Algorithms of Information Security: Error-correcting codes I

Faculty of Information Technology
Czech Technical University in Prague

September 29, 2021

Basic definitions

Definition

Let $A = \{a_1, \ldots, a_q\}$ be an alphabet; we call the a_i values symbols. A block code C of length n over A is a subset of A^n. A vector $c \in C$ is called a codeword. The number of elements in C, denoted $|C|$, is called the size of the code. A code of length n and size M is called an (n, M)-code.

Example. A code over $A = \{0, 1\}$ is called a binary code and a code over $A = \{0, 1, 2\}$ is called a ternary code.
Example. The set $\{(0, 0, 0), (1, 1, 1)\}$ is the binary (3,2)-code.

Basic definitions

Definition

The Hamming distance between two strings x and y of the same length over a finite alphabet A is defined as the number of positions at which the two strings differ. Let $x = x_1, \ldots, x_n$ and $y = y_1, \ldots, y_n$, then for every i defined

$$\delta(x_i, y_i) = \begin{cases} 1, & x_i \neq y_i, \\ 0, & x_i = y_i \end{cases}$$

Hamming distance is defined by

$$d(x, y) = \sum_{i=1}^{n} \delta(x_i, y_i).$$

Example. In the space F_2^5 the Hamming distance satisfies $d(10111, 11001) = 3$ and in F_3^4 we have $d(1122, 1220) = 2$.

Note. Hamming distance d defines a metric on A^n. That is, for every $x, y, z \in A^n$:

1. $0 \leq d(x, y) \leq n$
2. $d(x, y) = 0$ if and only if $x = y$
3. $d(x, y) = d(y, x)$
4. (triangle inequality) $d(x, z) \leq d(x, y) + d(y, z)$.

Note. We stress that the Hamming distance is not dependent on the actual values of x_i and y_i but only if they are equal to each other or not equal.

Definition

Let C be a code of length n over an alphabet A. The *nearest neighbor* decoding rule states that every $x \in A^n$ is decoded to $c_x \in C$ that is closest to x. That is, $D(x) = c_x$ where c_x is such that $d(x, c_x) = \min_{c \in C} d(x, c)$.

Definition

Let C be a code. The distance of the code, denoted $d(C)$, is defined by

$$d(C) = \min\{d(c_1, c_2) \mid c_1, c_2 \in C, c_1 \neq c_2\}$$

An (n, M)-code of distance d is called an (n, M, d)-code. The values n, M, d are called the parameters of the code.

Restating what we have discussed above, the aim of coding theory is to construct a code with a short n, and large M and d. We now show a connection between the distance of a code and the possibility of detecting and correcting errors.

Definition

Let C be a code of length n over alphabet A.

- C detects u errors if for every codeword $c \in C$ and every $x \in A^n$ with $x \neq c$, it holds that if $d(x, c) \leq u$ then $x \notin C$.
- C corrects v errors if for every codeword $c \in C$ and every $x \in A^n$ it holds that if $d(x, c) \leq v$ then nearest neighbor decoding of x outputs c.

Theorem

- *A code C detects u errors if and only if $d(C) > u$.*
- *A code C corrects v errors if and only if $d(C) \geq 2v + 1$.*

Linear code

We denote by F_q a finite field of size q. Recall that there exists such a finite field for any q that is a power of a prime. In this course, we will just assume that we are given such a field. In linear codes, the alphabet of the code are the elements of some finite field F_q.

Definition

A linear code with length n over F_q is a vector subspace of F_q^n.

Example. The repetition code $C = \{\underbrace{(x, \ldots, x)}_{n} \mid x \in F_q\}$ is a linear code.

Notation. A linear code of length n and dimension k is denoted as $[n, k]$-code (or an $[n, k, d]_q$-code when the distance d and the size of the alphabet q are specified).

Note. Dimension k is not M, i.e., the size of the code.

Linear code

Definition

Let C be a linear $[n, k]_q$ code over F_q^n. Then

1. The *dual code* of C is $C^\perp$ (the orthogonal complement of C in $F_q^n, C^\perp = \{x \in F_q^n \mid \langle x, c\rangle = 0$ for all $c \in C\}$) Notice that $C^\perp$ is an $[n, n-k]_q$ code.
2. The *dimension* of C is the dimension of C as a vector subspace of F_q^n, denoted dim(C).

Theorem

Let C be a linear code of length n over F_q. Then

1. $\mid C \mid = q^{dim(C)}$ *(dim$(C) = k$, i.e., dimension of a code).*
2. $C^\perp$ *is a linear code, and dim(C) + dim$(C^\perp) = n$.*
3. $(C^\perp)^\perp = C$.

Linear code

Definition

Let C be a linear code. Then

1. C is *self orthogonal* if $C \subseteq C^{\perp}$.
2. C is *self dual* if $C = C^{\perp}$.

The following theorem is an immediate corollary of the fact that $\dim(C) + \dim(C^{\perp}) = n$.

Theorem

1. *Let C be a self-orthogonal code of length n. Then dim$(C) \leq \frac{n}{2}$.*
2. *Let C be a self-dual code of length n. Then dim$(C) = \frac{n}{2}$.*

Definitions

Definition

Let $x \in F_q^n$. The Hamming weight of x, denoted $\mathrm{wt}(x)$ is defined to be the number of coordinates that are not zero. That is, $\mathrm{wt}(x) = d(x, 0)$.

Definition

Let C be a code (not necessarily linear). The weight of C, denoted $\mathrm{wt}(C)$, is defined by

$$\mathrm{wt}(C) = \min_{c \in C; c \neq 0} \{\mathrm{wt}(c)\}.$$

The following theorem only holds for linear codes:

Theorem

Let C be a linear code over F_q^n. Then $d(C) = wt(C)$.

Generator and Parity-Check Matrices

Definition

1. *A generator matrix* G for a linear code C is a matrix whose rows form a basis for C.
2. *A parity check matrix* H for C is a generator matrix for the dual code $C^{\perp}$.

Remarks:

1. If C is a linear $[n, k]$-code then $G \in F_q^{k \times n}$(recall that k denotes the number of rows and n the number of columns), and $H \in F_q^{(n-k) \times n}$.
2. The rows of a generator matrix are linearly independent.
3. In order to show that a $k-$ by $-n$ matrix G is a generator matrix of a code C it suffices to show that the rows of G are codewords in C and that they are linearly independent.

Definition

1. A generator matrix is said to be in standard form if it is of the form $(I_k \mid X)$, where I_k denotes the $k-$by$-k$ identity matrix.
2. A parity check matrix is said to be in standard form if it is of the form $(Y \mid I_{n-k})$.

Lemma

Let C be a linear $[n,k]$-code with generator matrix G. Then for every $v \in F_q^n$ it holds that $v \in C^{\perp}$ if and only if $v \cdot G^T = 0$. In particular, a matrix $H \in F_q^{(n-k)\times n}$ is a parity check matrix if and only if its rows are linearly independent and $H \cdot G^T = 0$.

An equivalent formulation: Let C be a linear $[n,k]$-code with a parity check matrix H. Then $v \in C$ if and only if $v \cdot H^T = 0$.

Theorem

Let C be a linear code and let H the parity check matrix for C. Then

1. *$d(C) \geq d$ if and only if every subset of $d-1$ columns of H are linearly independent.*
2. *$d(C) \leq d$ if and only if there exists a subset of d columns of H that are linearly dependent.*

Corollary. Let C be a linear code and let H be a parity check matrix for C. Then $d(C) = d$ if and only if every subset of $d-1$ columns in H are linearly independent and there exists a subset of d columns that are dependent in H.

Theorem

If $G = (I_k \mid X)$ is the generator matrix in standard form for a linear $[n,k]$-code C, then $H = (-X^T \mid I_{n-k})$ is a parity check matrix for C.

Equivalence of Codes

Definition

Two (n, M)-codes are equivalent if one can be derived from the other by a permutation of the coordinates and multiplication of any specific coordinate by a non-zero scalar.

Theorem

Every linear code C is equivalent to a linear code C' with a generator matrix in standard form.

Polynomial code

Fix a finite field F_q. For the purpose of constructing polynomial codes, we identify a word of n elements $c = (c_0, \ldots, c_{n-1})$ with its representing polynomial $c(x) = \sum_{i=0}^{n-1} c_i x^i$.

Definition

Fix some integer n and let $g(x)$ be some fixed polynomial of degree $m \leq n - 1$. The polynomial code generated by $g(x)$ is the code whose codewords are the polynomials of degree less than n that are divisible (without remainder) by $g(x)$.

Example. Let $n = 5, m = 2$ and consider the polynomial $g(x) = x^2 + x$ over F_2. Using $g(x)$, we generate the polynomials of degree ≤ 4, i.e., polynomials in the form $p(x) \cdot g(x)$, where $p(x) \in \{0, 1, x, (x+1), x^2, (x^2+1), (x^2+x), (x^2+x+1)\}$. Written explicitly:
$0, x^2 + x, x^3 + x^2, x^3 + x, x^4 + x^3, x^4 + x^3 + x^2 + x, x^4 + x^2, x^4 + x.$
And we can represent them as strings of binary digits:
00000, 00110, 01100, 01010, 11000, 11110, 10100, 10010.

Cyclic code

Definition

A code C is cyclic if every cyclic shift of a codeword in C is also a codeword. That is, $(c_0, c_1, \ldots, c_{n-1}) \in C$ implies that $(c_{n-1}, c_0, \ldots, c_{n-2}) \in C$.

In the notation of representing polynomials, a code C is cyclic if and only if $c(x) \in C$ implies

$$x \cdot c(x) \text{ mod } (x^n - 1) \in C.$$

If a code is linear, then equivalently we can say that $c(x) \in C$ implies

$$u(x) \cdot c(x) \text{ mod } (x^n - 1) \in C$$

for every $u(x) \in F_q[x]$. Hence, C is a linear cyclic code if and only if C is an ideal in the ring $F_q[x]/(x^n - 1)$.

Cyclic code

Theorem

Let C be a cyclic code over F_q and g the monic polynomial in C of minimal positive degree (prove that it is unique!). Then g generates C, i.e., $c \in C$ iff $g \mid c$.

Theorem

A polynomial code is cyclic if and only if its generator polynomial divides $x^n - 1$, where n is length of the code.

Dual Codes of Cyclic Codes

Let C be a cyclic $[n, k]$-code with a generator $g(x) = \sum_{i=0}^{n-k} g_i x^i$.
We know that g divides $x^n - 1$, and therefore, there exists $h(x) = \sum_{i=0}^{k} h_i x^i$ such that $gh = x^n - 1$.
Let $c \in C$. As g generates C we have $c = ga$ for some $a \in F_q[x]$.
Therefore

$$hc \text{ mod } (x^n - 1) = hga \text{ mod } (x^n - 1) = 0$$

This translates to the constraints:

$$c_0 h_i + c_1 h_{i-1} + \ldots + c_{n-k} h_{i-n+k} = 0,$$

for every $0 \leq i \leq n - 1$, where the indices are modulo n.

Dual Codes of Cyclic Codes

It follows that

$$H = \begin{pmatrix} h_k & h_{k-1} & \dots & h_0 & & & & \\ & h_k & h_{k-1} & \dots & h_0 & & & \\ \vdots & & \ddots & & & & & \\ & & & & h_k & h_{k-1} & \dots & h_0 \end{pmatrix}$$

is a $(n-k) \times n$ matrix of parity checks of C, and because it has the correct rank $n-k$ it is a parity check matrix of C.

Theorem

Let C be an $[n, k]$ cyclic code generated by $g(x)$ and let $h(x) = \frac{x^n - 1}{g(x)}$. Then, the dualcode of C is a cyclic $[n, n-k]$code whose generator polynomial is $x^k h(x^{-1})$. The polynomial $h(x)$ is called the check polynomial of C.

The Binary Hamming Code

Definition

Let $r \geq 2$ and let C be a binary linear code with $n = 2^r - 1$ whose parity check matrix H is such that the columns are all of the non-zero vectors in F_2^r. This code C is called a binary Hamming code of length $2^r - 1$, denoted Ham$(r, 2)$.

Propositions.

1. All binary Hamming codes of a given length are equivalent.
2. For every $r \in \mathbb{N}$, the dimension of Ham$(r, 2)$ is $k = 2^r - r - 1$.
3. For every $r \in \mathbb{N}$, the distance of Ham$(r, 2)$ is $d = 3$ and so the code can correct exactly one error.

Example. A generator matrix for Ham$(r,2)$, where $r=3$, is as follows:

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

Size of the code is
$M = \mid C \mid = \mid \{\sum_{i=1}^{4} u_i v_i, u_i \in \{0,1\}\} \mid = 2^4 = 16.$
The parity check matrix for Ham$(r,2)$ is as follows:

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

As can be seen, for Ham$(r,2)$ we have $n=7, k=4$ and H is the matrix of type (3×7) over F_2.

The Hamming code is cyclic

Any binary Hamming code is equivalent to a cyclic code.

Theorem

Fix a field F_{2^r} and let $n = 2^r - 1$. Then, there exists a $[n, k = n - r, 3]_2$ cyclic code. Since the only code with such length, dimension and distance is the Hamming code, the Hamming code is cyclic.