$$H(X,Y) = H(Y) + H(X|Y) \quad \| \quad H(X|Y) = H(X,Y) - H(Y)$$
$$= H(X) + H(Y|X)$$

$$H(X,Y) \begin{cases} \leq H(Y) \\ \leq H(X) \end{cases} \quad \Big| \quad H(X|Y) \overset{\text{ind}}{\leq} H(X)$$
$$H(X|YZ) \leq H(X|Z)$$

$$H(X) + H(Y) \overset{\text{ind}}{\geq} H(X,Y)$$

$$H(U|V) \leq H(U|V,W) + H(W)$$

tip: start from complicated side first

"fundamental lemma"

$$-\sum_i p_i \log p_i \geq -\sum_i p_i \log q_i$$

$$H(X,Y) = H(Y) + H(X|Y)$$

$$H(X|Y) = H(X,Y) - H(Y)$$

$H(X,Y) \geq H(Y)$
$\geq H(X)$

$$I(X|Y) = H(X) - H(X|Y)$$
$$= H(Y) - H(Y|X)$$

$H(X|X) = 0$
$H(X|ZY) \leq H(X|Z)$
$H(X|Y) \leq H(X)$
ind.

$$H(X) + H(Y) \geq H(X,Y)$$
ind.

$$H(U|V) \leq H(U|V,W) + H(W)$$

$$I(X|Y) \geq 0$$

Hamming props:

metric ≥ 0, symmetric, triangle ineq.

Chebyshev

$$Pr(|X - E(X)| \geq a) \leq \frac{var(X)}{a^2}$$

eg. ${}^nC_r$

$$\sum_{r=0}^{\lfloor pn \rfloor} \binom{n}{r} \leq 2^{nh(p)}$$

$$h(p) = -p \log p - (1-p)\log(1-p)$$

Fano $X = Y$; $Z = (X \neq Y)$ "error var"

$$H(X|Y) \leq H(Z) + P(Z=1) \cdot \log(m-1)$$

ramp scheme from $[n,k,d]$ code d^*

$(d^* - s - 1, n - d + 1, n - s)$ inf rate s

$d^* - 1$ cols of parcheck matrix...

Shannon entropy: - cts $\Rightarrow H_{n+1} \geq H_n$

- $H(X|Y) = \leq$ expected $H(X|Y=y)$ over all y

- $0 \leq H \leq \log n$.

Fundamental lemma:

$$-\sum_i p_i \log p_i \geq -\sum_i p_i \log q_i$$

e.g. uniform dist $\Rightarrow H = \log 8 = 3$

$\Rightarrow$ $\frac{1}{8}\ \frac{1}{8}\ \frac{1}{8}\ \frac{1}{8}\ \frac{1}{8}\ \frac{1}{8}\ \frac{1}{8}\ \frac{1}{8}$

$\frac{1}{8}+\epsilon, \frac{1}{8}+2\epsilon, \frac{1}{8}-2\epsilon, \frac{1}{8}-2\epsilon$

useful in proofs when rearranging has $a_i \log b_i$ can then say $\leq H(A)$

Shannon's noiseless — w's are symbols of source alpha.

$$\frac{H(W)}{\log(D)} \leq \bar{n} \leq \frac{H(W)}{\log(D)} + 1$$

ideal observer $\max P(t_i|r_j)$ need a priori $P(t_i)$

$\Updownarrow$ same if uniform prob t_i

max likelihood $\max P(r_j|t_i)$ only need knowledge of channel

= NN in binary sym. channel

Binary symmetric channel bit flip prob $p < \frac{1}{2}$

Hamming distance between two words

weight from $\underline{0}$

min weight = min dist for linear codes

Hamming Codes

$$[2^r - 1,\ n-r,\ 3] \quad \text{perfect} \quad \left[\frac{q^r - 1}{q-1},\ n-r,\ 3\right]$$

($2^r - 1 = n$; $q - 1 = 2 - 1$)

Simplex

$$[2^r - 1,\ r,\ 2^{r-1}] \quad \left[\frac{q^{r-1}}{q-1},\ r,\ q^{r-1}\right]$$

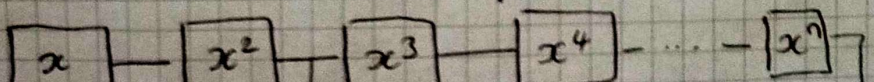
Griesmer ($\frac{q-1}{q-1} = n$)

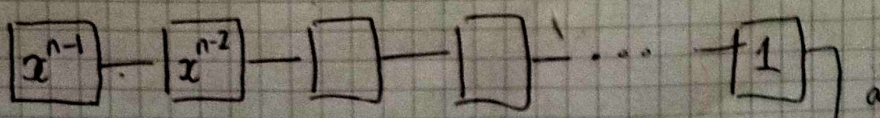
R-S codes

$$[n, k, n-k+1]$$ and dual is $$[n, (n-k), k+1]$$

MDS code.

LFSRs

Galois 

Fib 

always tap at 1.

$x^n = \sum \text{taps} \Rightarrow x^n + \sum \text{taps} = 0$

set up Fib LFSR as gen mat.

m-seq thms
- # of strings of 1s/0s
  r = 1s / no 0s
  r-1 = 0s / no 1s
- autocorrelation $-1/2^{r-1}$
- m-seq: simplex code

Equivalence for block codes

NB: can start with a linear block code & make an equivalent nonlinear code

operations that do not change Hamming dists
⇒ permute columns
⇒ apply permutation in col i. scalar α] to Gen mat

(i.e. to preserve linearity as well)

for linear codes: + row ops; replace perm with $\times\alpha$

Capacity

trade off between code length
words avail.
accuracy / error corr.

$$\sup_{\text{dist}(R)} I(T|R)$$

↳ basically dist(T)

of binary channel: $0 \to 1$
of nth extension $= nC$ $(0 \to n)$

Rate

$$\frac{1}{n}\log_2(M)$$

binary

n = length
M = number of codewords.

measures "redundancy"

Shannon's noisy coding

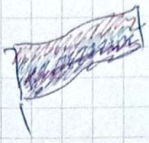
given a channel with capacity C
choose ε_{max} desired error prob ε_{max}
choose rate R | $R < C$

⇒ if you make your codes long enough, can achieve | R and ε_{max} | (surprising result?!)

Thm: $R > C \Rightarrow$ cannot do it.

Systematic : gen $\leftrightarrow$ par

$$(I_k \mid A) \longleftrightarrow (-A^T \mid I_{n-k})$$



relevant to nonbinary codes!

Dual code $[n,k] \leftrightarrow [n, n-k]$

(dual distance : harder)

BOUNDS

$$A_q(n,d) \leq \frac{q^n}{\sum_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i} (q-1)^i}$$

Hamming

perfect

d must be odd

$$\alpha_q(\delta) \leq 1-\delta \qquad \leq q^{n-d+1}$$

R-S codes

MDS

a limit on rate

$$\geq \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i}$$

thm:
$[n,k,d] \to [n,k-1,d']$
$d' \geq d - w + \lceil \frac{w}{q} \rceil$

$$n \geq \sum_{i=}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil$$

Griesmer

Simplex

puncture on i : delete col. i

on support: reduce word to $\underline{0}$

Watermark

- hard to remove
- doesn't affect user's perception but removing it does.

} assume we can.

block marking — only need q watermarks for n^q distinct versions

⇓

raises piracy problems by combining blocks

⇓

security → against trace pirates → ID copy blocks came from

→ prevent framing (weaker)

desc(S) all combos

interested in M_q & biggest code with these props.

C-TA 10 min 1 pirate

nearest neighbour decoding is a pirate (of anything in desc)

C-FP if $\underline{w} \in$ desc is a codeword, it's in S.

Results

$q \leq c \Rightarrow M_q \leq c$

C-TA: $q^{\lceil \frac{n}{c^2} \rceil} \leq M_q \leq q^{\lceil \frac{n}{c} \rceil} + 2c - 2$

(n,d) code with $d > n - \lceil \frac{n}{c^2} \rceil$ eg $[n, \lceil \frac{n}{c^2} \rceil, n - \lceil \frac{n}{c^2} \rceil + 1]$ RS

C-FP: $q^{\lceil \frac{n}{c} \rceil} \leq M_q \leq c q^{\lceil \frac{n}{c} \rceil}$ ← see notes, fancy max if $c \nmid n$.

eg $[n, \lceil \frac{n}{c} \rceil, n - \lceil \frac{n}{c} \rceil + 1]$ RS

USEFUL THMs

Ex U1#2.6. no. of prim. elts of $GF(q = p^m)$ is number of things coprime to $q-1$: no. of prim polys is $\frac{\text{prim elts}}{m}$

Ex U1#2.7. Char $p \implies (a+b)^p = a^p + b^p$

Ex U1#2.8. $GF(q^n)$ is a vector space over $GF(q)$ (any prime power q)

Thm U2#2.7. Shannon entropy Any function satisfying Shannon entropy properties (cts. $H_{n-1} \geq H_n$: $0 \leq H \leq \log n$: $\mathrm{H}(\mathbf{X}, \mathbf{Y}) = \mathrm{H}(\mathbf{X}) + \mathrm{H}(\mathbf{X}|\mathbf{Y})$) is a constant multiple of Shannon entropy

Thm U3#2.15. (Shannon's Noiseless Coding)

$$\frac{\mathrm{H}(\mathbf{W})}{\log D} \underbrace{\leq}_{\text{Any u.d}} \bar{n} \underbrace{\leq}_{\text{u.d. must exist}} \frac{\mathrm{H}(\mathbf{W})}{\log D} + 1$$

Ex U4#2.11. Hamming dist is a **metric**

error prob

$|X| = |Y|$

Fano $\| \quad H(X|Y) \leq H(Z) + \Pr(Z=1) \log(m-1)$

Thm U5#1.1. For the binary symmetric channel. NN decoding is equivalent to max. likelihood decoding

Thm U5#1.4. NN decoding of a block code with min. dist d can correct up to $\lfloor \frac{d-1}{2} \rfloor$ errors

Thm U5#2.2. Capacity of binary symmetric channel with error prob p:

$$1 + p \log p + (1-p) \log (1-p)$$

Page U5#5. If channel has cap. C. its nth extension has cap nC

Ex U9#1.11. $\mathrm{H}(\mathbf{K}) \leq H(\mathbf{S}_i)$

ramp sch. const.

Thm U9#1.19. Given $[n, m, d]$ code with dual dist. d^*, then for any $1 \leq s \leq d^* - 2$, there exists a $(t_1, t_2, n-s)$ ramp scheme with $t_1 = d^* - s - 1$ and $t_2 = n - d + 1$

Ex U9#1.21. Define average information rate $= \frac{n \log_2(|K|)}{\log_2(|S|)}$; for any (t_1, t_2, n) ramp scheme this rate is at most $t_2 - t_1$

Ex U10#1.13. for any $c \geq n$, the set of elements of $\{0, 1, \ldots, q-1\}^n$ with exactly one non-zero component is a c-FP code with $n(q-1)$ elements and is (see Thm 1.12) the largest possible c-FP code with these parameters

$(k-1, k, n)$ ramp $=$ $(k; n)$ threshold

Probability: d.r.v. → distribution. joint. conditional. Bayes

$H(X,Y) = H(X|Y) + H(Y)$

entropy: def. properties. conditional. joint etc.:

(fundamental lemma) *from U5: Chebyshev. Fano.* ${}^nC_r$ *lemma*

mutual information

Channel encoding: instantaneous. u.d.. ($\Longrightarrow$ compact) noiseless. noisy ⇓ (Shannon thms)

Huffman ← *existence. inequalities*

→ (noisy) decoding: ideal obs vs max likelihood

capacity. rate. ...

as channel extension

Block codes: Hamming dist. min. dist. error corr.. *puncturing. support. residual code*

⇓

Linear codes: gen/par mats $\Longrightarrow$ properties

$\Longleftrightarrow$ dual codes —— orthogonal fun

☆ Syndrome decoding ☆

$\rightsquigarrow$ Use of vector space features =

sizes. lin. independence. combo properties etc.

= over a vector space
= over a field:
see GF content
& polys in GFs

Bounds

(egs are linear but bounds are not specific to linear codes)

$A_q \leq \ldots$
$\geq \ldots$
$\geq \ldots$

$n \geq$

egs = Hamming, Simplex, R-S

as

☆ LFSRs → as linear codes

$\rightsquigarrow$ *properties*

Security: definitions. Kerckhoff principle. basic ciphers

Ciphers // encryption — Perfect secrecy

MDS/linear error.corr codes as:
- Secret-sharing — threshold schemes. ramp schemes
- Piracy — c-TA. c-FP