

Reversible codes

A code $\mathcal{C}$ is reversible if $(c_0, c_1, \dots, c_{n-2}, c_{n-1}) \in \mathcal{C}$ implies $(c_{n-1}, c_{n-2}, \dots, c_1, c_0) \in \mathcal{C}$.

Example: $\{000, 110, 101, 011\}$ is a reversible code.

Example: $[15, 6, 6]$ binary BCH code of length 15 with

$g(x) = M^{(-1)}(x)M^{(0)}(x)M^{(1)}(x)$ is reversible.

Exercise

Show that a cyclic code is reversible if the reciprocal of every zero of $g(x)$ is also a zero of $g(x)$.

$\alpha \rightarrow \text{zero of } g(x)$
if
 $\frac{1}{\alpha} \rightarrow \text{zero of } g(x)$

Reed-Solomon Codes (RS)

- An RS code over $GF(q)$ is a BCH code of length $N = q-1$.
- length $N = \#$ of non-zeros in the ground field $GF(q)$, $N = |GF(q) \setminus \{0\}|$
- $N, K, D \rightarrow$ length, dimension, min. dist.

Since $x^{q-1} - 1 = \prod (\alpha - \beta)$ the minimal polynomial

of α^i is simply $M^{(i)}(x) = (x - \alpha^i)$

Therefore, an RS code of length $q-1$ & designated dist. d has generator poly. (for some integer $b \geq 0$)

$$g(x) = (x - \alpha^b)(x - \alpha^{b+1}) \dots (x - \alpha^{b+d-2})$$

Usually, but not always, $b=1$.

$a(x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1} \in \mathcal{C}$, $g(x) = 0$
 $\Rightarrow a(x) = 0 \quad (a_0, a_1, \dots, a_{n-1})$
 $a(\frac{1}{\alpha}) = \frac{1}{\alpha^{n-1}} (a_{n-1} + \frac{a_{n-2}}{\alpha} + \dots + \frac{a_0}{\alpha^{n-1}})$

Summary of the properties of RS codes

number of elements in $GF(q)$ is $q^n = q^{m-1}$.

- An RS code of length $N=q-1$ over $GF(q)$ is a cyclic code with generator polynomial $g(x) = (x - \alpha^b)(x - \alpha^{b+1}) \dots (x - \alpha^{b+\delta-2})$ with designed dist. δ .

Where α is a primitive element of $GF(q)$ & b, δ is some +ve integer.

The dimension $K = N - \delta + 1$.

The min. dist. $D = \delta$

$D = N - K + 1$

- Often $b=1$.

This is a BCH code, & is MDS.

It may be extended to $[q+1, K, q-K+2]$, and (if $q=2^m$) $[2^m+2, 3, 2^m]$ and $[2^m+2, 2^{m-1}, 4]$ codes.

~~The idempotent is~~

Example: $GF(4) = \{0, 1, \alpha, \beta = \alpha^2\}$ with $\alpha^2 + \alpha + 1 = 0$.

An RS code over $GF(4)$ has $N=q-1=3$, designated

distance $\delta=2$ and $b=2$ has $D=\delta=2, K=N-\delta+1=2$

$[N=q-1=3, K=N-\delta+1=2, D=2]$ code with generator polynomial $g(x) = x^2 - \beta$.

All codewords ($4^2=16$ codewords)

$G_{2 \times 3} = \begin{bmatrix} \beta & 1 & 0 \\ 0 & \beta & 1 \end{bmatrix}$

000	000	000
001	001	001
010	010	010
011	011	011
100	100	100
101	101	101
110	110	110
111	111	111
000	000	000
001	001	001
010	010	010
011	011	011
100	100	100
101	101	101
110	110	110
111	111	111

$[3, 2, 2]$ RS code over $GF(4) = \{0, 1, \alpha, \beta = \alpha^2\}$.

$(q=4)$

Example: $GF(4) = \{0, 1, \alpha, \beta = \alpha^2\}$ with $\alpha^2 + \alpha + 1 = 0$.

An RS code over $GF(4)$ with designated distance $\delta = 2$, and $b = 2$ is an

$[N = q - 1 = 4 - 1 = 3, K = N - \delta + 1 = 3 - 2 + 1 = 2, D = \delta = 2]$ code with generator polynomial $g(x) = (x - \beta)(x - \alpha)$

i.e. the generator matrix $G = \begin{pmatrix} \beta & 1 & 0 \\ 0 & \beta & 1 \end{pmatrix}$. $K \times N = 2 \times 3$. $\Delta \beta = d(1+1) = \alpha^2 + 1 = 1$

The 4^2 Codewords are

000	01 α	$\beta\beta\beta$
$\beta 1 0$	$\alpha 0 1$	$\alpha \alpha \alpha$
0 $\beta 1$	1 $\alpha 0$	
10 β	111	
$\beta \alpha 1$	$\alpha \beta 0$	
1 $\beta \alpha$	0 $\alpha \beta$	
$\alpha 1 \beta$	$\beta 0 \alpha$	

$u \times \text{row}_1 + v \times \text{row}_2$
 $u, v \in \{0, 1, \alpha, \beta\}$
 $u=1, v=1 \rightarrow \beta, 1+\beta, 1$
 $\beta, \alpha, 1$
 $u=0, v=\alpha \rightarrow 0, \alpha\beta, \alpha$
 $0, 1, \alpha$
 $u=\alpha, v=1 \rightarrow \alpha\beta, \alpha+\beta, 1$
 $1, 1, 1$
 $u=\alpha^2, v=0 \rightarrow \alpha^2\beta, \alpha^2, 0$
 $\alpha, \beta, 0$
 $u=1, v=\alpha^2 \rightarrow \beta, 1+\alpha^2\beta, \alpha^2$
 ~~β, β, β~~
 $\beta, 1+\alpha, \alpha^2$
 β, β, β
 $u=\alpha^2, v=\alpha \rightarrow \alpha^2\beta, \alpha^2+\alpha\beta, \alpha$
 $\alpha, \alpha^2+1, \alpha$
 α, α, α

Example: the RS code over $GF(5)$ of designated distance $\delta = 3$ is a

$[N = q - 1 = 4, K = N - \delta + 1 = 4 - 3 + 1 = 2, D = 3]$ code. $b = 2$

Take $\alpha = 2$ as the primitive element of $GF(5)$, so that $(x - \alpha)(x - \alpha^2) = (x - 2)(x - 4) = x^2 + 4x + 3$

$g(x) = (x - \alpha)(x - \alpha^2) = (x - 2)(x - 4) = x^2 + 4x + 3$

Some of the $q^k = 5^2 = 25$ Codewords are

3410,	0341,	3201,	2140,
3×3	4×3	1×3	0
4	2	3	0
3×4	4×4	1×4	0
2	1	4	0

$G = \begin{pmatrix} 3 & 4 & 1 & 0 \\ 0 & 3 & 4 & 1 \end{pmatrix}$
 2×4
 $\begin{pmatrix} 3 & 7 & 5 & 1 \\ 3 & 2 & 0 & 1 \end{pmatrix}$

$$\text{Dimension } K = N - b + 1 = 2 \\ = N - \deg g(x)$$

$$4 - 2 = 2$$

By the BCH bound, min dist. $D \geq \delta = N - K + 1$

But by Singleton bound $D \leq N - K + 1$.

$\therefore D = N - K + 1 \Rightarrow$ This is MDS code

Importance of RS Codes

- They are natural codes to use when a code is required of length less than the size of the field. For being MDS, they have the highest possible min. dist.

- They are convenient for building other codes. For example, they can be ~~not~~ mapped into binary codes with surprisingly high min. dist. They are also used in constructing concatenated and Justesen codes

- They are useful for correcting bursts of errors.

Extended RS code (Adding an overall parity check over $GF(q^m)$)

Theorem $\mathcal{C} \rightarrow [N = q^m - 1, K, D]$ RS code, with generator poly.

$$g(x) = (x - \alpha)(x - \alpha^2) \cdots (x - \alpha^{D-1}).$$

Then extending each codeword $c = c_0 c_1 \cdots c_{N-1}$ of $\mathcal{C}$ by adding an overall parity check $c_N = -\sum_{i=0}^{N-1} c_i$

produces an $[N+1, K, D+1]$ code.

Proof. Suppose c has weight D , $c \in \mathcal{C}$.
 (The min. wt. is increased to $D+1$ provided

$$C(1) = -C_N = \sum_{i=0}^{N-1} C_i \neq 0$$

$$C(x) = C_0 + C_1x + \dots + C_{N-1}x^{N-1}$$

But $C(x) = a(x)g(x)$ for some $a(x)$

$$C_N = -\sum_{i=0}^{N-1} C_i$$

$$\therefore C(1) = a(1)g(1)$$

Now $g(1) \neq 0$. ~~$\alpha \neq 1$ as α generates $GF(q^m)$~~
 ~~$\alpha \neq 1$ as α generates $GF(q^m)$~~

furthermore, $a(1) \neq 0$ or else $C(x)$ has a factor $(x-1)g(x)$ & ~~if the code $\mathcal{C}$ becomes~~

So by the BCH bound min. dist. $> D+1$.
 So $C(1) \neq 0$ & hence min. dist. $= D+1$

Example:

$[3, 2, 2]$ RS

encode $\beta 10 \rightarrow \beta 10\alpha$

$1\alpha 0 \rightarrow 1\alpha 0\beta$

$\beta 0\alpha \rightarrow \beta 0\alpha 1$

$\beta\alpha 1 \rightarrow \beta\alpha 1 0$

extended β
 $[4, 2, 3]$ RS Code

$$\begin{aligned} \alpha^2 + 1 + \alpha &= 0 \\ C_N &= -\sum_{i=1}^{N-1} C_i \\ &= C(\beta+1) = -\alpha \\ &= -(1+\alpha) = \beta \\ &= -(\beta+\alpha) = -1 \\ &= -(\beta+\alpha+1) = 0 \end{aligned}$$

Mapping $GF(2^m)$ codes into binary codes

- $GF(q)$, $q = p^m \rightarrow m$ -tuples from $GF(p)$
 $\rightarrow$ a vector space of dim. m over $GF(p)$.
- $[N, k, D]$ RS code over $GF(q) \rightarrow [n = mN, k = mk, d \geq D]$ code over $GF(p)$.
- $q = 2^m \rightarrow$ binary codes obtained in this way often have high min. dist.

- $\xi_1, \xi_2, \dots, \xi_m \rightarrow$ basis of $GF(2^m)$ over $GF(2)$
 $(1, \alpha, \alpha^2, \dots, \alpha^{m-1})$ is a basis when α is a zero of primitive irreducible poly. $\pi(\alpha)$ used to generate $GF(p^m)$.
- if $\beta = \sum_{i=1}^m b_i \xi_i \in GF(2^m)$, $b_i \in GF(2)$, the mapping is

$$\beta \rightarrow \beta_1 \beta_2 \dots \beta_m \text{ from } b_1, b_2, \dots, b_m$$

- (This mapping sends linear codes to linear codes (but cyclic codes need not go into cyclic codes).)

Example:

$1, \alpha \rightarrow$ basis of $GF(4) = \{0, 1, \alpha, \beta = \alpha^2\}$ over $GF(2)$.

mapping

$$0 \rightarrow 00$$

$$1 \rightarrow 10$$

$$\alpha \rightarrow 01$$

$$\beta \rightarrow 11$$

$$\alpha^2 + \alpha + 1 = 0$$

$$\beta = \alpha^2 = \alpha + 1 \Rightarrow 11$$

N, k, D
 $[3, 2, 2]$ RS

over $GF(4)$

$\rightarrow$

$[6, 4, 2]$

binary code

over $GF(2)$

000

$\rightarrow$

000 000

$\beta 10$

$\rightarrow$

111 000

$1\alpha 0$

$\rightarrow$

100 100

$\beta 0 \alpha$

$\rightarrow$

110 001

$\beta \alpha 1$

$\rightarrow$

110 110

$\beta 10 \rightarrow 11 10 00$
 $\rightarrow 110 101 00$

Example:

Let $c = (c_0, c_1, \dots, c_{N-1})$ belong to an $[N, k, D]$ RS code over $GF(2^m)$.

- Replace each c_i by the corresponding binary m -tuple, and add an overall parity check on each m -tuple. This yields an

$[n = (m+1)(2^m-1), k = mk, d \geq 2D = 2(2^m - k)]$ code
 for any $k = 1, \dots, 2^m - 2$.

- The same construction applied to the extended RS code gives an

$[n' = (m+1)(2^m), k' = mk, d' \geq 2(D+1) = 2(2^m - k + 1)]$ binary code,
 for any $k = 1, \dots, 2^m - 1$.

- $[15, 10, 6]$ RS code over $GF(2^4)$ $\rightarrow [75, 40, 12]$ code over $GF(2)$
- $[16, 10, 7]$ extended RS code over $GF(2^4)$ $\rightarrow [80, 40, 14]$ code over $GF(2)$

Example: Using the basis $1, \alpha, \alpha^6$ for $GF(2^3)$ over $GF(2)$, the mapping is

$0 \rightarrow 000$	$\alpha^2 \rightarrow 101$	$\alpha^5 \rightarrow 011$
$1 \rightarrow 100$	$\alpha^3 \rightarrow 110$	$\alpha^6 \rightarrow 001$
$\alpha \rightarrow 010$	$\alpha^4 \rightarrow 111$	

$\{0, 1, \alpha, \dots, \alpha^6\}$
 $\alpha^7 = 1$
 $\alpha^3 + \alpha^4 + 1 = 0$ or
 $\alpha^3 + \alpha + 1 = 0$

Consider the $[7, 5, 3]$ RS code over $GF(2^3)$ with generator polynomial $g(x) = (x + \alpha^5)(x + \alpha^6)$.

$$= x^4 + \alpha x + x^2 + 0x^3 + \dots + 0x^6$$

$\alpha^5 + \alpha^6 = \alpha$
 $\alpha^5 \alpha^6 = \alpha^4 = \alpha^7 = 1$

This is mapped onto $[21, 15, 3]$ binary BCH code with generator polynomial $g_2(y) = M^{(1)}(y) = 1 + y + y^2 + y^4 + y^6$.

$\begin{matrix} N & K & D \\ [7, 5, 3] & m=3 \end{matrix}$
 $\downarrow$
 $\begin{matrix} mN & mK & d_{2D} \\ [21, 15, 3] \end{matrix}$

for $g_1(x)$ itself is mapped onto the vector

111 010 100 000 000 000 000

which is $g_2(x)$.

Also $\alpha g_1(x) \rightarrow y g_2(y)$

$\alpha^2 g_1(x) \rightarrow y^2 g_2(y)$

$\alpha^3 g_1(x) \rightarrow y^3 g_2(y)$, & so on.

(This is the only known, nontrivial, example of a cyclic code mapping in this way onto a cyclic code.)

Justesen codes

- ~~good binary~~ an infinite family of good binary codes ~~is possible to~~ obtained from RS codes.

Defⁿ (good code) A family of codes over $GF(q)$, for q fixed, is said to be good if it contains an infinite ~~family~~ sequence of codes $C_1, C_2, \dots$, where C_i is an $[n_i, k_i, d_i]$ code, such that both the rate $R_i = \frac{k_i}{n_i}$ and $\frac{d_i}{n_i}$ approach a nonzero limit as $i \rightarrow \infty$.

Construction of Justesen codes

- Consider an RS code $\mathcal{R}$ over $GF(2^m)$ with parameters $[N=2^m-1, K, D=N-K+1]$.
- Let α be a primitive element of $GF(2^m)$.
- Let $a = (a_0, a_1, \dots, a_{N-1})$, $a_i \in GF(2^m)$, be a typical codeword of $\mathcal{R}$.
- Let b be a vector $b = (a_0, a_0; a_1, \alpha a_1; a_2, \alpha^2 a_2; \dots; a_{N-1}, \alpha^{N-1} a_{N-1})$.
- finally, replacing each component of b by the corresponding α binary m -tuple, we obtain a binary vector c of length $2mN$.

Defⁿ for any N and K , with $0 < K < N$, the Justesen code $J_{N,K}$ consists of all such vectors c which are obtained from the $[N, K]$ RS code $\mathcal{R}$.

- BCH codes are of great practical importance for error correction, particularly if the expected no. of errors is small compared with the length.
- Double error-correcting BCH codes were constructed as $\rightarrow$ generalization of Hamming code.
- BCH codes are cyclic.
- t -error correcting BCH codes were introduced.
- BCH code over $GF(q)$ of length n & designated distance δ is the largest possible cyclic code having zeros

$$\alpha^b, \alpha^{b+1}, \alpha^{b+2}, \dots, \alpha^{b+\delta-2}$$

where $\alpha \in GF(q^m)$ is a primitive n -th root of unity, b is a non-negative integer & m is the multiplicative order of q modulo n .

- Important special cases are
 - $b=1$ (narrow sense BCH code).
 - $n = q^m - 1$ (primitive BCH code)
- BCH codes with $n = q - 1$ (ie. $m=1, \alpha \in GF(q)$) $\rightarrow$ Reed-Solomon codes
- Bounds on the dimension k and min. dist. d of any BCH code:

Theorem (a) The BCH code over $GF(q)$ of length n and designated distance δ has dimension $k \geq n - m(\delta - 1)$ & min. dist. $d \geq \delta$.

Theorem (b) The binary BCH code of length n & designated distance $\delta = 2t + 1$ has dimension $k \geq n - mt$ & min. dist. $d \geq \delta$.

Goppa code

- cyclic codes are specified in terms of a generator polynomial $g(x)$.
- Goppa codes are described in terms of a Goppa polynomial $G(z)$.
- estimate min. dist. d from $g(x) \rightarrow$ difficult of a cyclic code.
- for Goppa code, $d \geq \deg G(z) + 1$.
- Two ~~things~~^{things} are needed to define a Goppa code
 - (1) Goppa polynomial — A poly. $G(z)$ is called a Goppa poly., having coefficients from $GF(q^m)$, for some fixed m .
 - (2) A subset $L = \{\alpha_1, \dots, \alpha_n\}$ of $GF(q^m)$ s.t.
 $G(\alpha_i) \neq 0 \quad \forall \alpha_i \in L$.
- Usually L is taken to be all the elements of $GF(q^m)$ which are nonzeros of $G(z)$.
- With any vector $a = (a_1, \dots, a_n)$ over $GF(q)$, we associate the rational fun.

$$R_a(z) = \sum_{i=1}^n \frac{a_i}{z - \alpha_i} = \frac{a_1}{z - \alpha_1} + \frac{a_2}{z - \alpha_2} + \dots + \frac{a_n}{z - \alpha_n}$$

Defⁿ. The Goppa code $\Pi(L, G)$ (or Π) consists of all vectors a s.t. $R_a(z) \equiv 0 \pmod{G(z)}$ — (1)
or equivalently, such that $R_a(z) = 0$ in the polynomial ring $GF(q^m)[z] / (G(z))$.
If $G(z)$ is irreducible then Π is called an irreducible Goppa code.

Properties of the Goppa code $\Pi(L, G)$

1. $\Pi(L, G)$ is a linear code over $GF(q)$, defined by eqn. (1).

$$\text{length } n = |L|$$

$$\text{dimension } k \geq n - mr, \quad r = \deg G(z)$$

$$\text{min dist. } d \geq r + 1$$

2. In the binary case, if $G(z)$ has no multiple zeros, then $d \geq 2r + 1$.
- * 3. There exist long Goppa codes which meet the Gilbert-Varshamov bound.
- * 4. Extended binary double-error-correcting Goppa codes are cyclic.

MDS codes

- for a linear code over any field, $d \leq n - k + 1$ (Singleton bound).
- Codes with $d = n - k + 1 \rightarrow$ Maximum Distance Separable (MDS) codes.
 - $\rightarrow$ max. possible distance between codewords; Codewords may be separated into message symbols and check symbols.
- $\rightarrow$ any k symbols can be taken as message symbols
- Problem of finding the longest possible MDS code with a given dimension $=$ to a surprising list of combinatorial problems.
- $[n = q - 1, k, d = n - k + 1]$ RS codes exist over $GF(q)$
 - $\forall k = 1, 2, \dots, n \rightarrow$ MDS codes
- Adding an overall parity check
 - $\rightarrow [n+1, k, n-k+2]$ extended RS codes
 - $\rightarrow$ also MDS codes.

Q: $\rightarrow$ Whether more parity checks can be added, while preserving the property of being MDS.

(Ans. - ~~probably~~ one or two further parity checks can be added, but ~~probably~~ no more.)

Research Problem Given k, q , find largest n for which $[n, k, n-k+1]$ MDS code exists over $GF(q)$.

- In all the known cases, when an $[n, k, d]$ MDS code exists, then an $[n, k, d]$ RS ~~code~~ or ~~extended RS code~~ with the same parameters also exists.
- RS & extended RS codes are the most important class of MDS codes.

Generator and Parity Check matrices

- $C \rightarrow [n, k, d]$ code over $GF(q)$ with parity check matrix H and generator matrix G .

Theorem 1 C is MDS iff ~~not~~ every $n-k$ columns of H are linearly independent.

proof:- C contains a codeword of wt. w .
iff w columns of H are linearly dependent.
 $\hookrightarrow$ minimum d

$\therefore C$ has $d = n - k + 1$ iff no $d-1 = n-k$ columns or fewer columns of H are linearly dependent.

i.e. every C is MDS iff every $n-k$ columns of H are linearly independent.

Theorem 2 If C is MDS so is the dual code $C^\perp$.

proof:- $H_{n-k \times n}$ is the generator matrix of $C^\perp$.
 $n-k$ columns of H

By Theorem 1, every $n-k$ columns of H are linearly independent.

$$i.e. a_1 H_{i_1} + a_2 H_{i_2} + \dots + a_{n-k} H_{i_{n-k}} = 0 \Rightarrow a_1 = a_2 = \dots = a_{n-k} = 0$$

where $H_{i_1}, H_{i_2}, \dots, H_{i_{n-k}} \rightarrow$ columns of H .

Thus since any row of H is a codeword of C ,

if $0 \neq a_1 a_2 \dots a_{n-k} \in C^\perp$ then $a_1 = a_2 = \dots = a_{n-k} = 0$ for some $n-k$ or fewer coordinates in $a_1 a_2 \dots a_n$ are 0.

$\therefore \exists n - (n-k) + 1 = k+1$ or more ^{non-zero} co-ordinates in $a_1 a_2 \dots a_n \neq 0$

$\therefore$ min. dist. of $C^\perp$ is $d \geq k+1$.

$\therefore C^\perp$ is an $[n, n-k, k+1]$ code as $d = n - (n-k) + 1 = k+1$
 C is an MDS code

Example: $\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & \alpha & \beta \end{pmatrix}$ is the generator matrix for a $[4, 2, 3]$ MDS code C over $GF(4) = \{0, 1, \alpha, \beta = \alpha^2\}$.

The dual code $C^\perp$ has generator matrix $\begin{pmatrix} 1 & \alpha & 1 & 0 \\ 1 & \beta & 0 & 1 \end{pmatrix}$ is also a $[4, 2, 3]$ MDS code.

Corollary Let C be an $[n, k, d]$ code over $GF(q)$.

The following statements are equivalent:

- (i) C is MDS;
- (ii) every k columns of a generator matrix G are linearly independent (i.e. any k symbols of the codewords may be taken as message symbols);
- (iii) every $n-k$ columns of a parity check matrix H are linearly independent.

⊙ Proof - Follows from Theorems 1 & 2.

- RS $\rightarrow$ MDS code construction
- Orthogonal Array $\leftrightarrow$ MDS code construction
- n in P.C. $\leftrightarrow$ MDS code.