Some slides for 1st Lecture, Coding theory

Diego Ruano

Department of Mathematical Sciences
Aalborg University
Denmark

8-02-2011

Block code

A block code C is a set of M codewords, where all the codewords are n-tuples and we refer to n as the length of the code.

- Where do the codewords live?
- This is not enough, we will work with block linear codes.

Recall:

- Let $\mathbb{F}$ be a field. Then $\mathbb{F}^n$ is a vector space.
- Vector subspace
- Basis of a vector space.
- Dimension of a vector space: number of elements of a basis
- Inner product $x \cdot y = \sum x_i y_i \in \mathbb{F}$

Example: $(1, 1) \cdot (1, 1)$?

Linear code

A linear (n, k) block code C is a k-dimensional vector subspace of $\mathbb{F}^n$.

Note:

- $(0, \ldots, 0) \in C$
- $M = q^k$.

Systematic encoding: $G = (I, A)$

Generator matrix

A generator matrix G of an (n, k) code C is $k \times n$ matrix whose rows are linearly independent.

Encoding rule: $c = uG$, where u is the information vector of length k.
Systematic encoding: $G = (I, A)$

A parity check is a vector h of length n such that

$$Gh^T = 0$$

Parity check matrix

A parity check matrix H for an (n, k) code is an $(n-k) \times n$ matrix whose rows are linearly independent parity checks.

- $GH^T = 0$
- $H = (-A^T, I)$ if $G = (I, A)$

How do we detect an error?:

Syndrome

For $r \in \mathbb{F}^n$

$$s = Hr^T$$

$$H(c+e)^T = He^T$$

Dual code

$$C^\perp = \{x \in \mathbb{F}^n : x \cdot c = 0, \forall c \in C\}$$

Rate $R = k/n$.
Rate of $C^\perp = 1 - R$.

How many errors can we correct?

Hamming weight

Let $x \in \mathbb{F}^n$, $w(x) = \#\{i : x_i \neq 0\}$

***t*-error correcting**

A code is t-error correcting if for all codeword c_1, c_2 and for any errors e_1, e_2 with weight $\leq t$, we have

$$c_1 + e_1 \neq c_2 + e_2$$

Hamming distance

Let $x, y \in \mathbb{F}^n$, $d(x, y) = \#\{i : x_i \neq y_i\}$

$\mathbb{F}^n$ is a metric space with this distance.

Hamming distance

$$d(C) = \min\{d(x, y) : x, y \in C\}$$

How many errors can we correct?

For linear codes: it is easier to compute d

Lemma 1.2.1

In an (n, k) code the minimum distance is equal to the minimum weight of a nonzero codeword.

Note: $w(x) = d(0, x)$ and $d(x, y) = w(x - y)$

Theorem 1.2.1

An (n, k) code is t-error correcting if and only if $t < d/2$. That is, if $t \leq \lfloor \frac{d-1}{2} \rfloor$

Theorem 1.2.1

An (n, k) code is t-error correcting if and only if $t < d/2$. That is, if $t \leq \lfloor \frac{d-1}{2} \rfloor$

Proof $\Rightarrow$:

- Let $c_1 + e_1 = c_2 + e_2$ with $c_i \in C$ and $w(e_i) \leq t$
- $w(c_1 - c_2) = w(e_2 - e_1) \leq w(e_2) + w(e_1) \leq 2t < d$. Contradiction.

Proof $\Leftarrow$:

- Let $t \geq d/2$ and $w(c) = d$.
- Change $\lceil \frac{d}{2} \rceil$ positions (of the non-zero positions) of c to zero.
- Then, $0 + y = c + (y - c)$ (think in $c_1 + e_1 = c_2 + e_2$)
- Hence, it is not t-error correcting because
- $d(0, y) \leq d - \lceil \frac{d}{2} \rceil \leq t$
- $d(c, y) = \lceil \frac{d}{2} \rceil \leq \lceil \frac{2t}{2} \rceil = t$

Lemma 1.2.1

Let C be an (n, k) code and H a parity check matrix for C.

- If j columns are linearly dependent, C contains a codeword with non-zero elements in some of the corresponding positions
- If C contains a word of weight j, then there exist j linearly dependent columns of H.

Proof: Think in $Hc^T = 0$

Lemma 1.2.3

Let C be an (n, k) code with parity check matrix H. Then minimum distance of C equals the minimum number of linearly dependent columns of H.

For a binary code $d \geq 3$ if and only if the columns of H are distinct and nonzero.

Theorem 1.2.2. Gilbert-Varshamov bound

There exists a binary linear code of length n, with at most m linearly independent parity checks and minimum distance at least d, if

$$1+\binom{n-1}{1}+\cdots+\binom{n-1}{d-2}<2^m$$

- For n large, good binary codes exist. How to construct them?
- For n large, can we get even better codes?
- Short codes, can have better minimum distances.

- Binary Hamming code
- Extended binary Hamming code