IERG 6120 Coding Theory for Storage Systems Lecture 3 - 15/09/2016

Finite Field and Linear Codes

Lecturer: Kenneth Shum *Scribe: Yulin Shao*

Non-linear block code is not easy to construct and describe, and there is no fast decoding in general. Thus, we study the linear code in this lecture.

1 Finite field

Definition 1.1. A finite field (or Galois field) is a field that contains a finite number of elements. It is a set on which the operations of addition, subtraction, multiplication and division are defined and satisfy certain basic rules.

In this lecture we consider finite field whose size is a prime number. This is usually called the *prime field*, and is denoted by $\mathbb{Z}_p$ (or $GF(p)$, $\mathbb{F}_p$), where p is a prime number. The elements are $0, 1, ..., p-1$. Further, we define the following operations on $\mathbb{Z}_p$:

- addition: $\forall x, y \in \mathbb{Z}_p$, $x + y \stackrel{def}{=} (x + y) \bmod p$.
- multiplication: $\forall x, y \in \mathbb{Z}_p$, $x \cdot y \stackrel{def}{=} (x \cdot y) \bmod p$.
- additive inverse: $\forall x \in \mathbb{Z}_p$, $(p - x) \bmod p$ is the additive inverse of x.
- multiplicative inverse: $\forall x \in \mathbb{Z}_p, x \neq 0$, the multiplicative inverse of x, denoted by y, $y \in \mathbb{Z}_p$, satisfies $x \cdot y \equiv 1 \bmod p$.

(We will use the notation $x \equiv y \bmod p$ to mean that $x - y$ is an integral multiple of p.)

Example 1.1. If $p = 5$, then all the elements in $\mathbb{Z}_5$ are given by $\{0, 1, 2, 3, 4\}$. $2+3 \equiv 0 \bmod 5$, $2 \cdot 3 \equiv 1 \bmod 5$. The additive inverse of 2 is 3 and the multiplicative inverse of 2 is also 3.

Lemma 1.1 (No zero divisor)**.** *If $x \cdot y \equiv 0 \bmod p$, then $x \equiv 0 \bmod p$ or $y \equiv 0 \bmod p$.*

Proof Since $x \cdot y \equiv 0 \bmod p$, we have: "$xy - 0$ is divisible by p", that is, "p divides xy".
Thus, "p is a factor of x" or "p is a factor of y".
Finally, we have: $x \equiv 0 \bmod p$ or $y \equiv 0 \bmod p$. □

Proposition 1.2. *If $a \not\equiv 0 \bmod p$, then multiplication by a is a bijection from $\mathbb{Z}_p$ to $\mathbb{Z}_p$.*

Proof Let $f(x) = ax \bmod p$.
Suppose that $\exists x, y \in \mathbb{Z}_p$, $x \neq y$, but $f(x) = f(y)$, namely, $ax \equiv xy \bmod p$.
We have $ax - ay \equiv 0 \bmod p$, hence $a(x - y) \equiv 0 \bmod p$.
Since $a \not\equiv 0 \bmod p$, by Lemma 1.1, we have $x - y \equiv 0 \bmod p$, namely, $x \equiv y \bmod p$.
Thus, f is an injection.
Since f is an injection from a finite set $\mathbb{Z}_p$ to itself, f is also a surjection.
Finally, f is a bijection. □

In **Proposition 1.2**, we have used the following elementary fact.

Theorem 1.3. *For finite sets S and T with the same cardinality, if g is a one-to-one mapping from S to T, then g is also a surjection.*

Proof We prove by contradiction and suppose that g is not surjective. Then there is an element in T, say t, which does not have any pre-image. The function g can be regarded as a function from S to $T \setminus \{t\}$. Since the cardinality of $T \setminus \{t\}$ is strictly less than that of S, by the Pigeonhole Principle, there are two elements s and s' in the domain S which are mapped to the same element in $T \setminus \{t\}$, contradicting the assumption that g is injective. □

Proposition 1.4. *Multiplicative inverse of a nonzero element exists in $\mathbb{Z}_p$, and is unique.*

Proof Let $x \in \mathbb{Z}_p$, $a \not\equiv 0 \bmod p$. Multiply all the elements in $\mathbb{Z}_p$ by a, then there is one and only one element y satisfies $a \cdot y \equiv 1 \bmod p$ (**Proposition 1.2**). The element y is exactly the multiplicative inverse of the nonzero element x. □

Proposition 1.5 (Bezout's theorem)**.** *Given two integers a and b, we can find two integers r and s such that the greatest common devisor of a, b, denoted by $gcd(a, b)$, is equal to $ra + sb$.*

The integers r and s in **Proposition 1.5** can be computed by extended Euclidean algorithm.
Implication: $\forall a \not\equiv 0 \bmod p$, $gcd(a, p) = 1$. By **Proposition 1.5**, we have $ra + sp = 1$ from some integers r and s. Thus, $ra \equiv 1 \bmod p$. This implies that r is the multiplicative inverse of a, and we can utilize the extended Euclidean algorithm to calculate the multiplicative inverse.

Example 1.2. Given $p = 11$, find $3^{-1} \bmod 11$ in $\mathbb{Z}_{11}$.

- Scheme 1 (The general way when p is not large): We try all the nonzero elements in $\mathbb{Z}_{11}$. Specifically, $3 \cdot 1 \equiv 3 \bmod 11$, $3 \cdot 2 \equiv 6 \bmod 11$, $3 \cdot 3 \equiv 9 \bmod 11$, $3 \cdot 4 \equiv 1 \bmod 11$. Thus, 4 is the multiplicative inverse of 3.
- Scheme 2 (Utilize extended Euclidean algorithm):

r	s	$ra + sp$	
0	1	$p = 11$	① Initialization
1	0	$a = 3$	② Initialization
–3	1	2	③ 11 mod 3 =2
4	–1	1	④ 3 mod 2 =1

In the above table, row 3 is obtained by subtracting 3 times row 2 from row 1. Row 4 is obtained by subtracting row 3 from row 2. By the extended Euclidean algorithm, we have $4 \cdot 3 - 11 = gcd(3, 11) = 1$. Thus, $3^{-1} \equiv 4 \bmod 11$.

2 Vector space

Definition 2.1 (Vector space)**.** Let $\mathbb{Z}_q$ be the finite field of order q, q is a prime number. A nonempty set $V \subseteq \mathbb{Z}_q^n$, together with some (vector) addition '+' and scalar multiplication by elements of $\mathbb{Z}_q$, is a vector space over $\mathbb{Z}_q$ if it satisfies the following conditions. $\forall\ \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\forall\ \lambda, \mu \in \mathbb{Z}_q$:

(1) $\mathbf{u} + \mathbf{v} \in V$;
(2) $\lambda\mathbf{u} \in V$;
(3) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$, $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$;
(4) $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{u}$, $(\lambda + \mu)\mathbf{u} = \lambda\mathbf{u} + \mu\mathbf{u}$, $(\lambda\mu)\mathbf{u} = \lambda(\mu\mathbf{u})$;
(5) There is an element $\mathbf{0} \in V$ with the property $\mathbf{0} + \mathbf{u} = \mathbf{u}, \forall\ \mathbf{u} \in V$;
(6) $\forall\ \mathbf{u} \in V$, there is an element $-\mathbf{u} \in V$, s.t. $-\mathbf{u} + \mathbf{u} = \mathbf{0}$.

Example 2.1. For instance, $V_1 = \{(0,0,0,0), (1,0,1,0)(0,1,0,1)(1,1,1,1)\} \subseteq \mathbb{Z}_2^4$ is a vector space over $\mathbb{Z}_2$; $V_2 = \{(0,0,0), (0,1,2)(0,2,1)\} \subseteq \mathbb{Z}_3^3$ is a vector space over $\mathbb{Z}_3$.

Definition 2.2 (Subspace)**.** A nonempty subset C of a vector space V is a subspace of V if it is itself a vector space under the same vector addition and scalar multiplication as V.

Proposition 2.1. *A nonempty subset C of a vector space V over $\mathbb{Z}_q$ is a subspace if and only if the following condition is satisfied:*
if $\mathbf{x}, \mathbf{y} \in C$ and $\lambda, \mu \in \mathbb{Z}_q$, then $\lambda\mathbf{x} + \mu\mathbf{y} \in C$.

Example 2.2. For instance, the V_1 in **Example 2.1** is a vector space itself, while it is also a subspace of $\mathbb{Z}_2^4$; similarly, V_2 is a subspace of $\mathbb{Z}_3^3$.

3 Linear codes

Definition 3.1 (Linear code)**.** A linear code C of length n over $\mathbb{Z}_q$ is a subspace of $\mathbb{Z}_q^n$.

Example 3.1. The following are linear codes:
(i) The repetition code: $C = \{(\lambda, \lambda, ..., \lambda) : \lambda \in \mathbb{Z}_q, \forall\ q\}$;
(ii) $C = \{000, 001, 010, 011\}$ over $\mathbb{Z}_2$ is a subspace of $\mathbb{Z}_2^3$;
(iii) $C = \{0000, 1100, 2200, 0001, 0002, 1101, 1102, 2201, 2202\}$ over $\mathbb{Z}_3$ is a subspace of $\mathbb{Z}_3^4$.

Definition 3.2 (Basis of C)**.** Let C be a linear code over $\mathbb{Z}_q$. A nonempty subset $B = \{\mathbf{g_1}, \mathbf{g_2}, ..., \mathbf{g_k}\}$ of C is called a *basis* for C if B is linearly independent and any codeword $\mathbf{c} \in C$ can be expressed as a unique linear combination of vectors in B. i.e.,

$$\forall \mathbf{c} \in C,\ \mathbf{c} = \{\alpha_1\mathbf{g_1} + \alpha_2\mathbf{g_2} + ... + \alpha_k\mathbf{g_k} :\ \alpha_1, \alpha_2, ..., \alpha_k \in \mathbb{Z}_q\}.$$

Example 3.2. For the linear code $C = \{0000, 0001, 0010, 0100, 0011, 0101, 0110, 0111\}$ over $\mathbb{Z}_2$, one of its bases is $B = \{0001, 0010, 0100\}$.

Definition 3.3 (Dual code, dimension of C)**.** Let C be a linear code in $\mathbb{Z}_q^n$,

(i) The *dual code* of C, denoted by $C^\perp$, is defined as the orthogonal complement of the subspace C of $\mathbb{Z}_q^n$. Specifically, each element in $C^\perp$ is orthogonal to all the elements in C. $C^\perp \triangleq \{\mathbf{v} \in \mathbb{Z}_q^n :\ \forall\ \mathbf{c} \in C, \mathbf{v} \cdot \mathbf{c} = 0\}$.

(ii) The *dimension* of the linear code C is the number of elements in its bases, and is denoted by $\dim(C)$.

Theorem 3.1. *Let C be a linear code of length n over $\mathbb{Z}_q$. Then*
(i) $\dim(C) = \log_q |C|$*;*
(ii) $C^\perp$ *is a linear code and* $\dim(C) + \dim(C^\perp) = n$*;*
(iii) $(C^\perp)^\perp = C$.

Example 3.3. (i) For code $C = \{0000, 1010, 0101, 1111\}$ over $\mathbb{Z}_2$, we have $C^\perp = C = \{0000, 1010, 0101, 1111\}$, and $\dim(C)$=$\dim(C^\perp)$= $\log_2 4 = 2$

(ii) For code $C = \{000, 001, 002, 010, 020, 011, 012, 021, 022\}$ over $\mathbb{Z}_3$, we have $C^\perp = \{000, 100, 200\}$, $\dim(C)$=2 and $\dim(C^\perp)$=1.

4 Generator matrix and parity-check matrix

Knowing a basis for a linear code enables us to describe its codewords explicitly.

Definition 4.1.
(i) A *generator matrix* G for a linear code C is a matrix G whose rows form a basis for C.
(ii) A *parity-check matrix* H for a linear code C is a generator matrix for the dual code $C^{\perp}$.

Lemma 4.1. (i) $GH^T = 0$;
(ii) *The code C can also defined as* $C = \{\mathbf{v} \in \mathbb{Z}_q^n : \mathbf{v}H^T = \mathbf{0}\}$.

Example 4.1. Let C be the linear code over $\mathbb{Z}_3$ defined by the generator matrix $G = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 1 & 1 & 1 \end{bmatrix}$.

(i) We can write down all the codewords by the linear combination of the two rows in G, and we have $C = \{00000, 11100, 22200, 22111, 00211, 11011, 11222, 22022, 00122\}$.

(ii) We can obtain the parity-check matrix H as follows. First, we transform the matrix G to the reduced row echelon form by elementary row operations.

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 1 & 1 & 1 \end{bmatrix} \xrightarrow{\text{row1+row2}} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 \end{bmatrix} \xrightarrow{\text{row2*2}} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 2 \end{bmatrix} \xrightarrow{\text{row1-row2}} \begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 & 2 \end{bmatrix}$$

Then, we could obtain a parity-check matrix H using the defining property that $GH^T = 0$,

$$H = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 2 & 0 & 1 & 1 & 0 \\ 2 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

5 Cosets and standard array decoding

Definition 5.1 (Coset)**.** Let C be a linear code of length n over $\mathbb{Z}_q$, and let $\mathbf{u} \in \mathbb{Z}_q^n$ be any vector of length n; we define the *coset* of C determined by $\mathbf{u}$ to be the set

$$C + \mathbf{u} = \{\mathbf{v} + \mathbf{u} : \mathbf{v} \in C\}$$

Example 5.1. For a code $C = \{000, 101, 010, 111\}$ over $\mathbb{Z}_2$, we have $C + 101 = \{101, 000, 111, 010\}$.

Theorem 5.1. *Let C be a linear code of length n over $\mathbb{Z}_q$, $\dim(C) = k$. Then*
(i) *every vectors of $\mathbb{Z}_q^n$ is contained in some coset of C;*
(ii) *two cosets are either identical or they have empty intersection;*
(iii) *there are q^{n-k} different cosets of C;*
(iv) $\forall\ \mathbf{u}, \mathbf{v} \in \mathbb{Z}_q^n, \mathbf{u} - \mathbf{v} \in C$ *if and only if* $\mathbf{u}$ *and* $\mathbf{v}$ *are in the same coset;*

Implication: Let C be a linear code. Assume the codeword $\mathbf{v}$ is transmitted and the word $\mathbf{u}$ is received. Define the *error pattern* (or error string) as $\mathbf{e} = \mathbf{u} - \mathbf{v}$. Then, $\mathbf{u} - \mathbf{e} = \mathbf{v} \in C$. This implies that error pattern $\mathbf{e}$ is contained in the coset $\mathbf{u} + C$(part (iv) in **Theorem 5.1**). On the other hand, any vector in coset $\mathbf{u} + C$ may be an error pattern leading to the received word $\mathbf{u}$.

Since error patterns $\mathbf{e}$ with smaller Hamming weight are the more likely to occur, nearest neighbour decoding works for a linear code C in the following manner. Upon receiving the word $\mathbf{u}$, we choose a word $\mathbf{e}$ of least weight in the coset $\mathbf{u} + C$ and declare that $\mathbf{u} - \mathbf{e}$ was the codeword transmitted.

Example 5.2. Consider a linear code $C = \{00000, 11100, 00111, 11011\}$ over $\mathbb{Z}_2$, $\dim(C)$=2. We first write down a *standard array* of C as

$$\begin{aligned}
&00000 + C : 00000, 11100, 00111, \underline{11011}\\
&10000 + C : 10000, 01100, 10111, 01011\\
&01000 + C : 01000, 10100, 01111, 10011\\
&00100 + C : 00100, 11000, 00011, 11111\\
&00010 + C : 00010, 11110, 00101, 11001\\
&\underline{00001} + C : 00001, 11101, 00110, \underline{11010}\\
&10010 + C : 10010, 01110, 10101, 01001\\
&10001 + C : 10001, 01101, 10110, 01010
\end{aligned}$$

Each row in the standard array are the vectors in a coset. The first vector is the one with smallest Hamming weight in the corresponding coset, and is called the *coset leader*.

Standard array decoding: We declare that the coset leader of the coset containing the received word as the error pattern, and the transmitted codeword is the difference between the received word and the associated coset leader.

For example, if the received word $\mathbf{u} = 11010$, then we decode as follows: the error pattern $\mathbf{e}$ is 00001, and transmitted codeword is $\mathbf{u} - \mathbf{e} = 11011$, the the vector in the first row of the standard array lying above 11010.

We note that in general there are more than one way to write down a standard array. Whenever there are two or more vectors with the smallest Hamming weight in a coset, we can choose one of them arbitrarily as the coset leader. For example, the following is also a standard array of C in Example 5.2,

$$\begin{aligned}
&00000 + C : 00000, 11100, 00111, 11011\\
&10000 + C : 10000, 01100, 10111, 01011\\
&01000 + C : 01000, 10100, 01111, 10011\\
&00100 + C : 00100, 11000, 00011, 11111\\
&00010 + C : 00010, 11110, 00101, 11001\\
&00001 + C : 00001, 11101, 00110, 11010\\
&10010 + C : 10010, 01110, 10101, 01001\\
&01010 + C : 01010, 10110, 01101, 10001
\end{aligned}$$

The coset leader of the last coset is 01010.

Exercises:

1. Recall that a relation $\sim$ on a set E is called an *equivalence relation* if it satisfies

   (i) for all $x \in E$, we have $x \sim x$;

   (ii) for all $x, y \in E$, $x \sim y$ implies $y \sim x$;

   (iii) for all $x, y, z \in E$, $x \sim y$ and $y \sim z$ implies $x \sim z$;

   For a fixed integer m, show that the relation $x \sim y$ on the set of integers defined by $x \equiv y \mod m$ is an equivalence relation, and integers with the same remainder after division by m form an equivalence class.

For a given linear code C in $\mathbb{Z}_p^n$, show that the relation $\mathbf{u} \sim \mathbf{v}$ defined by $\mathbf{u} - \mathbf{v} \in C$ is an equivalence relation, and the cosets of the code C are the equivalence classes. (This proves parts (i) and (ii) in **Theorem 5.1**.)

2. Let C be a linear code over $\mathbb{Z}_p$ of dimension k and length n, and H be a parity-check matrix of C. We define the *syndrome* of a vector $\mathbf{u} \in \mathbb{Z}_p^n$ by the vector-matrix product $\mathbf{u}H^T$. Show that the vector in a coset of C has the same syndrome. Thus, the syndrome is indeed a function from the collection of cosets of C to the vector space $\mathbb{Z}_p^{n-k}$. Show that this is a bijection between the cosets and the vectors in $\mathbb{Z}_p^{n-k}$, i.e., no two cosets has the same syndrome, and every vector in $\mathbb{Z}_p^{n-k}$ is the syndrome of some coset of C.

3. Using the first standard array in p.5, decode the word (i) $\mathbf{y} = 11110$, (ii) $\mathbf{y} = 01101$.

   Using the second standard array in p.5, decode the word (i) $\mathbf{y} = 11110$, (ii) $\mathbf{y} = 01101$.

4. We index the components of a linear code C of length n by $1, 2, \ldots, n$. A collection of indices is called an *information set* of C if we can determine the codeword uniquely from these indices. It is easy to see that the size of an information set equals the dimension of C. If a generator matrix G of C is given, then a set of indices $\mathcal{I}$ is an information set if and only if the columns of G with indices in $\mathcal{I}$ form a square non-singular matrix. For instance, the information sets of the linear code in Example 4.1 are $\{1,3\}$, $\{1,4\}$, $\{1,5\}$, $\{2,3\}$, $\{2,4\}$, $\{2,5\}$, $\{3,4\}$, and $\{3,5\}$.

   Show that if $\mathcal{I}$ is an information set of C, then the complement of $\mathcal{I}$ in $\{1, 2, \ldots, n\}$ is an information set of $C^\perp$.