Linear codes II

Take q a prime power, let $K = GF(q)$.

Recall a linear code is a linear subspace $C \subseteq K^n$

If C has basis $\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_k$, then C has dimension k, $|C| = q^k$ & say C is an (n,k)-code.

The dual code of C is

$$C^\perp = \{\vec{w} \in K^n \mid \vec{w} \cdot \vec{a} = 0 \quad \forall \vec{a} \in C\}$$

$$= \{\vec{w} \in K^n \mid \vec{w} \cdot \vec{g}_i = 0 \quad i = 1, 2, \ldots, k\}$$

If C is (n,k)-code, $C^\perp$ is $(n, n-k)$-code.

Given C, generator matrix is

$$G = \begin{pmatrix} \vec{g}_1 \\ \vec{g}_2 \\ \vdots \\ \vec{g}_k \end{pmatrix}$$

parity checksum matrix is ($\vec{h}_1, \ldots \vec{h}_{n-k}$ a basis of $C^\perp$)

$$H = \begin{pmatrix} \vec{h}_1 \\ \vec{h}_2 \\ \vdots \\ \vec{h}_{n-k} \end{pmatrix}.$$

Ex Parity code $C \subseteq GF(2)^n$, $G = \begin{pmatrix} 1 & \cdots & 1 \\ & 1 & 0 \\ & & 1 \\ 0 & & 1 \\ & \cdots & 1 \end{pmatrix}$, $H = (1\ 1\ \cdots\ 1)$.

Given linear code C & $d \geq 1$,

Prop The distance $d(C) \geq d \iff$ set of every $d-1$ columns of H are linearly independent.

Ex Parity code: $d = 1$, every 1 column is $(1) \neq 0$, so lin. indep $\rightsquigarrow d(C) \geq 2$

$d = 3$: first two columns are (1) & (1), not lin. indep $\rightsquigarrow d(C) < 3$

$\Rightarrow d(C) = 2$.

Pf. Recall: $d(C) = \min_{0 \neq \vec{a} \in C} w(\vec{a}) = \min_{0 \neq \vec{a} \in C} (\#\text{non-0 entries of } \vec{a})$.

- $C = (C^\perp)^\perp$, so $\vec{a} \in C \iff \vec{a} \cdot \vec{w} = 0 \quad \forall \vec{w} \in C^\perp$

$\iff \vec{a} \cdot \vec{h}_i = 0 \quad i = 1, 2, \ldots, n-k$

$\iff H\vec{a} = 0 \iff \sum_{i=1}^{n} a_i \vec{c}_i = 0,$

where $H = [\vec{c}_1 \ \vec{c}_2 \ \cdots \ \vec{c}_n] \quad \vec{c}_i \in K^{n-k}$.

~~Every linear relation between rows of the $\vec{a}_i$ gives~~

An $\vec{a} \in C$ of weight m gives a linear relation between m of the $\vec{c}_i$ & vice versa, so

$d(C) \geq d \iff \nexists \vec{a} \in C,\ w(a) \leq d-1 \iff \nexists \vec{c}_{i_1}, \ldots, \vec{c}_{i_{d-1}}$ linearly dependent. □

Ex (Parity code). $H = [\vec{c}_1, \ldots, \vec{c}_{n-1}]$, all $\vec{c}_i = (1)$. Lin. relation $\vec{c}_1 + \vec{c}_2 = 0$ gives $\vec{a} = [1 \ 1 \ 0 \ 0 \ \cdots \ 0] \in C$.

Special cases: $d(C) \geq 1$ always true

$d(C) \geq 2 \iff$ all columns of H non-zero

$d(C) \geq 3 \iff$ for any two columns $\vec{c}_i, \vec{c}_j$, these are linearly independent.

This means $\nexists\ a, b \in K$ s.t. $a\vec{c}_i + b\vec{c}_j = 0$

$\Updownarrow$

$\vec{c}_i, \vec{c}_j \neq \vec{0}$, and there is no $\lambda \in K \setminus \{0\}$ such that $\vec{c}_j = \lambda \vec{c}_i$. $\left(\lambda = -\frac{a}{b}\right)$

Ex (Hamming codes) In K^m, choose a maximal set S of $\vec{c}_i$ s.t. no two are ~~linearly dependent~~ rescalings of each other. (& no $\vec{c}_i$ is $\vec{0}$)

There are $|K^m \setminus \{\vec{0}\}| = q^m - 1$ to choose, each is equivalent by rescalings to $|K \setminus \{0\}| = q - 1$ others, find $|S| = \frac{q^m - 1}{q - 1}$.

Take $H = [\vec{c}_1, \ldots, \vec{c}_{|S|}]$, a $(m \times |S|)$-matrix. Then G is a $(|S| - m) \times |S|$-matrix, & code C has $d(C) \geq 3$, i.e. is 1-error correcting. Hamming bound for 1-error correcting code in $K^{|S|}$ is

$$|C| \leq \frac{q^{|S|}}{1 + (|S|)(q-1)} = \frac{q^{|S|}}{1 + (q^m - 1)} = q^{|S| - m} = |S|_q, \quad |C|,$$

so C is 1-perfect.

Ex: $q = 2$, $m = 2 \rightsquigarrow H = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ $G = (1 \quad 1 \quad 1)$

$q = 2$, $m = 3 \rightsquigarrow H = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}$

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

equivalent to Fano code

Corollary If $C \subseteq K^n$ is a code of dimension k, then $d(C) \leq n - k + 1$

Pf H has $n-k$ rows, so a collection of $n-k+1$ column vectors must be lin. dependant. Thus $d(C) < (n-k+1)+1$.

Encoding & decoding with linear codes

Given (n,k) code C with generator matrix G, encoding is given by function $\phi_G : K^k \to K^n$

$$\phi_G(\vec{w}) = \vec{w}G.$$

To decode, given received $\vec{y} \in K^n$ solve the equation $\vec{w}G = \vec{y}$ for $\vec{w}$.

A code C is <u>systematic</u> if $G = [I_k, G']$, where G' is a $k \times (n-k)$-matrix

E.g. Parity code $G = [I_{n-1}, G']$, $G' = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$.

If C is systematic, $\phi_G(\vec{w}) = \vec{w}G = (\vec{w}I_k, \vec{w}G') = (\vec{w}, \vec{w}G')$, so decoding is simple.

<u>Error-detection & correction</u>: Assume C is t-error correcting, & that a message is changed by noise in at most t components.

Want to send $\vec{w} \in K^k$, transmit $\vec{y} = \vec{w}G$. Received is $\vec{z} = \vec{y} + \vec{e}$, where $\vec{e}$ is the error, & we know $w(\vec{e}) \leq t$.

Then $H\vec{z} = H\vec{y} + H\vec{e} = H\vec{e}$, since $\vec{y} \in C \Leftrightarrow H\vec{y} = \vec{0}$.

Moreover, $w(\vec{e}) \leq t$ means $\vec{e} = 0$ or $\vec{e} \notin C$, so $H\vec{e} = 0 \Leftrightarrow \vec{e} \in C \Leftrightarrow \vec{e} = \vec{0}$, i.e. message was received correctly.

Call $H\vec{z}$ the <u>syndrome</u> of $\vec{z}$. If $H\vec{z} \neq \vec{0}$, find the unique $\vec{e} \in K^k$ of weight $\leq t$ s.t. $H\vec{e} = H\vec{z}$.

Then $\vec{z} - \vec{e} \in C$, since $H(\vec{z} - \vec{e}) = H\vec{z} - H\vec{e} = 0$, so know $\vec{y} = \vec{z} - \vec{e}$.