{min} \ge n - k + 1$$

Example : $\mathcal{X} = \{0, 1, 2\}$

$\mathcal{X}$ with mod-3 arithmetic is a field.

Suppose $n = 3$, $k = 2$

p-ch. $\alpha_1 = 0$, $\alpha_2 = 1$, $\alpha_3 = 2$.

$u_1 u_2$	$u(D)$	$x_1 x_2 x_3$
0 0	0	0 0 0
0 1	1	1 1 1
0 2	2	2 2 2
1 0	D	0 1 2
1 1	D+1	1 2 0
1 2	D+2	2 0 1
2 0	2D	0 2 1
2 1	2D+1	1 0 2
2 2	2D+2	2 1 0

3^2 codewords

$BL \Rightarrow$

$d = 2$

usual choices for $|\mathcal{X}|$ are powers of two

e.g. 2^8, 2^{16}.

Ex: $|\mathcal{X}| = 2^8$ elements of $\mathcal{X}$ are bytes
$\equiv$ 8 bit strings

Suppose we construct a RS code with
$k = 4$ & $n = 8$, & $d = 5$

# of codewords $= (2^8)^4 = 2^{32}$ codewords.

$\Rightarrow$ block length = 8 bytes = 64 bits

$d_{min} = 5$: any two codewords will differ
in ≥ 5 bytes.

min d_H in bits may be as small as 5 bits
as big as 40 bits.

Coming next: Polar Codes.