CHAPTER II

PRELIMINARIES

The standard terminologies and properties for linear codes are cursorily given in section 2.1. In section 2.2, a linear error-block code which is a generalization of the classical linear error-correcting code is introduced. The terminologies used for linear error-block codes follow those which are initiated by K. Feng, L. Xu and F. J. Hickernell in [3].

2.1 Linear Codes

Let $\mathbb{F}_q^n$ denote the vector space of all $n-$tuples over the finite field $\mathbb{F}_q$. The vector $(v_1, v_2, \ldots, v_n)$ in $\mathbb{F}_q^n$ is written in the form $v_1v_2 \ldots v_n$ and called a **word of length** n. Each v_i in the word $v_1v_2 \ldots v_n$ is called a **digit**.

$\mathbb{F}_q^n$ is also a metric space under the Hamming distance d_H. The **Hamming distance** $d_H(u, v)$ between words $u = u_1u_2 \ldots u_n$ and $v = v_1v_2 \ldots v_n$ is defined to be the number of digits which u and v differ. In other words,

$$d_H(u, v) = d_H(u_1, v_1) + d_H(u_2, v_2) + \cdots + d_H(u_n, v_n),$$

where

$$d_H(u_i, v_i) = \begin{cases} 1 & \text{if } u_i \neq v_i, \\ 0 & \text{if } u_i = v_i. \end{cases}$$

For each $v \in \mathbb{F}_q^n$, the **Hamming weight** of v, denoted $w_H(v)$, is defined to be the number of nonzero digits of v. That is, for $v = v_1v_2 \ldots v_n$,

$$w_H(v) = w_H(v_1) + w_H(v_2) + \cdots + w_H(v_n),$$

where

$$w_H(v_i) = \begin{cases} 1 & \text{if } v_i \neq 0, \\ 0 & \text{if } v_i = 0. \end{cases}$$

A **linear code $\mathcal{C}$ of length** n over $\mathbb{F}_q$ is defined to be a subspace of the vector space $\mathbb{F}_q^n$. An element v in $\mathcal{C}$ is called a **codeword**. If the dimension of $\mathcal{C}$ is k, $\mathcal{C}$ is said to be an $[n, k]$ **(linear) code** over $\mathbb{F}_q$.

The **minimum Hamming distance** of a linear code $\mathcal{C}$ containing at least two codewords, denoted $d_H(\mathcal{C})$, is defined by

$$d_H(\mathcal{C}) = \min\{d_H(u, v) \mid u, v \in \mathcal{C}, u \neq v\}.$$

Similarly, the **minimum Hamming weight** of a linear code $\mathcal{C}$, denoted $w_H(\mathcal{C})$, is defined by

$$w_H(\mathcal{C}) = \min\{w_H(v) \mid v \in \mathcal{C} \setminus \{0\}\}.$$

To avoid the tedious computation of $d_H(\mathcal{C})$ when the dimension (size) of a code is large, the relationship between $d_H(\mathcal{C})$ and $w_H(\mathcal{C})$ is given.

Theorem 2.1.1. *If $\mathcal{C}$ is a linear code, then $d_H(\mathcal{C}) = w_H(\mathcal{C})$.*

A process of decoding is an algorithm to determine which codeword was sent when a word is received. Here, the nearest neighbor decoding rule will be applied. If a word r is received, r will be decoded to $c_r \in \mathcal{C}$ if the distance between r and c_r is minimal among all the codewords in $\mathcal{C}$.

A linear code $\mathcal{C}$ is said to be $t-$**error-correcting** if the nearest neighbor decoding is able to correct t or fewer errors. The minimum Hamming distance of a linear code plays very important role in its error-correcting capability. They are related as stated in the next theorem.

Theorem 2.1.2. *A linear code $\mathcal{C}$ is t-error-correcting if and only if*

$$d_H(\mathcal{C}) \geq 2t + 1.$$

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

As an $[n, k]$ code $\mathcal{C}$ is a $k-$dimensional subspace of the vector space $\mathbb{F}_q^n$, every codeword in $\mathcal{C}$ can be uniquely written as a linear combination of a fixed basis of $\mathcal{C}$. Each codeword in $\mathcal{C}$ can represent one information word, so each message $u \in \mathbb{F}_q^k$ is encoded to the word $u\mathbf{G}$ where $\mathbf{G}$ is a matrix whose rows form a basis for $\mathcal{C}$. Equivalently,

$$\mathcal{C} = \{u\mathbf{G} \mid u \in \mathbb{F}_q^k\}.$$

$\mathbf{G}$ is then called a **generator matrix** for $\mathcal{C}$.

Since a linear code is a subspace of the vector space $\mathbb{F}_q^n$, it is the kernel of a linear transformation. In particular, there is an $(n-k) \times n$ matrix $\mathbf{H}$, called a **parity-check matrix** for the $[n, k]$ code $\mathcal{C}$,

$$\mathcal{C} = \{v \in \mathbb{F}_q^n \mid \mathbf{H}v^T = 0\}.$$

Notice that the rows of $\mathbf{H}$ will also be linear independent.

The minimum Hamming distance of a linear code can be determined via its parity-check matrix as follows :

Theorem 2.1.3. *Let $\mathcal{C}$ be a linear code over $\mathbb{F}_q$ with parity-check matrix $\mathbf{H}$. Then $d_H(\mathcal{C}) = d$ if and only if any $d-1$ columns of $\mathbf{H}$ are linearly independent and $\mathbf{H}$ has d columns that are linearly dependent.*

2.2 Linear Error-Block Codes

For a given positive integer n, $\pi = [n_1][n_2]\dots[n_s]$ is called a **partition** of n if $n = n_1 + n_2 + \cdots + n_s$ with $n_1 \geq n_2 \geq \cdots \geq n_s$. In case

$$\pi = \underbrace{[m_1]\dots[m_1]}_{l_1 \text{ copies}}\underbrace{[m_2]\dots[m_2]}_{l_2 \text{ copies}}\cdots\underbrace{[m_r]\dots[m_r]}_{l_r \text{ copies}},$$

we write $\pi = [m_1]^{l_1}[m_2]^{l_2}\dots[m_r]^{l_r}$ where $m_1 > m_2 > \cdots > m_r$.

Given a partition $\pi = [n_1][n_2]\dots[n_s]$ of n, the vector space $\mathbb{F}_q^n$ over $\mathbb{F}_q$ can be viewed as a product of $\mathbb{F}_q^{n_i}$'s, i.e.,

$$\mathbb{F}_q^n = \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2} \times \cdots \times \mathbb{F}_q^{n_s}.$$

Each word $v \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2} \times \cdots \times \mathbb{F}_q^{n_s}$ can be uniquely written as $v = v_1 v_2 \dots v_s$ where $v_i \in \mathbb{F}_q^{n_i}$ and each v_i in the word $v_1 v_2 \dots v_s$ is called a **block** in v.

In 2006, K. Feng, L. Xu and F. J. Hickernell [3] initiated a concept of $\pi-$distance which is a natural generalization of the Hamming distance. For a given partition $\pi = [n_1][n_2]\dots[n_s]$ of a positive integer n, the $\pi-$**distance** between $u = u_1 u_2 \dots u_s$ and $v = v_1 v_2 \dots v_s$ in $\mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2} \times \cdots \times \mathbb{F}_q^{n_s}$ is defined to be the number of blocks which u and v differ. In other words,

$$d_\pi(u, v) = d_{[n_1]}(u_1, v_1) + d_{[n_2]}(u_2, v_2) + \cdots + d_{[n_s]}(u_s, v_s)$$

where

$$d_{[n_i]}(u_i, v_i) = \begin{cases} 1 & \text{if} \quad u_i \neq v_i \in \mathbb{F}_q^{n_i}, \\ 0 & \text{if} \quad u_i = v_i \in \mathbb{F}_q^{n_i}. \end{cases}$$

Clearly, d_π forms a metric on $\mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2} \times \cdots \times \mathbb{F}_q^{n_s}$.

Similarly, $\pi-$**weight** of v in $\mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2} \times \cdots \times \mathbb{F}_q^{n_s}$ is defined to be the number of nonzero blocks in v. That is, for $v = v_1 v_2 \dots v_s$,

$$w_\pi(v) = w_{[n_1]}(v_1) + w_{[n_2]}(v_2) + \cdots + w_{[n_s]}(v_s)$$

where

$$w_{[n_i]}(v_i) = \begin{cases} 1 & \text{if } v_i \neq 0 \in \mathbb{F}_q^{n_i}, \\ 0 & \text{if } v_i = 0 \in \mathbb{F}_q^{n_i}. \end{cases}$$

Remark 2.2.1. For given $u, v \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2} \times \cdots \times \mathbb{F}_q^{n_s}$ and $\lambda \in \mathbb{F}_q \setminus \{0\}$, the following relations can be easily shown :

i) $w_\pi(u) - w_\pi(v) \leq w_\pi(u+v) \leq w_\pi(u) + w_\pi(v)$,

ii) $w_\pi(u-v) = d_\pi(u,v)$,

iii) $w_\pi(v) = w_\pi(\lambda v)$, and

iv) $w_\pi(\lambda u + v) = w_\pi(u + \lambda^{-1}v)$.

For each partition $\pi = [n_1][n_2]\dots[n_s]$ of n, a **linear error-block code with type π over $\mathbb{F}_q$** is defined to be a subspace $\mathcal{C}$ of $\mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2} \times \cdots \times \mathbb{F}_q^{n_s}$.

A $k-$dimensional linear error-block code with type π over $\mathbb{F}_q$ is shortly called an $[n,k;\pi]$ **(linear error-block) code** over $\mathbb{F}_q$.

Remark 2.2.2. A linear code of length n is a linear error-block code with type $[1]^n$.

The **minimum $\pi-$distance** of a linear error-block code $\mathcal{C}$, denoted $d_\pi(\mathcal{C})$, is defined by

$$d_\pi(\mathcal{C}) = \min\{d_\pi(u,v) \mid u,v \in \mathcal{C},\ u \neq v\}.$$

Similarly, the **minimum $\pi-$weight** of a linear error-block code $\mathcal{C}$, denoted $w_\pi(\mathcal{C})$, is defined by

$$w_\pi(\mathcal{C}) = \min\{w_\pi(v) \mid v \in \mathcal{C} \setminus \{0\}\}.$$

To rapidly compute $d_\pi(\mathcal{C})$ when the size of code is large, the relationship between $d_\pi(\mathcal{C})$ and $w_\pi(\mathcal{C})$ is given.

Theorem 2.2.1. *If $\mathcal{C}$ is a linear error-block code with type π, then*

$$d_\pi(\mathcal{C}) = w_\pi(\mathcal{C}).$$

An $[n.k;\pi]$ code $\mathcal{C}$ with minimum $\pi-$distance d will be denoted by $[n,k,d;\pi]$ **code** where n,k,π and d are called **parameters** of the linear error-block code $\mathcal{C}$.

A linear error-block code $\mathcal{C}$ is said to be **t-error-block-correcting** if the nearest neighbor decoding is able to correct t or fewer error-blocks.

The minimum $\pi-$distance of a linear error-block code performs its error-correction.

Theorem 2.2.2. *A linear error-block code $\mathcal{C}$ is t-error-block-correcting if and only if $d_\pi(\mathcal{C}) \geq 2t+1$.*

As in the classical case, a generator matrix for an $[n, k; \pi]$ code $\mathcal{C}$ is also defined to be any $k \times n$ matrix $\mathbf{G}$ whose rows form a basis for $\mathcal{C}$. Since $\mathcal{C}$ is an error-block code, a generator matrix for $\mathcal{C}$ can be viewed in blocks depending on the partition π as follows : Let $\pi = [n_1][n_2]\cdots[n_s]$. A **generator matrix** for $\mathcal{C}$ is a matrix $[\ G_1 \ \ G_2 \ \ \cdots \ \ G_s\ ]$ where each G_i is a $k \times n_i$ matrix and

$$[\ G_1 \ \ G_2 \ \ \cdots \ \ G_s\ ] = \mathbf{G}.$$

The G_i is called the i^{th} **block** in $\mathbf{G}$.

A parity-check matrix for $\mathcal{C}$ is defined in term of a parity-check matrix for a linear code in the same fashion. That is for a parity-check matrix $\mathbf{H}$ for $\mathcal{C}$ considered as a linear code, $[\ H_1 \ \ H_2 \ \ \cdots \ \ H_s\ ]$ is a **parity-check matrix** for the linear error-block code $\mathcal{C}$ if each H_i is an $(n-k) \times n_i$ matrix and $[\ H_1 \ \ H_2 \ \ \cdots \ \ H_s\ ] = \mathbf{H}$. Each H_i is called the i^{th} **block** in $\mathbf{H}$. Consequently,

$$0 = \mathbf{H}\mathbf{G}^T = H_1{G_1}^T + H_2{G_2}^T + \cdots + H_s{G_s}^T.$$

As we study parity-check matrix in blocks, the set of blocks $H_{i_1}, H_{i_2}, \ldots, H_{i_l}$ in $\mathbf{H} = [\ H_1 \ \ H_2 \ \ \cdots \ \ H_s\ ]$ is said to be **linearly independent** if the union of all column vectors in $H_{i_1}, H_{i_2}, \ldots, H_{i_l}$ is a linearly independent set. Otherwise, we say that the set of blocks $H_{i_1}, H_{i_2}, \ldots, H_{i_l}$ is **linearly dependent**.

Remark 2.2.3. *i)* Blocks $H_{i_1}, H_{i_2}, \ldots, H_{i_l}$ in $\mathbf{H}$ are linearly independent if

$$H_{i_1}{v_{i_1}}^T + H_{i_2}{v_{i_2}}^T + \cdots + H_{i_l}{v_{i_l}}^T = 0$$

implies that $v_{i_j} = 0 \in \mathbb{F}_q^{n_{i_j}}$ for $j = 1, 2, \ldots, l$.

ii) Blocks $H_{i_1}, H_{i_2}, \ldots, H_{i_l}$ in $\mathbf{H}$ are linearly dependent if there is $r \in \{1, 2, \ldots, l\}$ such that $v_{i_r} \neq 0 \in \mathbb{F}_q^{n_{i_r}}$ and

$$H_{i_1}v_{i_1}^T + \cdots + H_{i_r}v_{i_r}^T + \cdots + H_{i_l}v_{i_l}^T = 0$$

where $v_{i_j} \in \mathbb{F}_q^{n_{i_j}}$ for $j = 1, 2, \ldots, l$.

The minimum $\pi-$distance of a linear error-block code is determined using a parity-check matrix as follows :

Theorem 2.2.3. *Let $\mathcal{C}$ be an $[n, k; \pi]$ code over $\mathbb{F}_q$ with type $\pi = [n_1][n_2]\ldots[n_s]$ and $\mathbf{H} = [\ H_1 \quad H_2 \quad \cdots \quad H_s\ ]$ parity-check matrix for $\mathcal{C}$. Then $d_\pi(\mathcal{C}) = d$ if and only if any $d-1$ blocks of $\mathbf{H}$ are linearly independent and $\mathbf{H}$ has d blocks that are linearly dependent.*

Proof. It suffices to show the following two assertions :

i) $d_\pi(\mathcal{C}) \geq d$ if and only if any $d-1$ blocks in $\mathbf{H}$ are linearly independent.

ii) $d_\pi(\mathcal{C}) \leq d$ if and only if $\mathbf{H}$ has d blocks that are linearly dependent.

$\mathcal{C}$ contains a nonzero word $v = v_1 v_2 \ldots v_s$ of $\pi-$weight l (where $0 \neq v_{i_j} \in \mathbb{F}_q^{n_{i_j}}$ for $j = 1, 2, \ldots, l$ and $v_{i_j} = 0 \in \mathbb{F}_q^{n_{i_j}}$ otherwise) if and only if

$$\begin{aligned} 0 = \mathbf{H}v^T &= H_1v_1^T + H_2v_2^T + \cdots + H_sv_s^T \\ &= H_{i_1}v_{i_1}^T + H_{i_2}v_{i_2}^T + \cdots + H_{i_l}v_{i_l}^T \end{aligned}$$

which is true if and only if there are l blocks in $\mathbf{H}$ that are linearly dependent, say $H_{i_1}, H_{i_2}, \ldots, H_{i_l}$ are linearly dependent.

Thus $d_\pi(\mathcal{C}) \geq d$ if and only if $\mathcal{C}$ does not contain any nonzero words of $\pi-$weight $\leq d-1$ if and only if $\leq d-1$ blocks in $\mathbf{H}$ are linearly independent. This proves i).

Similarly, $d_\pi(\mathcal{C}) \leq d$ if and only if $\mathcal{C}$ contains a nonzero word of $\pi-$weight $\leq d$ if and only if $\mathbf{H}$ has $\leq d$ blocks (and hence d blocks) that are linearly dependent. This proves ii). □

Theorem 2.2.2 indicates that error-correcting capacity depends deeply on how the length of code words is partitioned. Next example illustrates this fact.

Example 2.2.1. Consider a parity-check matrix

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

of a linear code $\mathcal{C}$. By Theorems 2.1.2 and 2.1.3, $\mathcal{C}$ is a $1-$error correcting linear code of length 7. Thus $\mathcal{C}$ cannot correct any $2-$errors. But by viewing $\mathcal{C}$ as an $[7, 3; [2]^2[1]^3]$ codes, it can correct $2-$errors which occur in the 1^{st} and 2^{nd} positions even though this new code is also $1-$error-block-correcting code. This is because they are in the same first block. Similarly, it can also correct $2-$errors which occur in the 3^{rd} and 4^{th} positions.