MBNA Publishing House Constanta 2025

Proceedings of the International Scientific Conference SEA-CONF

SEA-CONF PAPER • OPEN ACCESS

On the Properties of Linear Block Codes

ISSN: 2457-144X; ISSN-L: 2457-144X

doi: 10.21279/2457-144X-25-002

On the Properties of Linear Block Codes

Bianca-Liana Bercea (Straton)

Ovidius University from Constanta
bianca.bercea@365.univ-ovidius.ro

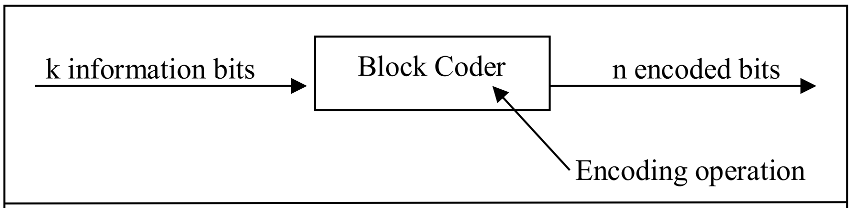
Abstract. *Linear codes* are algebraic codes, typically over a finite field, where the sum of two codewords is always a codeword and the multiplication of a codeword by a field element is also a codeword. Linear codes that are also block codes are *linear block codes*. These codes are used for error control coding, satellite and deep space communications, and they are used for magnetic and optical data storage in hard disks and magnetic tapes and single error correcting and double error correcting code used to improve semiconductor memories. An advantage of linear block codes is that they are easiest to detect and correct errors. Another advantage is that extra parity bit does not convey any information but detects and correct errors. Among the disadvantages of linear block codes are that the transmission bandwidth is more and that extra bit reduces the bit rate of transmitter and also its power. Cyclic codes are special linear block codes with one extra property. In a cyclic code, if a codeword is cyclically shifted (rotated), the result is another codeword.

1. Theoretical considerations

Definition 1. Let F be a finite set called alphabet, having $q = card(F) = |F|$ elements. A subset $C \neq \emptyset, C \subset F^n$ is called code over the alphabet F. The elements of F^n are called words, the elements of C are called codewords, and n is the length of the code. For $q = 2$ the code is called binary.

Definition 2. A block code is said to be linear code if its codewords satisfy the condition that the sum of any two codewords gives another codeword, i.e. $c_p = c_i + c_k$.

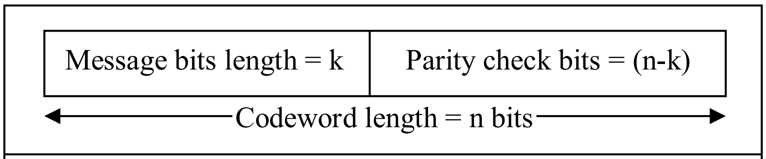
Example 3. The block code $\{0000, 0101, 1010, 1111\}$ is linear.
Indeed, if we consider the codewords 0101 and 1010, then their sum $0101 + 1010 = 1111$ gives us a codeword in the block code.



The general representation of the codeword

This representation of the codeword will consist in two section: first k bits and second part contains (n-k) bits. So, n-digit codeword made up of k information digits and (n-k) redundant parity check digits. The rate of efficiency for this code is $r = \frac{k}{n} = \frac{number\ of\ information\ bits}{total\ number\ of\ bits\ in\ codeword}$.

A linear block code is a linear combination of parity bits and message bits. First portion of k bits is always identical to the message sequence to be transmitted and the second portion of (n-k) bits are computed from message bits according to the encoding rule and is called parity bits.



Structure of a codeword

Remark 4. In linear block code, each block containing k messages bits is encoded into a block of n bits by adding (n-k) parity check bits.

Consider a pair of code vectors $x = (x_1, \dots, x_n) \in F^n$ and $y = (y_1, \dots, y_n) \in F^n$ that have the same number of elements.

Definition 5. Hamming distance is an application $d: F^n \times F^n \to \mathbb{R}$ given by

$$d(x, y) = |\{i \mid x_i \neq y_i\}|,$$

$d(x, y)$ is defined as the number of locations in which their respective elements differ.

Example 6. Let's consider the linear block code $\{0000, 0101, 1010, 1111\}$. Let's calculate the Hamming distance.

$d(0000, 0101) = 2;$
$d(0101, 1010) = 4;$
d(1010, 1111) = 2.

Definition 7. Hamming weight $w(x)$ is defined as the number of elements in the code vector. In other words, the weight of a codeword is the number of nonzero elements.

Example 8. Let's consider the linear block code $\{0000, 0101, 1010, 1111\}$. Let's calculate the Hamming weight.

$w(0101) = 2;$
$w(1010) = 2;$
$w(1111) = 4.$

Definition 9. The minimum distance of a block code C is defined as

$$d_{min} = \min_{x,y \in C, x \neq y} d(x, y)$$

the smallest hamming distance between any pair of code vectors in the code or smallest hamming weight of the nonzero vectors in the code.

Efficient codes have both d_{min} and $card\ C$. It is very important to know the minimum distance of a code.

Example 10. For the linear block code $\{0000, 0101, 1010, 1111\}$, the minimum distance of codewords is 2.

Theorem 11. A code C with minimal distance $d = d_{min}$ can detect $d - 1$ errors and correct $\left[\frac{d-1}{2}\right]$ errors. Moreover, these margins are the best possible.

Proof.

Suppose that a word x is obtained from a word in code $c \in C$ by inserting at most $d - 1$ errors. So, $d(x, c) \leq d - 1$ which means that x cannot be a codeword because the minimum distance is d. On the other hand, if $d_{min} = d$, then there are words in code c, c' such that $d(c, c') = d$, that is, we could obtain c' from c by making exactly d errors. The code cannot detect such errors.

To prove the assertion about error correction, we first assume that $d = 2t + 1$. Let $c \in C$ and suppose that x is obtained from c after introducing t errors. Can x be obtained from another word c' in the code, with at most t errors? If that were to happen, we would have $d(c, c') \leq d(c, x) + d(x, c') \leq t + t =$

$2t$. But $c \neq c'$ and so $d(c,c') \geq d(C)$, from which it follows that $d(C) \leq 2t$, which contradicts the hypothesis. If c, c' were codewords such that $d(c,c') = 2t+1$, then making $t+1$ errors in c we could get c'. ■

Remark 12. An (n,k) linear block code has the powers to correct all error patterns of weight t or less if and only if $d(x_i, x_j) \leq 2t+1$. An (n,k) linear block code of minimum distance d_{min} can correct up to 1 error if and only if $t = \frac{d_{min}-1}{2}$.

2. Properties of linear block codes

i) The all zero codeword $[0,0,0,\dots,0]$ is a valid codeword.

ii) Given any 3 codewords c_i, c_j, c_k such that $c_k = c_i + c_j$ then the distance between two codewords will be equal to the weight of the codewords i.e. $d(c_i, c_j) = w(c_k)$.

Let's consider $c_i = 0101$, $c_j = 1010$.

$c_i + c_j = 0101 + 1010 = 1111 = c_k$.

$d(c_i, c_j) = 4$, $w(c_k) = 4$.

So, this block code follows the property of linear block codes.

iii) The minimum distance of a linear block code is equal to the minimum weight of its nonzero codewords.

$$\begin{aligned} d_{min} &= \min\{w_t(x+y) | x, y \in C, x \neq y\} \\ &= min\{w_t(v) | v \in C, v \neq 0\} \end{aligned}$$

Example 13. Let's consider (7,4)-Hamming code.

Let's choose the codewords $c_1 = 0001011$ and $c_{10} = 1010011$.

From the definition of linear block codes we have that:

$$\begin{aligned} c_1 + c_{10} &= 0001011 + 1010011 \\ &= 1011000 = c_{11} \end{aligned}$$

So, $c_{11} = c_1 + c_{10}$.

$c_0 = 0000000$ is a codeword in the block code.

Hamming distance for $c_1 = 0001011$ is 3 and Hamming distance for $c_{10} = 1010011$ is 4.

So, minimum Hamming distance for c_1 and c_{10} is $d(c_1, c_{10}) = 3$.

The weight of c_{11} is $w(c_{11}) = 3$.

So, the condition is satisfied: $d(c_1, c_{10}) = w(c_{11})$.

3. Coding and decoding of linear block codes

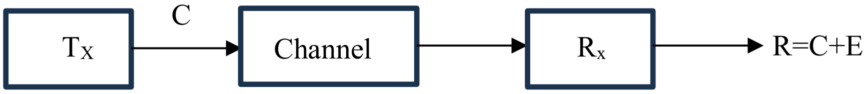
(i) Parity check matrix (H)

The parity check matrix is represented as

$$H = [P^t : I_{n-k}]$$

where P^t is the transpose parity matrix and I_{n-k} is the identity matrix of dimension $n-k$.

Parity check matrix is used to verify the codeword c generated by generator matrix.



where C is the codeword, R is the received codeword and E is the error vector.

(ii) Steps of coding and decoding

1) C is correct codeword if $C \cdot H^t = 0$.

2) Received codeword R is a combination of codeword C and an error vector E

$$R = C + E$$

3) The error syndrome is given as

$$S = R \cdot H^t$$

If $s = 0$, then the codeword C is valid.
If $s \neq 0$, then the codeword C is invalid.

Example 14. Consider a $(7,4)$ block code generated by the generator matrix

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

We'll find out the error vector and let's suppose that the received codeword R is 1001001.

Since we know that $G = [I_k : P]$, it results that $P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ and $P^t = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$.

The parity check matrix is given by $H = [P^t : I_{n-k}]$.
Since we know that number of message bit is $k = 4$ and the length of codeword is $n = 7$, it result that $n - k = 7 - 4 = 3$.

So, $H = \begin{bmatrix} 1 & 1 & 1 & 0 & : & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & : & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & : & 0 & 0 & 1 \end{bmatrix}$ and $H^t = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

First, we'll check if C is valid and for that we need the calculation of C.
Since number of message bit is $k = 4$, we have 2^4 combination of codeword that are possible.

$$C = [1\ 0\ 1\ 1] \cdot \begin{bmatrix} 1 & 0 & 0 & 0 & : & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & : & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & : & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & : & 0 & 1 & 1 \end{bmatrix} = [1\ 0\ 1\ 1\ 0\ 0\ 1].$$

Let's verify if $C \cdot H^t = 0$.

$$C \cdot H^t = [1\ 0\ 1\ 1\ 0\ 0\ 1] \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [0\ 0\ 0].$$

So, C is a valid codeword.
For the calculation of error syndrome S, we have $S = R \cdot H^t$, where $R = 1001001$.

$$S = [1\ 0\ 0\ 1\ 0\ 0\ 1] \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [1\ 0\ 1].$$

It means that third bit from the received codeword will have error.
So, the error vector is $E = 0010000$.
We can verify if the error vector is correct by the relation $C = R + E$.
$C = 1001001 + 0010000 = 1011001$.

References

[1] A. Atanasiu. Error correction and detection codes. University of Bucharest 2001.
[2] C. Gheorghe, D. Popescu. Criptografie. Coduri. Algoritmi. University of Bucharest 2005.