0.1 Introduction

Definition 0.1. A ***linear code*** C over $GF(q)$ is a subspace of $V(n, q)$.

Notation: If linear code C is a k-dimensional subspace of $V(n, q)$ with minimum distance d, C is denoted by $[n, k]$-code or $[n, k, d]$-code depending on the context.

The ***weight*** $\mathrm{w}(\mathbf{x})$ of a vector $\mathbf{x}$ in $V(n, q)$ is the number of non-zero entries of $\mathbf{x}$.

Lemma 0.2. If $\mathbf{x}$ and $\mathbf{y} \in V(n, q)$ then

$$d(\mathbf{x}, \mathbf{y}) = w(\mathbf{x} - \mathbf{y}).$$

Proof. $\mathbf{x} - \mathbf{y}$ will be non-zero exactly at those places where $\mathbf{x}$ and $\mathbf{y}$ differ. □

Theorem 0.3. Let C be a linear code and let $w(C)$ be the smallest of the non-zero weights. Then $d(C) = w(C)$.

Proof. Let $\mathbf{x}$ and $\mathbf{y}$ be codewords with distance $d(C)$. Then by Lemma 0.2,

$$d(C) = w(\mathbf{x} - \mathbf{y}) \geq w(C)$$

as $\mathbf{x} - \mathbf{y} \in C$. Let $\mathbf{z}$ be a codeword such that $w(\mathbf{z}) = w(C)$. Since $\mathbf{0} \in C$, by Lemma 0.2 again,

$$w(C) = w(\mathbf{z} - \mathbf{0}) = d(\mathbf{z}, \mathbf{0}) \geq d(C).$$

Therefore, $d(C) = w(C)$. □

This theorem allows us to calculate $d(C)$ in M iterations, which otherwise would have been $\binom{M}{2} = \frac{M(M-1)}{2}$, one of the reasons we focus on linear codes. One more reason being that every codeword can be expressed in terms of k codewords, in the case of $[n, k]$-code, we don't have to list all the codewords.

Definition 0.4. A $k \times n$ matrix whose rows form a basis of a linear $[n, k]$-code is called a ***generator matrix*** of the code.

Definition 0.5. Two linear codes are said to be ***equivalent*** if one can be obtained from another by combinations of operations of the types,

(i) permuations of the positions of the code.
(ii) multiplication of the symbols of fixed position by a non-zero scalar.

Theorem 0.6. Two $k \times n$ matrices generate equivalent linear $[n, k]$-codes over $GF(q)$ if one can be obtained from other by a sequence of these operations.

(i) Permutations of the rows.
(ii) Multiplication of a row by a non-zero scalar.
(iii) Addition of the scalar multiple of one row to another.
(iv) Permutations of the columns.
(v) Multiplication of any column by a non-zero scalar.

Proof. The first three operations are called ***row operations***, and they will preserve the linear independence of the row vectors and will replace one basis by another of the same code. Last two are called ***column operations***, they convert the basis into basis of a equivalent code. □

Theorem 0.7. Let G be a generator matrix of an $[n, k]$-code. Then by operations described in Theorem 0.6, G can be converted into ***standard form***

$$[I_k \mid A],$$

where I_k is the $k \times k$ identitiy matrix, and A is a $k \times (n - k)$ matrix.

Proof. Let $G = [g_{ij}]$ and $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_k$ and $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ be its rows and columns respectively, then the following scheme will convert it to standard form.
Suppose G has already been transformed to

$$\begin{bmatrix} 1 & \cdots & 0 & g_{1j} & \cdots & g_{1n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & g_{j-1,j} & \cdots & g_{j-1,n} \\ 0 & \cdots & 0 & g_{jj} & \cdots & g_{jn} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & g_{kj} & \cdots & g_{kn} \end{bmatrix}$$

where $j \in \{1, 2, \ldots, k\}$.

1. If $g_{jj} \neq 0$, proceed to next step. If $g_{ij} = 0$, if for some $i > j$, $g_{ij} \neq 0$, replace $\mathbf{r}_j$ with $\mathbf{r}_i$. If not, replace $\mathbf{c}_j$ with $\mathbf{c}_p$, where $g_{jp} \neq 0$. (We could have sone the column step directly without looking for the row first, but we are trying that the matrix remains basis of the original code C, if possible.)
2. Multiply $\mathbf{r}_j$ with g_{jj}^{-1}, so that g_{jj} becomes 1.
3. For each $i \in \{1, 2, \ldots, k\}, i \neq j$, convert $\mathbf{r}_i \to \mathbf{r}_i - g_{ij}\mathbf{r}_j$.

Now after these steps, we have converted $\mathbf{c}_j$ to the the required form. So repeating for all values of j will give the standard form. □

Note: *Standard form* of a generator matrix is not unique, for instance we can interchange the columns of A and still satisfy all the conditions. Also, the standard form will remain generator matrix of the original code C iff we don't use the column operations. It is possible if and only if the first k columns of generator matrix are *linearly independent*.

0.2 Encoding with a linear code

Let C be an $[n, k]$-code over $GF(q)$ with generator matrix G. It can be used to communicate q^k distinct messages. If the rows of G are $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_k$, then we *encode* a message vector $\mathbf{u} = u1u2 \cdots u_k \in V(k, q)$ by a function that maps $V(k, q) \to k$-dimensional subspace of $V(n, q)$ (the code C).

$$\mathbf{u}G = \sum_{i=1}^{k} u_i \mathbf{r}_i$$

In particular, if $G = [I_k \mid A]$ (standard form, $A = [a_{ij}]$), then encoding is

$$\mathbf{x} = \mathbf{u}G = x_1 x_2 \cdots x_k \cdots x_n,$$

where $x_i = u_i$, $1 \leq i \leq k$ are ***message digits***, and

$$x_{k+i} = \sum_{j=1}^{k} a_{ij} u_j \quad 1 \leq i \leq n - k,$$

are the ***check digits***. The check digits represent *redundancy* added to protect against noise.

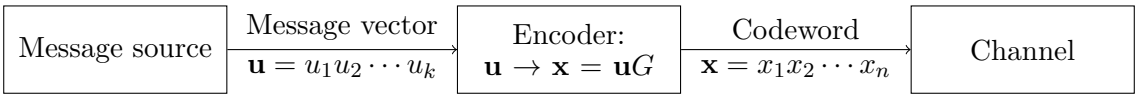

Figure 1: Encoding a message vector.

0.3 Decoding with a linear code

Suppose the codeword sent is $\mathbf{x} = x_1x_2\cdots x_n$ and the recieved vector is $\mathbf{y} = y_1y_2\cdots y_n$. We define ***error vector*** $\mathbf{e}$ to be

$$\mathbf{e} = \mathbf{y} - \mathbf{x} = e_1e_2\cdots e_n.$$

Definition 0.8. Suppose that C is an $[n,k]$-code over and $\mathbf{a}$ is any vector in $V(n,q)$. Then set

$$\mathbf{a} + C = \{\mathbf{a} + \mathbf{x} \mid \mathbf{x} \in C\}$$

is called a ***coset*** of C.

Lemma 0.9. Suppose that $\mathbf{x} + C$ is a coset of C and $\mathbf{b} \in \mathbf{a} + C$. Then $\mathbf{b} + C = \mathbf{a} + C$.

Proof. Since $\mathbf{b} \in \mathbf{a} + C$, we have $\mathbf{b} = \mathbf{a} + \mathbf{x}$ for some $\mathbf{x} \in C$. Now if $\mathbf{b} + \mathbf{y} \in \mathbf{b} + C$, then

$$\mathbf{b} + \mathbf{y} = (\mathbf{a} + \mathbf{x}) + \mathbf{y} = \mathbf{a} + (\mathbf{x} + \mathbf{y}) \in \mathbf{a} + C.$$

Hence, $\mathbf{b} + C \subseteq \mathbf{a} + C$. Similarly, if $\mathbf{a} + \mathbf{z} \in \mathbf{a} + C$, then

$$\mathbf{a} + \mathbf{z} = (\mathbf{b} - \mathbf{x}) + \mathbf{z} = \mathbf{b} + (\mathbf{z} - \mathbf{x}) \in \mathbf{b} + C.$$

Hence, $\mathbf{a} + C \subseteq \mathbf{b} + C$. Therefore, $\mathbf{b} + C = \mathbf{a} + C$. □

Theorem 0.10. Suppose C is an $[n,k]$-code over $GF(q)$. Then

(i) every vector of $V(n,q)$ is in some coset of C,
(ii) every coset contains exactly q^k vectors,
(iii) two cosets are either disjoint or coincide.

Proof.

(i) If $\mathbf{a} \in V(n,q)$, then $\mathbf{a} = \mathbf{a} + 0 \in \mathbf{a} + C$.
(ii) It can be easily seen that the mapping $C \to \mathbf{a} + C$ is one-one. Therefore, $|C| = |\mathbf{a} + C|$.
(iii) Suppose cosets $\mathbf{a} + C$ and $\mathbf{b} + C$ overlap and $\mathbf{v} \in (\mathbf{a} + C) \cap (\mathbf{b} + C)$. Thus, for some $\mathbf{x}, \mathbf{y} \in C$,

$$\begin{aligned} &\mathbf{v} = \mathbf{a} + \mathbf{x} = \mathbf{b} + \mathbf{y} \\ \implies &\mathbf{a} = \mathbf{b} + \mathbf{y} - \mathbf{x} \\ \implies &\mathbf{a} \in \mathbf{b} + C \end{aligned}$$

Therefore, by Lemma 0.9, $\mathbf{a} + C = \mathbf{b} + C$.

□

Definition 0.11. The vector having minimum weight in a coset is called the ***coset leader***. (If there are more than one vectors with minimum weight then we choose anyone.)

Theorem 0.10 implies:

$$V(n,q) = (\mathbf{0} + C) \cup (\mathbf{a}_1 + C) \cup \cdots \cup (\mathbf{a}_s + C)$$

where $s = q^{n-k} - 1$, and by Lemma 0.9, we can take $\mathbf{a}_i$ to be the coset leaders.

Definition 0.12. A ***standard array*** for an $[n,k]$-code C is a $q^{n-k} \times q^k$ array of all the vectors of $V(n,q)$. First row will consist of code C in any order except $\mathbf{0}$ which will be in in the first column. Following rows will be cosets with coset leaders being the first element of each row, remaining elements in each row will be sum of its coset leader with corresponding codeword in the same column.

For example, let C be a binary $[4,2]$-code with generator matrix

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

then $C = \{0000, 1011, 0101, 1110\}$.
The standard array of C will be

codewords →	0000	1011	0101	1110
	1000	0011	1101	0110
	0100	1111	0001	1010
	0010	1001	0111	1100
	↑			
	coset leaders			

Decoding scheme using standard array: When $\mathbf{y}$ is recieved its position is found in array. Then the *decoder* assumes that the error vector $\mathbf{e}$ is the coset leader, and $\mathbf{y}$ is decoded as $\mathbf{x} = \mathbf{y} - \mathbf{e}$ at the top of the column containing $\mathbf{y}$. By choosing the minimum weight vector as the coset leader, we ensure that standard array decoding scheme is a nearest neighbour decoding scheme.

Theorem 0.13. Let C be a binary $[n, k]$-code, and for $i \in \{1, 2, \ldots, n\}$, let α_i denote the number of cosets leaders of weight i. Then *probability* $P_{\text{corr}}(C)$ of correctly recognizing the codeword is given by

$$P_{\text{corr}}(C) = \sum_{i=0}^{n} \alpha_i p^i (1-p)^{n-i}.$$

Proof. The probability of error code to be vector of weight i is $p^i(1-p)^{n-i}$. Therefore, the result follows directly from the definition of decoding scheme. □

Note: If $d(C) = 2t + 1$ or $2t + 2$, then $\alpha_i = \binom{n}{i}$ for $0 \geq i \geq t$. In particular, if C is a perfect code, in addition to the general result, $\alpha_i = 0$ for $i > t$ also holds.

Theorem 0.14. Let C be a binary $[n, k]$-code, and let A_i denote the number of codewords of weight i. Then *probability* $P_{\text{corr}}(C)$ of an error going undetected is given by

$$P_{\text{undetec}}(C) = \sum_{i=1}^{n} A_i p^i (1-p)^{n-i}.$$

Proof. The error will only go undetected if and only if $\mathbf{y} - \mathbf{x}$ is a non-zero codeword, where $\mathbf{y}$ and $\mathbf{x}$ are received and sent vectors respectively. The probability of $\mathbf{y} - \mathbf{x} = \mathbf{z} \in C$ is $p^{w(\mathbf{z})}(1-p)^{n-w(\mathbf{z})}$, thus summing over non-zero $\mathbf{z}$ gives the result. □

Definition 0.15. For a linear $[n, k]$-code C, ***rate*** is defined as the ratio of number of message symbols to the total number of symbols sent, i.e. $R(C) = \dfrac{k}{n}$.

Good code generally will have high rate.

Definition 0.16. The ***capacity*** $\mathscr{C}(p)$ of a binary symmetric channel with symbol error probability p is

$$\mathscr{C}(p) = 1 + p \log_2 p + (1-p) \log_2 (1-p).$$

Theorem 0.17. Shannon's Theorem Suppose a channel is binary symmetric with symbol error probability p. Let $R \in \mathbb{R}$ satisfying $R < \mathscr{C}(p)$. Then for all $\epsilon > 0$, there exists, for some large n, an $[n, k]$-code C of rate $\dfrac{k}{n} \geq R$ such that $P_{\text{err}}(C) < \epsilon$. (where $P_{\text{err}}(C) = 1 - P_{\text{corr}}(C)$)