INTRODUCTION TO LINEAR CODES

Patrick Fitzpatrick

University College Cork

S^3Cm, 2–11 July 2008

Codes and Linear Codes

Let p be a prime number, $q = p^r$ and $\mathbb{F}_q$ a finite field with q elements.

- $\mathbb{F}_q$ is called the **alphabet**
- An **information sequence** is a finite sequence $m = x_1 x_2 \cdots x_k$ where $x_i \in \mathbb{F}_q$.
- Each information sequence is encoded as the image of an injective map $i : \{sequences\ of\ length\ k\} \to \{sequences\ of\ length\ n\}$
- The set $\mathcal{C} = Im(i)$ is called a **(block) code.** The elements of $\mathcal{C}$ are called **codewords** and n is the **length** of the code.

The most useful codes have some structure. If the sets of sequences of length k, n are regarded as the vector spaces $\mathbb{F}_q^k, \mathbb{F}_q^n$ and the map i is linear then the code then $\mathcal{C}$ a **linear code** and its **dimension** is k. From now on all our codes are linear.

The most useful codes have some structure. If the sets of sequences of length k, n are regarded as the vector spaces $\mathbb{F}_q^k, \mathbb{F}_q^n$ and the map i is linear then the code then $\mathcal{C}$ a **linear code** and its **dimension** is k. From now on all our codes are linear.

Given $x, y \in \mathbb{F}_q^n$ we define **Hamming distance** between x and y as

$$d(x, y) = \sharp\{i \mid 1 \leq i \leq n, x_i \neq y_i\}.$$

The most useful codes have some structure. If the sets of sequences of length k, n are regarded as the vector spaces $\mathbb{F}_q^k, \mathbb{F}_q^n$ and the map i is linear then the code then $\mathcal{C}$ a **linear code** and its **dimension** is k. From now on all our codes are linear.

Given $x, y \in \mathbb{F}_q^n$ we define **Hamming distance** between x and y as

$$d(x, y) = \sharp\{i \mid 1 \leq i \leq n, x_i \neq y_i\}.$$

Given $z \in \mathbb{F}_q^n$ we define **Hamming weight** of z as $wt(z) = d(z, 0)$. Thus, $d(x, y) = wt(x - y)$.

Minimum distance

The **minimum distance** of $\mathcal{C}$ is,

$$d = d(\mathcal{C}) = min\{d(x, y) \mid x, y \in \mathcal{C}, x \neq y\}$$

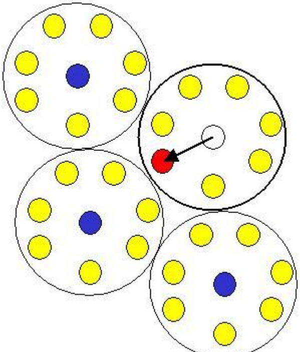
and obviously this is also $min\{wt(c) \mid c \in \mathcal{C}, c \neq 0\}$.

Maximum likelihood decoding

If the received word is $x \in \mathbb{F}_q^n$ then we decode it as an element $c \in \mathcal{C}$ which minimizes $d(x, c)$. This is not necessarily uniquely defined. The balls of radius $t = \lfloor (d-1)/2 \rfloor$ centred on codewords are disjoint. If the number of errors is $\leq t$ then there is a unique codeword closest to x.

Maximum likelihood decoding

If the received word is $x \in \mathbb{F}_q^n$ then we decode it as an element $c \in \mathcal{C}$ which minimizes $d(x, c)$. This is not necessarily uniquely defined. The balls of radius $t = \lfloor (d-1)/2 \rfloor$ centred on codewords are disjoint. If the number of errors is $\leq t$ then there is a unique codeword closest to x.



The corrupted word still lies in its original sphere. The center of this sphere is the corrected word.

$\mathcal{C}$ is a vector space of dimension k, so there exist linearly independent codewords $c_1, \ldots, c_k \in \mathcal{C}$ forming a basis of $\mathcal{C}$. The $k \times n$ matrix G whose rows are the c_i is called a **generator matrix** of $\mathcal{C}$. Notice that there are as many generator matrices as there are bases of $\mathcal{C}$.

$\mathcal{C}$ is a vector space of dimension k, so there exist linearly independent codewords $c_1, \ldots, c_k \in \mathcal{C}$ forming a basis of $\mathcal{C}$. The $k \times n$ matrix G whose rows are the c_i is called a **generator matrix** of $\mathcal{C}$. Notice that there are as many generator matrices as there are bases of $\mathcal{C}$.

Given $a \in \mathbb{F}_q^k$, the encoding map is just $a \to aG$.

$\mathcal{C}$ is a vector space of dimension k, so there exist linearly independent codewords $c_1, \ldots, c_k \in \mathcal{C}$ forming a basis of $\mathcal{C}$. The $k \times n$ matrix G whose rows are the c_i is called a **generator matrix** of $\mathcal{C}$. Notice that there are as many generator matrices as there are bases of $\mathcal{C}$.

Given $a \in \mathbb{F}_q^k$, the encoding map is just $a \to aG$.

A **parity check matrix** is an $(n-k) \times n$ matrix H such that $Hx^t = 0$ for every $x \in \mathcal{C}$. Note that by definition the rows of H are also linearly independent. Sometimes it is convenient to consider matrices H satisfying the given property whose rows are not linearly independent.

$\mathcal{C}$ is a vector space of dimension k, so there exist linearly independent codewords $c_1, \ldots, c_k \in \mathcal{C}$ forming a basis of $\mathcal{C}$. The $k \times n$ matrix G whose rows are the c_i is called a **generator matrix** of $\mathcal{C}$. Notice that there are as many generator matrices as there are bases of $\mathcal{C}$.

Given $a \in \mathbb{F}_q^k$, the encoding map is just $a \to aG$.

A **parity check matrix** is an $(n-k) \times n$ matrix H such that $Hx^t = 0$ for every $x \in \mathcal{C}$. Note that by definition the rows of H are also linearly independent. Sometimes it is convenient to consider matrices H satisfying the given property whose rows are not linearly independent.

The code with generator matrix H is called the **dual code** of $\mathcal{C}$ and we denote it as $\mathcal{C}^\perp$.

Goal

A code $\mathcal{C}$ may be described by three parameters $[n, k, d]$. The main goal is to maximize k (which gives the most efficient use of the communication channel) **and**d (which means the code corrects the highest number of errors). However these are competing objectives.

Goal

A code $\mathcal{C}$ may be described by three parameters $[n, k, d]$. The main goal is to maximize k (which gives the most efficient use of the communication channel) **and** d (which means the code corrects the highest number of errors). However these are competing objectives.

For example, if we fix n the best code in the sense of distance is the **repetition code** with generator matrix $(1, 1, \ldots, 1)$. This is a $[n, 1, n]$ code.

Goal

A code $\mathcal{C}$ may be described by three parameters $[n, k, d]$. The main goal is to maximize k (which gives the most efficient use of the communication channel) **and** d (which means the code corrects the highest number of errors). However these are competing objectives.

For example, if we fix n the best code in the sense of distance is the **repetition code** with generator matrix $(1, 1, \ldots, 1)$. This is a $[n, 1, n]$ code.

In the other hand the best code in the sense of dimension is generated by Id_n, which is a $[n, n, 1]$ code.

Goal

A code $\mathcal{C}$ may be described by three parameters $[n, k, d]$. The main goal is to maximize k (which gives the most efficient use of the communication channel) **and**d (which means the code corrects the highest number of errors). However these are competing objectives.

For example, if we fix n the best code in the sense of distance is the **repetition code** with generator matrix $(1, 1, \ldots, 1)$. This is a $[n, 1, n]$ code.

In the other hand the best code in the sense of dimension is generated by Id_n, which is a $[n, n, 1]$ code.

Singleton bound

For any $[n, k, d]$ code we have $d \leq n - k + 1$.

Cyclic codes

We say that a linear code $\mathcal{C}$ of length n over $\mathbb{F}_q$, is **cyclic**, if for every $(c_0, c_1, \ldots, c_{n-1}) \in \mathcal{C}$ we have that$(c_{n-1}, c_0, \ldots, c_{n-2}) \in \mathcal{C}$.

Cyclic codes

We say that a linear code $\mathcal{C}$ of length n over $\mathbb{F}_q$, is **cyclic**, if for every $(c_0, c_1, \ldots, c_{n-1}) \in \mathcal{C}$ we have that$(c_{n-1}, c_0, \ldots, c_{n-2}) \in \mathcal{C}$.

The reason for the change of subscript notation is that we may identify every codeword $(c_0, c_1, \ldots, c_{n-1})$ with a univariate polynomial
$(c_0, c_1, \ldots, c_{n-1}) \leftrightarrow c(X) = c_0 + c_1 X + \cdots + c_{n-1})X^{n-1}$.

Cyclic codes

We say that a linear code $\mathcal{C}$ of length n over $\mathbb{F}_q$, is **cyclic**, if for every $(c_0, c_1, \ldots, c_{n-1}) \in \mathcal{C}$ we have that$(c_{n-1}, c_0, \ldots, c_{n-2}) \in \mathcal{C}$.

The reason for the change of subscript notation is that we may identify every codeword $(c_0, c_1, \ldots, c_{n-1})$ with a univariate polynomial
$(c_0, c_1, \ldots, c_{n-1}) \leftrightarrow c(X) = c_0 + c_1 X + \cdots + c_{n-1})X^{n-1}$.

Main property of cyclic codes

$\mathcal{C}$ is a cyclic code if and only if it is an ideal of $\mathbb{F}_q[X]/\langle X^n - 1\rangle$ (under the correspondence $c(x) \leftrightarrow \overline{c(x)} \in \mathbb{F}_q[X]/\langle X^n - 1\rangle$).

Cyclic codes

A nice property of $\mathbb{F}_q[X]/\langle X^n - 1\rangle$ is that every ideal I is generated by only one polynomial (i.e. it is a **principal ideal ring**), say $I = \langle g(X)\rangle$, with $g(X) \mid X^n - 1$.

Cyclic codes

A nice property of $\mathbb{F}_q[X]/\langle X^n - 1\rangle$ is that every ideal I is generated by only one polynomial (i.e. it is a **principal ideal ring**), say $I = \langle g(X)\rangle$, with $g(X) \mid X^n - 1$.

For any cyclic code of length n, there exists a unique monic polynomial $g(X) \in \mathbb{F}_q$ dividing $X^n - 1$ such that $C = \langle g(X)\rangle$.Therefore, the codewords are precisely the multiples of $g(X)$ with degree less than n. The polynomial $g(X)$ is called the **generator polynomial**.

Cyclic codes

Let $\mathcal{C}$ be a cyclic code of length n with generator polynomial $g(X)$ of degree $n-k$. Then

$$\{g(X), Xg(X), \ldots, X^{k-1}g(X)\}$$

is a basis (vector space) of $\mathcal{C}$. Thus, $\mathcal{C}$ has dimension k.

Cyclic codes

Let $\mathcal{C}$ be a cyclic code of length n with generator polynomial $g(X)$ of degree $n-k$. Then

$$\{g(X), Xg(X), \ldots, X^{k-1}g(X)\}$$

is a basis (vector space) of $\mathcal{C}$. Thus, $\mathcal{C}$ has dimension k.

In particular, if $g(X) = g_0 + g_1 X + \cdots + g_{n-k} X^{n-k}$ then the matrix G is a generator matrix of $\mathcal{C}$:

$$G = \begin{pmatrix} g_0 & g_1 & g_2 & \cdots & g_{n-k} & & & & \\ & g_0 & g_1 & g_2 & \cdots & g_{n-k} & & & \\ & & \ddots & \ddots & & \ddots & & & \\ & & & & & g_0 & g_1 & g_2 & \cdots & g_{n-k} \end{pmatrix}$$

Let $\mathcal{C}$ a cyclic code of length n with generator polynomial $g(X)$ of degree $n-k$. The polynomial

$$h(X) = \frac{X^n - 1}{g(X)} = h_0 + h_1 X + \cdots + h_k X^k$$

is called the **parity check polynomial.**

Let $\mathcal{C}$ a cyclic code of length n with generator polynomial $g(X)$ of degree $n-k$. The polynomial

$$h(X) = \frac{X^n - 1}{g(X)} = h_0 + h_1 X + \cdots + h_k X^k$$

is called the **parity check polynomial.**

The $(n-k) \times n$ matrix H is a parity check matrix of $\mathcal{C}$,

$$H = \begin{pmatrix} & & & & h_k & h_{k-1} & h_{k-2} & \cdots & h_0 \\ & & & h_k & h_{k-1} & h_{k-2} & \cdots & h_0 & \\ & & \cdots & \cdots & & \cdots & & & \\ h_k & h_{k-1} & h_{k-2} & \cdots & h_0 & & & & \end{pmatrix}$$

Zeros of a cyclic code

Suppose that $f_1(X), \ldots, f_r(X)$ are the irreducible factors of the generator polynomial of $\mathcal{C}$, and $\{\alpha_i\}_1^s$ is the set of all roots of the f_i lying in a splitting field $\mathbf{F}_{q^m}$ of $X^n - 1$ over $\mathbf{F}_q$. Then,

$$\mathcal{C} = \langle g(X) \rangle = \{c(X) \mid c(\alpha_1) = \cdots = c(\alpha_s) = 0\}.$$

Zeros of a cyclic code

Suppose that $f_1(X), \ldots, f_r(X)$ are the irreducible factors of the generator polynomial of $\mathcal{C}$, and $\{\alpha_i\}_1^s$ is the set of all roots of the f_i lying in a splitting field $\mathbf{F}_{q^m}$ of $X^n - 1$ over $\mathbf{F}_q$. Then,

$$\mathcal{C} = \langle g(X) \rangle = \{c(X) \mid c(\alpha_1) = \cdots = c(\alpha_s) = 0\}.$$

Notice that we may think in the opposite direction, i.e. a set $\{\alpha_1, \ldots, \alpha_s\} \in \mathbb{F}_{q^m}$ defines a cyclic code.

Parity check matrix over extension field

The $s \times n$ matrix H' extends the definition of parity check matrix because $c(X) \in \mathcal{C}$ if and only if $H'C(X) = 0$.

$$H' = \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha_s & \alpha_s^2 & \cdots & \alpha_s^{n-1} \end{pmatrix}$$

BCH and RS Codes

We fix a field $\mathbb{F}_q$, natural numbers n, b and δ with $2 \leq \delta \leq n$. Let m satisfy $q^m \equiv 1 (\mod n)$, and let $\alpha \in \mathbb{F}_{q^m}$ be a primitive n^{th} root of unity.

BCH and RS Codes

We fix a field $\mathbb{F}_q$, natural numbers n, b and δ with $2 \leq \delta \leq n$. Let m satisfy $q^m \equiv 1 \pmod{n}$, and let $\alpha \in \mathbb{F}_{q^m}$ be a primitive n^{th} root of unity.

Bose, Ray-Chaudhuri, Hocquenghem

A **BCH code** of length n and **designed distance** δ is the cyclic code with generator polynomial having roots $\alpha^b, \alpha^{b+1}, \ldots, \alpha^{b+\delta-2}$.

BCH and RS Codes

We fix a field $\mathbb{F}_q$, natural numbers n, b and δ with $2 \leq \delta \leq n$. Let m satisfy $q^m \equiv 1 (\bmod n)$, and let $\alpha \in \mathbb{F}_{q^m}$ be a primitive n^{th} root of unity.

Bose, Ray-Chaudhuri, Hocquenghem

A **BCH code** of length n and **designed distance** δ is the cyclic code with generator polynomial having roots $\alpha^b, \alpha^{b+1}, \ldots, \alpha^{b+\delta-2}$.

- If $b = 1$ the code is called **strict sense**.
- If $n = q^m - 1$ the code is called a **primitive** BCH code.
- If $n = q - 1$ then the code is a **Reed-Solomon (RS)** code. Note that in this case $\alpha \in \mathbb{F}_q$.

Theorem

A BCH code of designed distance δ has minimum distance $d \geq \delta$. In general we do not know the exact value of the minimum distance.

Theorem

A BCH code of designed distance δ has minimum distance $d \geq \delta$. In general we do not know the exact value of the minimum distance.

Theorem

An RS code of length n, dimension k and designed distance δ satisfies $d = \delta = n - k + 1$ and therefore has parameters $[n, k, n - k + 1]$. For this reason (in view of the Singleton bound) these codes are said to be **maximum distance separable (MDS)**.

Theorem

A BCH code of designed distance δ has minimum distance $d \geq \delta$. In general we do not know the exact value of the minimum distance.

Theorem

An RS code of length n, dimension k and designed distance δ satisfies $d = \delta = n - k + 1$ and therefore has parameters $[n, k, n - k + 1]$. For this reason (in view of the Singleton bound) these codes are said to be **maximum distance separable (MDS)**.

RS codes are among the most commonly used. The only inconvenience is the small length. Any element of $\mathbb{F}_q^r$ may be identified with a vector in $\mathbb{F}_q^r$. Then an RS code of length n is converted to a code of length rn. Those new codes are very good for decoding bursts, especially when **interleaved**.