Linear Codes

Generator Matrix: Let $\{c_1, \cdots, c_k\}$ be a basis for $[n,k]_q$ code $\mathscr{C}$. Then generator matrix

$$G = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}_{k \times n}$$

Thus this matrix representation is a way to get a code by simply finding k indep vectors, and taking their rowspace.

Here $\text{rank}(G) = \dim(\text{rowspace}(G)) = \dim(\mathscr{C}) = k$.

Reed Solomon Codes

G = Vandermonde $= \begin{bmatrix} 1 & 1 & & 1 \\ \alpha_1 & \alpha_2 & & \alpha_3 \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_3^2 \\ \vdots & \vdots & & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & & \alpha_3^{k-1} \end{bmatrix}_{k \times n}$

$k < n$

If all α_i are distinct matrix is full rank.

Proof: If α_is are distinct, any k columns are lin. indep. It is sufficient to show first k columns are.

Consider $\begin{bmatrix} 1 & & 1 \\ \vdots & \cdots & \vdots \\ \alpha_1^{k-1} & & \alpha_k^{k-1} \end{bmatrix}$. Its determinant $= \prod_{i \neq j} (\alpha_i - \alpha_j)$ (check)

$\Rightarrow$ If every $\alpha_i \neq \alpha_j \Rightarrow$ each term is non zero $\Rightarrow$ prod non zero

$\Rightarrow$ Det $\neq 0$

$\Rightarrow$ vectors are lin. indep.

$\Rightarrow \text{rank}(G) = k. = \text{dimension}(\mathscr{C})$

What is the min distance?

Claim: $d_{min}(\mathcal{C}) = n-k+1$ (singleton bound with equality!)

Proof:

Recall linear combinations of the rows is equ to multiplying the G with a coefficient vector.

i.e. $\mathcal{C} = \{ c \mid c = \underset{1\times k}{m} \underset{k\times k}{G} \}$

Any non zero codeword is given by $m \neq 0$.

$$[m_0 \ m_1 \ \cdots \ m_{k-1}] \begin{bmatrix} 1 & & 1 \\ \alpha_1 & \cdots & \alpha_k \\ \vdots & & \vdots \\ \alpha_1^{k-1} & & \alpha_k^{k-1} \end{bmatrix}$$

$$= \left[\sum_j M_j \alpha_1^j \quad \sum_j M_j \alpha_2^j \quad \cdots \quad \sum_j M_j \alpha_k^j \right]$$

For non zero polynomial.

$\Rightarrow$ A polynomial $m(x) = m_0 + m_1 x + m_2 x^2 + \cdots + m_{k-1} x^{k-1}$ is evaluated at $\alpha_1, \alpha_2, \cdots \alpha_k$.

Now a poly of deg d has $\leq d$ roots. (in $\mathbb{F}_q$)

$\Rightarrow$ $M(X)$ can have atmost $k-1$ distinct roots.

$\Rightarrow$ At most $k-1$ of the α_is can make $m(\alpha_i) = 0$.

$\Rightarrow$ $w_H(c) \geq n - (k-1) \quad \forall c \in \mathcal{C}$

$\Rightarrow$ $\min\limits_{\substack{c \in \mathcal{C} \\ c \neq 0}} w_H(c) \geq n-k+1$.

But by Singleton bound $\Rightarrow d_{min}(\mathcal{C}) \leq n-k+1$

$\Rightarrow$ $w_H(c) = d_{min}(\mathcal{C}) = n-k+1$.

Remark: α_i s are distinct hence $q \geq n$.

Can there be MDS codes with field size $q < n$? Very recently it was shown that you need $q = O(n)$ (or $O(\sqrt{n})$ not sure).

<u>Evaluation based code</u>: code is generated by evaluating a polynomial

polynomials on $\mathbb{F}_q$

$$\mathcal{C}_{RS} = \{ (m(\alpha_1), \dots, m(\alpha_k)) \mid \forall\, m(x) \in \mathbb{F}_q[x], \deg(m(x)) \leq k-1 \}$$

Decoding of RS Codes (Error Locator Polynomials)

Given $[n, k, d_{min}]_q$ RS code $\mathcal{C}_{RS}$, a channel with the worst case error model.

Needed for MHD to work.

Design the code s.t. $d_{min}(\mathcal{C}_{RS}) = n-k+1 = 2t+1$.

Trivial MHD decoder: Run a linear search over all codewords
$\Rightarrow O(|\mathcal{C}|) = O(q^k) \approx O(q^n)$ since $\frac{k}{n}$ is const

Can we do better?

Error Locator Polynomials: A polynomial $E(x)$ s.t. $E(x)\big|_{\alpha_i} = E(\alpha_i) = 0$ iff $y_i \neq m(\alpha_i)$ where y is the received vector, and m is the message polynomial

Suppose we have such an $E(x)$. We can evaluate it at each α_i. If it is zero then that position has an error.

Notice now that

$$y_i E(\alpha_i) = m(\alpha_i) E(\alpha_i), \quad \forall i.$$

since if $y_i \neq m(\alpha_i)$ both sides becomes zero.

Let $N(X) = m(X) E(X) = n_0 + n_1 X + \cdots + n_{deg(N)} X^{deg(N)}$

$E(X) = e_0 + e_1 X + \cdots + e_{deg(E)} X^{deg(E)}$

Now we have equations of the form-

$$y_i \left(e_0 + e_1\alpha_i + e_2\alpha_i^2 + \cdots + e_{deg(E)} \alpha_i^{deg(E)} \right)$$

$$= n_0 + n_1 \alpha_i + \cdots + n_{deg(N)} \alpha_i^{deg(N)} \quad \text{for each } i.$$

Each of these are linear equations in variables $e_0, \cdots, e_{deg(E)}$ and $n_0, \cdots, n_{deg(N)}$.

A solution to these gives us both the polynomials $N(X)$ and $E(X)$ (and hence $M(X)$).

Jan 27

To show: i) There exists a solution to the above system.
ii) If there are multiple solutions, they yield the same message. i.e. $\frac{N_1(x)}{E_1(x)} = m(x) = \frac{N_2(x)}{E_2(x)}$

and $\deg(E) \leq t$ and $\deg(N) \leq (k-1)+t$

it gives 0 wherever $y_i \neq m(\alpha_i)$ so needs t zeroes.

n equations: $y_i E(\alpha_i) = m(\alpha_i) E(\alpha_i) = N(\alpha_i)$ $\forall i \in [n]$.

Let $E(x) = e_0 + e_1 x + \dots + e_{\deg(E)} x^{\deg(E)}$
$N(x) = n_0 + \dots + n_{\deg(N)} x^{\deg(N)}$

Proof: Lets define $E(x)$ such that $E(\alpha_i)=0$ iff $m(\alpha_i) \neq y_i$

Choose $$E(x) = \prod_{i \in \{i \mid m(\alpha_i) \neq y_i\}} (x - \alpha_i)$$

we're just showing that such a polynomial exists. The decoder doesn't know this.

If we now define $N(x) = m(x) \cdot E(x)$ and this is a solution to the system. Notice $\deg(E) = |\{i \mid m(\alpha_i) \neq y_i\}| = t$.

What if there are multiple solutions? Does their ratio give us $m(x)$?

Claim: If $\deg(E(x)) \leq t = \frac{d-1}{2} = \frac{n-k+1-1}{2} = \frac{n-k}{2}$ (with $d = 2t+1$) we have a unique $m(x)$.

(Note $\Rightarrow \deg(N(x)) = \deg(E(x)) + \deg(m) = t+k-1$

Proof: consider $\frac{N_1(x)}{E_1(x)} = \frac{N_2(x)}{E_2(x)}$

$\Rightarrow N_1(x) E_2(x) = N_2(x) E_1(x)$

consider $b(x) = N_1(x) E_2(x) - N_2(x) E_1(x)$

To show: $b(x) = 0$.

$$\mathrm{Deg}(b) \leq (k+t-1) + t$$
$$\leq k+2t-1$$

But $\underset{\substack{\downarrow \\ d(C)\ (\text{singleton})}}{2t+1} = n-k+1$

$\Rightarrow \deg(b(x)) \leq n-1$

Also, $N_1(\alpha_i) = y_i \cdot E_1(\alpha_i)$ since it satisfies the system of equations

Similarly, $N_2(\alpha_i) = y_i\, E_2(\alpha_i)$

$$\Rightarrow b(\alpha_i) = N_1(\alpha_i) E_2(\alpha_i) - N_2(\alpha_i) E_1(\alpha_i)$$
$$= y_i E_1(\alpha_i) E_2(\alpha_i) - y_i E_2(\alpha_i) E_1(\alpha_i)$$
$$= 0 \quad \forall\ i \in [n].$$

But $n > n-1$

$\Rightarrow$ $n-1$ deg polynomial has more that $(n-1)$ distinct roots. $\therefore$ By the fundamental theorem of algebra, $b(x)$ is the zero polynomial.

$\Rightarrow$ The ratio is same for all solutions. Since we found one solution with ratio $= m(x)$, all solutions have ratio $m(x)$.

$\Rightarrow$ We have an $O(n^3)$ + complexity of poly division.

$O(n^3)$ ↓ solve the linear system with Gaussian Elim to find E and N.

complexity of poly division ↓ Find the ratio of E and N.

This is called **Berlekamp - Welch Algorithm**.