Codes and Cryptography

Jorge L. Villar

MAMME, Fall 2015

PART VII

Outline

Reed-Solomon Codes Revisited

For the $[q, k, q-k+1]_q$ Reed-Solomon code

$$G = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \alpha & \alpha^2 & \cdots & \alpha^{q-2} \\ 0 & 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2(q-2)} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 1 & \alpha^{k-1} & \alpha^{2(k-1)} & \cdots & \alpha^{(k-1)(q-2)} \end{pmatrix}$$

Puncturing at the first position,

$$G = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha & \alpha^2 & \cdots & \alpha^{q-2} \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2(q-2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha^{k-1} & \alpha^{2(k-1)} & \cdots & \alpha^{(k-1)(q-2)} \end{pmatrix}$$

Property: The rotation map $(x_1, x_2, \ldots, x_n) \mapsto (x_2, \ldots, x_n, x_1)$ preserves the punctured code.

Cyclic Codes

Definition

A code $\mathcal{C} \subset \{0, \ldots, r-1\}^n$ is called cyclic if the rotation map $(x_1, x_2, \ldots, x_n) \mapsto (x_2, \ldots, x_n, x_1)$ preserves it.

Polynomial interpretation: There is a natural bijection $\mathbb{F}_q^n \to \mathbb{F}_q[X]_{\leq n-1}$ that maps each vector $(x_1, \ldots, x_n)$ to the polynomial $x_1 X^{n-1} + \ldots + x_n X^0$.

Rotation corresponds to multiplication by X modulo $X^n - 1$.

Lemma

The linear q-ary cyclic codes of length n are the ideals of the quotient ring $\mathbb{F}_q[X]/(X^n - 1)$.

The trivial, repetition and $[q-1, k, q-k]_q$ Reed-Solomon codes are cyclic.

A Note on Equivalence of Linear Codes (I)

Scaling maps

$$(x_1, x_2, \ldots, x_n) \mapsto (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n) \qquad \lambda_i \neq 0$$

and permutation maps

$$(x_1, x_2, \ldots, x_n) \mapsto (x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}) \qquad \pi \in S_n$$

are isometries with respect to the Hamming distance, and can be considered as **isomorphisms of codes**.

Isomorphisms preserve the main properties of the code: length, size, minimum distance, covering radius... but they do not preserve cyclicity!

E.g., a linear code with $G = (1\,2\,\cdots\,n)$, $n < q$ is isomorphic to the repetition code but it is not cyclic.

A Note on Equivalence of Linear Codes (II)

From the point of view of the generating matrix, one can

- permute the columns of G
- multiply the columns of G by nonzero scalars
- perform "gaussian elimination" operations to the rows of G

without modifying the basic properties (not including cyclicity) of the linear code.

Several differently looking generating matrices essentially defining the same code often appear in the literature (e.g., for the Golay codes).

Generating Polynomial of a Cyclic Code

The quotient ring $\mathbb{F}_q[X]/(X^n - 1)$ is principal (because $\mathbb{F}_q[X]$ is principal). Therefore, any cyclic code is generated by a single (monic) polynomial g.

For a (linear) $[n, k, d]_q$ cyclic code, $\mathcal{C} = (g)$,

- The degree of g is $n - k$
- g has at least d nonzero coefficients
- g divides $X^n - 1$

If $g = X^{n-k} + g_{n-k-1}X^{n-k-1} + \ldots + g_0 X^0$, then

$$G = \begin{pmatrix} 1 & g_{n-k-1} & \cdots & g_0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & g_1 & g_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & g_1 & g_0 \end{pmatrix}$$

Example: a $[8, 5, ?]_{11}$ Cyclic Code

Splitting $X^8 - 1$ into irreducible factors in $\mathbb{F}_{11}[X]$:

$$(X-1)(X+1)(X^2+1)(X^2-3X-1)(X^2+3X-1)$$

Taking $g = (X+1)(X^2+1) = X^3 + X^2 + X + 1$:

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

NO! Not a good code, as $d = 2$. Indeed, $(X-1)g = X^4 - 1 \in (g)$.
Taking now $g = (X+1)(X^2 - 3X - 1) = X^3 - 2X^2 - 4X - 1$:

$$G = \begin{pmatrix} 1 & 9 & 7 & 10 & 0 & 0 & 0 & 0 \\ 0 & 1 & 9 & 7 & 10 & 0 & 0 & 0 \\ 0 & 0 & 1 & 9 & 7 & 10 & 0 & 0 \\ 0 & 0 & 0 & 1 & 9 & 7 & 10 & 0 \\ 0 & 0 & 0 & 0 & 1 & 9 & 7 & 10 \end{pmatrix}$$

YES! It is a $[8, 5, 4]_{11}$ cyclic MDS code!

A Note About Factoring $X^n - 1$ in $\mathbb{F}_q$ (I)

The roots of $X^n - 1$ are exactly the n n-th roots of unity in $\mathbb{F}_q$. But not all of them are contained in $\mathbb{F}_q$, but in some finite extension $\mathbb{F}_{q^e}$. They actually form a cyclic subgroup of $\mathbb{F}_{q^e}^\times$.

The extension degree e fulfils $X^n - 1 \mid X^{q^e-1} - 1$, that is, $n \mid q^e - 1$ or equivalently $q^e = 1 \bmod n$.

If $\alpha \in \mathbb{F}_{q^e}$ is primitive then $\beta = \alpha^{\frac{q^e-1}{n}}$ is a **primitive n-th root of unity**, which generates all the other n-th roots of unity.

Each power of β along with its conjugates in $\mathbb{F}_q$ defines each of the irreducible factors of $X^n - 1$ in $\mathbb{F}_q$.

Observe that the conjugates of β^i are $\beta^{iq}, \beta^{iq^2}, \ldots$, where the exponents can be computed modulo n.

A Note About Factoring $X^n - 1$ in $\mathbb{F}_q$ (II)

In the previous example $e = 2$, since $11^2 - 1 = 15 \cdot 8$. Then $\beta = \alpha^{15}$.

The only conjugate of β is $\beta^{11} = \beta^3$ (as $11 = 3 \bmod 8$) since $\beta^{11^2} = \beta$. Similarly, the conjugate of β^2 is β^6, the conjugate of β^5 is β^7, and β^4 and 1 have no conjugates.

The irreducible factors of $X^8 - 1$ in $\mathbb{F}_{11}$ are then

$$X - 1$$
$$X - \beta^4 = X + 1$$
$$(X - \beta)(X - \beta^3) = X^2 + 3X - 1$$
$$(X - \beta^2)(X - \beta^6) = X^2 + 1$$
$$(X - \beta^5)(X - \beta^7) = X^2 - 3X - 1$$

The last step requires building $\mathbb{F}_{11^2} = \mathbb{F}_{11}[i]/(i^2 + 1)$, then $\alpha = 4 + i$ is primitive and $\beta = 4(1 - i)$ is a primitive eight root of unity.

The Dual of a Cyclic Code

Given a $[n, k, d]_q$ cyclic code $\mathcal{C} = (g)$, define $h \in \mathbb{F}_q[X]$ such that $X^n - 1 = gh$, and $\overleftarrow{h} \in \mathbb{F}_q[X]$ such that $\overleftarrow{h}(X) = X^{n-1}h(\frac{1}{X})$. (The order of the coefficients is reversed.)

Lemma

$\mathcal{C}^\perp = (g^\perp)$ *where* $g^\perp = \frac{\overleftarrow{h}}{h(0)X^{n-k-1}} = -\frac{X^k}{g(0)}h(\frac{1}{X})$

For a polynomial $a = a_{n-1}X^{n-1} + \ldots + a_0X^0$, define $coef(a) = (a_{n-1}, \ldots, a_0)$, (i.e., the associated codeword). Thus, $coef(a) \cdot coef(b) = coef_{n-1}\left(a\overleftarrow{b}\right)$, where $coef_i(a) = a_i$, and $coef(X^i g) \cdot coef(X^j g^\perp) = coef_{n-1}(X^{n-1+i-j}gg^\perp(\frac{1}{X})) = -coef_{n-1}\left(\frac{X^{n-k-1+i-j}}{g(0)}gh\right) = -coef_{n-1}\left(\frac{X^{n-k-1+i-j}}{g(0)}(X^n - 1)\right) = 0$ since $0 \le n - k - 1 + i - j \le n - 2$.

Dual of Previous Example: a $[8, 3, 6]_{11}$ Cyclic Code

$X^8 - 1 = (X - 1)(X + 1)(X^2 + 1)(X^2 - 3X - 1)(X^2 + 3X - 1)$
$g = (X + 1)(X^2 - 3X - 1) = X^3 - 2X^2 - 4X - 1$

$$G = \begin{pmatrix} 1 & 9 & 7 & 10 & 0 & 0 & 0 & 0 \\ 0 & 1 & 9 & 7 & 10 & 0 & 0 & 0 \\ 0 & 0 & 1 & 9 & 7 & 10 & 0 & 0 \\ 0 & 0 & 0 & 1 & 9 & 7 & 10 & 0 \\ 0 & 0 & 0 & 0 & 1 & 9 & 7 & 10 \end{pmatrix}$$

$h = (X - 1)(X^2 + 1)(X^2 + 3X - 1) = X^5 + 2X^4 - 3X^3 + 3X^2 - 4X + 1$
$g^\perp = X^5 - 4X^4 + 3X^3 - 3X^2 + 2X + 1$

$$G^\perp = \begin{pmatrix} 1 & 7 & 3 & 8 & 2 & 1 & 0 & 0 \\ 0 & 1 & 7 & 3 & 8 & 2 & 1 & 0 \\ 0 & 0 & 1 & 7 & 3 & 8 & 2 & 1 \end{pmatrix}$$

It is a $[8, 3, 6]_{11}$ cyclic MDS code!

Encoding as a Systematic Code

Given a cyclic code $\mathcal{C} = (g)$ we can encode messages as in a systematic code by means of polynomial division.

A source message $(m_1, \ldots, m_k)$ is written as the polynomial $m = m_1 X^{n-1} + \ldots m_k X^{n-k}$.

Now dividing by g we obtain $m = sg + r$ where $\deg r \leq n - k - 1$. Therefore, $c = m - r \in (g)$.

The final codeword is $(m_1, \ldots, m_k, -r_{n-k-1}, \ldots, -r_0)$, where $r = r_{n-k+1} X^{n-k+1} + \ldots + r_0 X^0$.

Thus, we only need to store the polynomial ($n - k + 1$ coefficients) and not the whole generating matrix (nk elements).

Syndromes of a Cyclic Code

The cyclicity of the code can be used to save memory to store the syndromes.

Actually, the syndrome of $y \in \mathbb{F}_q[X]_{\leq n-1}$ can be computed as the remainder $s(y) = y \bmod g$ in $\mathbb{F}_q[X]$.

Thus, we do not need to store the whole parity check matrix.

Moreover, we can reduce the number of stored syndromes by computing not only $s(y)$ but also $s(X^i y)$ for $i = 1, \ldots, n-1$.

Then we only need to store the syndromes of

- 1 for $t \leq 1$
- $X + 1, X^2 + 1, \ldots, X^{n-1} + 1$ for $t \leq 2$
- $X^2 + X + 1, X^3 + X + 1, \ldots, X^{n-1} + X^{n-2} + 1$ for $t \leq 3$
- $\ldots$

Roots of a Cyclic Code

The roots of the generating polynomial g of a cyclic code $\mathcal{C} = (g)$ are also roots of all codewords in $\mathcal{C}$.

Since $g = \prod_{i=1}^{n-k} (X - \xi_i)$ for $\xi_1, \ldots, \xi_{n-k}$ in some extension field, the syndrome of a polynomial $y \in \mathbb{F}_q[X]_{\leq n-1}$ can be redefined as $s(y) = (y(\xi_1), \ldots, y(\xi_{n-k}))$.

The **syndrome polynomial** for y is $s_y = \sum_{i=1}^{n-k} y(\xi_i) X^i$.

E.g., for the Reed-Solomon cyclic code, $\xi_i = \alpha^{-i}$ and then $s(y) = (y(\alpha^{-1}), \ldots, y(\alpha^{-n+k}))$, or $s_y = y(\alpha^{-1})X + \ldots + y(\alpha^{-n+k})X^{n-k}$.

Other Reed-Solomon Cyclic Codes (I)

The cyclic $[10, 3, 8]_{11}$ Reed-Solomon code ($\alpha = 2 \in \mathbb{F}_{11}$):

$$G = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 5 & 10 & 9 & 7 & 3 & 6 \\ 1 & 4 & 5 & 9 & 3 & 1 & 4 & 5 & 9 & 3 \end{pmatrix}$$

with generating polynomial

$$g = X^7 + 7X^6 + 2X^5 + X^4 + 2X^3 + 5X^2 + 4X + 7 = \prod_{i=1}^{7} (X - \alpha^{-i})$$

The red columns form the cyclic $[5, 3, 3]_{11}$ shortened code

$$G_{short} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 9 & 3 \\ 1 & 5 & 3 & 4 & 9 \end{pmatrix}$$

with generating polynomial $g_{short} = (X - \alpha^{-2})(X - \alpha^{-4})$

Other Reed-Solomon Cyclic Codes (II)

More generally, let $\beta = \alpha^m$ where $m \mid q-1$ and $\alpha \in \mathbb{F}_q$ is a primitive element. We can define a $[n, k, n-k+1]_q$ cyclic MDS code, for $n = \frac{q-1}{m}$ by means of the generating matrix

$$G = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \beta & \beta^2 & \cdots & \beta^{n-1} \\ 1 & \beta^2 & \beta^4 & \cdots & \beta^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \beta^{k-1} & \beta^{2(k-1)} & \cdots & \beta^{(k-1)(n-1)} \end{pmatrix}$$

which is a shortened Reed-Solomon code.

Observe that $\beta^n = 1$, that is, β is a root of $X^n - 1$. Actually, the n roots of $X^n - 1$ are exactly β^i for $i = 0, \ldots, n-1$. The generating polynomial of the code is now $g = \prod_{i=1}^{n-k}(X - \beta^{-i})$

Correcting Errors With The Syndrome Polynomial (I)

Let $\mathcal{C}$ be a $[n, k, n-k+1]_q$ shortened Reed-Solomon cyclic code, with $n = \frac{q-1}{m}$ and $\beta = \alpha^m$, for some $m \mid q-1$.

Assume that a corrupted codeword $y \in \mathbb{F}_q[X]_{\leq n-1}$ contains exactly t errors, for some $t \leq \lfloor \frac{n-k+1}{2} \rfloor$.

The **error locator polynomial** is defined as

$$E_y = \prod_{j=1}^{t}(\beta^{-z_j}X - 1)$$

where $z_1, \ldots, z_t \in \{0, \ldots, n-1\}$ are the error positions (being 0 the rightmost one).

Let e_j be the error occurred at position z_j, that is

$$y - \sum_{j=1}^{t} e_j X^{z_j} \in \mathcal{C}$$

Correcting Errors With The Syndrome Polynomial (II)

Recall that $s_y = \sum_{i=1}^{n-k} y(\beta^{-i})X^i$ and $y(\beta^{-i}) = \sum_{j=1}^{t} e_j \beta^{-iz_j}$.

Now, $s_y = \sum_{j=1}^{t} e_j \sum_{i=1}^{n-k}(\beta^{-z_j}X)^i = \sum_{j=1}^{t} e_j \beta^{-z_j} X \frac{(\beta^{-z_j}X)^{n-k} - 1}{\beta^{-z_j}X - 1}$

and $E_y s_y = \sum_{j=1}^{t} e_j \beta^{-z_j} X \left((\beta^{-z_j}X)^{n-k} - 1\right) \prod_{\ell=1,\, \ell \neq j}^{t} (\beta^{-z_\ell}X - 1)$

which coefficients of degree $t+1, \ldots, 2t$ are zero. This results in a homogeneous linear system of t equations (independent of $e_1, \ldots, e_t$) that determines the error locator polynomial.

Berlekamp-Massey algorithm efficiently solves this step.

Correcting Errors With The Syndrome Polynomial (III)

Once E_y is known, it is factorized and $e_1, \ldots, e_t$ are computed with the formula

$$e_j = -\frac{\widehat{(E_y s_y)}(\beta^{z_j})}{\beta^{z_j} E_y'(\beta^{z_j})}$$

$\widehat{E_y s_y}$ contains only the monomials of $E_y s_y$ of degree $\leq n-k$. Indeed,

$$\widehat{(E_y s_y)}(\beta^{z_j}) = -e_j \prod_{\ell=1,\, \ell \neq j}^{t} (\beta^{z_j - z_\ell} - 1)$$

while

$$E_y'(\beta^{z_j}) = \beta^{-z_j} \prod_{\ell=1,\, \ell \neq j}^{t} (\beta^{z_j - z_\ell} - 1)$$

Example: Correcting Errors For $[10, 3, 8]_{11}$ Code

Primitive element: $\alpha = 2 \in \mathbb{F}_{11}$
Generating polynomial: $g = X^7 + 7X^6 + 2X^5 + X^4 + 2X^3 + 5X^2 + 4X + 7$
Codeword: $x = (2, 8, 9, 0, 4, 1, 6, 5, 3, 10)$
Error Vector: $e = (8, 0, 0, 5, 0, 3, 0, 0, 0, 0)$
Corrupted Codeword: $y = (10, 8, 9, 5, 4, 4, 6, 5, 3, 10)$
Syndrome Vector: $s = (2, 4, 5, 1, 0, 3, 6)$
Syndrome Polynomial: $s_y = 2X + 4X^2 + 5X^3 + X^4 + 3X^6 + 6X^7$

Trying $t = 1$: Error Locator Polynomial: $E_y = \lambda_0 + \lambda_1 X$

$$\text{rank} \begin{pmatrix} 6 & 3 & 0 & 1 & 5 & 4 \\ 3 & 0 & 1 & 5 & 4 & 2 \end{pmatrix} < 2? \text{ NO!}$$

Example: Correcting Errors For $[10, 3, 8]_{11}$ Code

Primitive element: $\alpha = 2 \in \mathbb{F}_{11}$
Generating polynomial: $g = X^7 + 7X^6 + 2X^5 + X^4 + 2X^3 + 5X^2 + 4X + 7$
Codeword: $x = (2, 8, 9, 0, 4, 1, 6, 5, 3, 10)$
Error Vector: $e = (8, 0, 0, 5, 0, 3, 0, 0, 0, 0)$
Corrupted Codeword: $y = (10, 8, 9, 5, 4, 4, 6, 5, 3, 10)$
Syndrome Vector: $s = (2, 4, 5, 1, 0, 3, 6)$
Syndrome Polynomial: $s_y = 2X + 4X^2 + 5X^3 + X^4 + 3X^6 + 6X^7$

Trying $t = 2$: Error Locator Polynomial: $E_y = \lambda_0 + \lambda_1 X + \lambda_2 X^2$

$$\text{rank} \begin{pmatrix} 6 & 3 & 0 & 1 & 5 \\ 3 & 0 & 1 & 5 & 4 \\ 0 & 1 & 5 & 4 & 2 \end{pmatrix} < 3? \text{ NO!}$$

Example: Correcting Errors For $[10, 3, 8]_{11}$ Code

Primitive element: $\alpha = 2 \in \mathbb{F}_{11}$
Generating polynomial: $g = X^7 + 7X^6 + 2X^5 + X^4 + 2X^3 + 5X^2 + 4X + 7$
Codeword: $x = (2, 8, 9, 0, 4, 1, 6, 5, 3, 10)$
Error Vector: $e = (8, 0, 0, 5, 0, 3, 0, 0, 0, 0)$
Corrupted Codeword: $y = (10, 8, 9, 5, 4, 4, 6, 5, 3, 10)$
Syndrome Vector: $s = (2, 4, 5, 1, 0, 3, 6)$
Syndrome Polynomial: $s_y = 2X + 4X^2 + 5X^3 + X^4 + 3X^6 + 6X^7$

Trying $t = 3$: Error Locator Polynomial: $E_y = \lambda_0 + \lambda_1 X + \lambda_2 X^2 + \lambda_3 X^3$

$$\text{rank} \begin{pmatrix} 6 & 3 & 0 & 1 \\ 3 & 0 & 1 & 5 \\ 0 & 1 & 5 & 4 \\ 1 & 5 & 4 & 2 \end{pmatrix} < 4? \text{ YES!}$$

Solving $\begin{pmatrix} 6 & 3 & 0 & 1 \\ 3 & 0 & 1 & 5 \\ 0 & 1 & 5 & 4 \\ 1 & 5 & 4 & 2 \end{pmatrix} \begin{pmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$:

Error Locator Polynomial: $E_y = 1 + 6X + 7X^2 + 9X^3$

Example: Correcting Errors For $[10, 3, 8]_{11}$ Code

Primitive element: $\alpha = 2 \in \mathbb{F}_{11}$
Generating polynomial: $g = X^7 + 7X^6 + 2X^5 + X^4 + 2X^3 + 5X^2 + 4X + 7$
Codeword: $x = (2, 8, 9, 0, 4, 1, 6, 5, 3, 10)$
Error Vector: $e = (8, 0, 0, 5, 0, 3, 0, 0, 0, 0)$
Corrupted Codeword: $y = (10, 8, 9, 5, 4, 4, 6, 5, 3, 10)$
Syndrome Vector: $s = (2, 4, 5, 1, 0, 3, 6)$
Syndrome Polynomial: $s_y = 2X + 4X^2 + 5X^3 + X^4 + 3X^6 + 6X^7$

Error Locator Polynomial: $E_y = 1 + 6X + 7X^2 + 9X^3$
Computing the roots of E_y: $\alpha^4, \alpha^6, \alpha^9$. Then $e = (*, 0, 0, *, 0, *, 0, 0, 0, 0)$.
Truncating $E_y s_y$: $\widehat{E_y s_y} = 10X^3 + 5X^2 + 2X$
Computing individual errors:

$$e_4 = -\left.\frac{10X^3+5X^2+2X}{X(5X^2+3X+6)}\right|_{X=\alpha^4=5} = 3$$

$$e_6 = -\left.\frac{10X^3+5X^2+2X}{X(5X^2+3X+6)}\right|_{X=\alpha^6=9} = 5$$

$$e_9 = -\left.\frac{10X^3+5X^2+2X}{X(5X^2+3X+6)}\right|_{X=\alpha^9=6} = 8$$

Recovered Error Vector: $e = (8, 0, 0, 5, 0, 3, 0, 0, 0, 0)$

Codes and Cryptography

Jorge L. Villar

MAMME, Fall 2015

END OF PART VII