Introduction to Algebraic Coding Theory

Hashem Bordbar

Center for Information Technologies and Applied Mathematics,
The University of Nova Gorica, Slovenia.
Hashem.bordbar@ung.si

Ring Theory Seminar,
Institute for Mathematics and Scientific Computing,
University of Graz.

Contents

1. History
2. Preliminaries
3. Results

Claude Shannon

In 1948, Claude Shannon published his seminal paper, *"A Mathematical Theory of Communication"*, which marked the foundation of both information theory and coding theory.

C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, 1948, pp. 379–423. Available at:
`https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf`

Claude Shannon

Shannon defined **channel capacity** and proved that reliable communication is possible at any rate below it, even with potential distortion.
Shannon's results guarantee that data can be encoded before transmission and decoded with the desired accuracy, despite any alterations.

Channel

A fundamental characteristic of any communication **channel** is that information originates from a source and is transmitted through the channel to a receiver on the other end.

Channel

A fundamental characteristic of any communication **channel** is that information originates from a source and is transmitted through the channel to a receiver on the other end. Communication channels encompass a wide range of systems, including magnetic storage devices, compact discs, and electronic communication technologies such as cellular phones

Noise

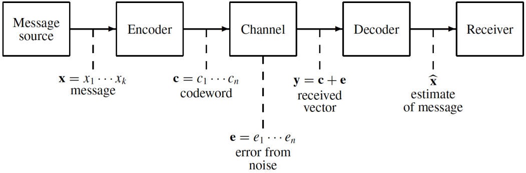
Take a compact disc as an example. The message could be music, speech, or data that's recorded on the disc. The disc itself acts as the channel, and the person listening is the receiver. But this channel isn't perfect, it's 'noisy,' meaning that what we read from the disc isn't always exactly what was written. For example, if the data is in binary form, a 0 might sometimes be read as a 1. That's why we need error-correcting codes to detect and fix these kinds of mistakes.

Coding Theory

Error-control codes are used to **detect** and **correct** errors that may occur when data is transmitted over noisy channels or stored on physical media. The study of these codes falls under the field of **coding theory.**

It is important to note that error-control coding represents just one component of the overall processing applied to messages before they are transmitted or stored.

Communication Channel



Overview of the Coding Process

The process begins with an **information source.** A **source encoder** transforms the source output into a sequence of symbols, which we denote as the message x.

This message is then encoded into a string of discrete symbols, usually binary, called the **codeword** c.

Due to noise in the channel or storage medium, an **error vector** e may be introduced, resulting in the **received vector**:

$$y = c + e$$

After decoding, we obtain a vector $\tilde{x}$, which is our estimate of the original message. The goal is to achieve:

$$\tilde{x} = x$$

ensuring the message is correctly recovered despite the presence of errors.

The ISBN Code

The International Standard Book Number (ISBN) is a coding system used globally by publishers to identify key properties of each book.

The first nine digits (from 0 to 9) encode information such as the language, publisher, and title. To guard against errors, a tenth digit is added as a **check digit**, forming a ten-digit codeword.

The ten-digit string $x_1 x_2 \cdots x_{10}$ is chosen to satisfy the condition:

$$\sum_{i=1}^{10} i \cdot x_i \equiv 0 \pmod{11}$$

If the computed check digit $x_{10} = 10$, it is represented by the letter X.

The ISBN Code

The ISBN code can:

- Detect any single-digit error,
- Detect any transposition of two digits.

The ISBN code's error-detecting capabilities can be compared with other check-digit systems used in practice, such as those found on airline tickets, in bank numbers on checks, in credit card numbers, in the Universal Product Code (UPC) found on groceries, etc.

Why Can the ISBN Code Detect Any Single-Digit Error?

The ISBN-10 code is designed so that the following condition must hold:

$$\sum_{i=1}^{10} i \cdot x_i \equiv 0 \pmod{11}$$

If a single-digit error occurs (e.g., one digit x_k is changed to x'_k), the weighted sum changes by:

$$k \cdot (x'_k - x_k)$$

This value is usually **not divisible by 11**, so the checksum condition fails.

Why Can the ISBN Code Detect Transposition of Two Digits?

Suppose two digits in the ISBN-10 code, at positions *i* and *j*, are accidentally swapped.

Originally, their contribution to the checksum is:

$$i \cdot x_i + j \cdot x_j$$

After swapping, the contribution becomes:

$$i \cdot x_j + j \cdot x_i$$

The difference in the checksum is:

$$(i \cdot x_j + j \cdot x_i) - (i \cdot x_i + j \cdot x_j) = (i - j)(x_j - x_i)$$

This value is generally nonzero and **not divisible by 11**.

Repetition Code and Nearest Neighbor Decoding

Suppose we want to send a `1` to mean “yes” and a `0` to mean “no.” If we transmit just a single bit, noise may corrupt it and the receiver may interpret the wrong message.
To improve reliability, we can use a **repetition code.** For example:

- Transmit `11111` to represent “yes”
- Transmit `00000` to represent “no”

If one or two bits are flipped during transmission, the receiver can still correctly decode the message by choosing the codeword that is **closest** (in Hamming distance) to the received 5-tuple.

This decoding strategy is known as **nearest neighbor decoding**.

Alphabets in Coding Theory

In coding theory, the alphabets used are typically **finite fields** with q elements, denoted by $\mathrm{GF}(q)$.

A code is called ***q*-ary** if its codewords are defined over the q-ary alphabet $\mathrm{GF}(q)$.

The most commonly used alphabets are the **binary extension fields**, $\mathrm{GF}(2^m)$, due to their suitability for digital systems.

Linear Error-Control Codes

We study **linear error-control codes**, which are special codes with rich mathematical structure.

Linear codes are widely used in practice for several reasons:

- They are easy to construct.
- Encoding is quick and efficient.
- Decoding is often simplified due to the linearity of the code.

Their algebraic properties make them especially suitable for both theoretical study and real-world applications.

Codes over Finite Fields

Let $\mathbb{F}_q^n$ denote the vector space of all n-tuples over the finite field $\mathbb{F}_q$. An (n, M) code C over $\mathbb{F}_q$ is a subset of $\mathbb{F}_q^n$ containing M codewords.

We typically write vectors $(a_1, a_2, \ldots, a_n) \in \mathbb{F}_q^n$ as $a_1 a_2 \cdots a_n$, and refer to them as **codewords.**

Codes are often named according to the field over which they are defined:

- Codes over $\mathbb{F}_2$ are called **binary codes.**
- Codes over $\mathbb{F}_3$ are called **ternary codes.**
- Codes over $\mathbb{F}_4$ are called **quaternary codes.**

Note: The term *quaternary* has also been used to describe codes over the ring $\mathbb{Z}_4$, the integers modulo 4.

Linear Codes, Generator and Parity-Check Matrices

If C is a k-dimensional subspace of $\mathbb{F}_q^n$, then C is called an $[n, k]$ linear code over $\mathbb{F}_q$.

A **generator matrix** G for an $[n, k]$ code C is any $k \times n$ matrix whose rows form a basis for C.

There also exists an $(n - k) \times n$ matrix H, called a **parity-check matrix** for the code C, defined by:

$$C = \left\{ x \in \mathbb{F}_q^n \,\middle|\, Hx^T = 0 \right\}.$$

Note: The rows of H are also linearly independent.

Example: A Hamming Code

The set

$$\{0000000, 0001111, 0010110, 0011001, 0100101, 0101010, 0110011,$$

$$0111100, 1000011, 1001100, 1010101, 1011010, 1100110, 1101001,$$

$$1110000, 1111111\}$$

is a binary linear code known as a **Hamming code.**

This is a $(7, 16, 3)$ code because:

- Each codeword has length 7.
- There are 16 codewords.
- The dimension is 4.

Generator and Parity-Check Matrices

Theorem:
If the generator matrix G of an $[n, k]$ linear code C is in standard form:

$$G = [I_k \mid A],$$

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

then a parity-check matrix H for C is given by:

$$H = [-A^T \mid I_{n-k}],$$

where A^T is the transpose of A, and I_{n-k} is the identity matrix of size $n-k$.

Generator and Parity-Check Matrices

- The **generator matrix** G of an $[n, k]$ linear code C is a $k \times n$ matrix whose rows are linearly independent and span the code.
- The **parity-check matrix** H has linearly independent rows and satisfies $Hx^T = 0$ for all $x \in C$.
- The rows of H generate a new code called the **dual code** (or **orthogonal code**) of C, denoted $C^\perp$.

Inner Product and Dual Code

Let $u = (u_0, u_1, \ldots, u_{n-1})$ and $v = (v_0, v_1, \ldots, v_{n-1}) \in \mathbb{F}_q^n$.

The **inner product** $u \cdot v$ is defined as:

$$u \cdot v = \sum_{i=0}^{n-1} u_i v_i$$

If C is a linear code of length n over $\mathbb{F}_q$, then the **(Euclidean) dual code** $C^\perp$ is defined by:

$$C^\perp = \left\{ v \in \mathbb{F}_q^n : u \cdot v = 0 \text{ for all } u \in C \right\}.$$

Linear Complementary Pair (LCP) and LCD Codes

Definition

Let (C, D) be a pair of linear codes of length n over a field $\mathbb{F}$. The pair (C, D) is called a **linear complementary pair (LCP)** if:

$$C \oplus D = \mathbb{F}^n \quad \text{and} \quad C \cap D = \{0\}.$$

If $D = C^{\perp}$, then C is called a **linear complementary dual (LCD)** code.

Characterization of LCD Codes

If C is an LCD code over $\mathbb{F}$ with generator matrix G and parity-check matrix H, then the following properties hold:

1. $GH^T = \mathbf{0}$ and $HG^T = \mathbf{0}$.
2. GG^T is invertible.
3. The matrix
$$\begin{pmatrix} G \\ H \end{pmatrix}$$
is invertible.

Hull and l-Intersection Codes

Definition

Let C be a linear code over $\mathbb{F}$. The **hull** of C is the linear code

$$\mathrm{Hull}(C) = C \cap C^{\perp}.$$

The definition of LCD codes was generalized to linear l-intersection codes as follows:

Definition

Let (C, D) be a pair of linear codes of length n over $\mathbb{F}$. Then (C, D) is called a **linear l-intersection pair** of codes if

$$\dim(C \cap D) = l.$$

If $D = C^{\perp}$, then

$$\dim(\mathrm{Hull}(C)) = l.$$

Cyclic Codes

Definition

A linear code C of length n over a field $\mathbb{F}$ is called a **cyclic code** if

$$\sigma(c) \in C \quad \text{for all } c \in C,$$

where σ is the *cyclic shift* operator. In other words, for every codeword

$$(c_0, c_1, \ldots, c_{n-1}) \in C,$$

the cyclic shift

$$(c_{n-1}, c_0, \ldots, c_{n-2})$$

also belongs to C.

Efficient Implementation of Cyclic Codes

Cyclic codes can be implemented efficiently using simple hardware devices called shift registers.
This is of great interest in applications involving fiber optics, where high-speed data rates are possible.

Example of a Cyclic Code

Example

The binary code $C = \{000,\ 110,\ 011,\ 101\}$ is a cyclic code.

- It is linear and of length $n = 3$.
- For each codeword, its cyclic shift is also in C. For example:

$$\text{Shift of } 110 \to 011,$$
$$\text{Shift of } 011 \to 101,$$
$$\text{Shift of } 101 \to 110.$$

- Hence, C is closed under cyclic shifts.

Codewords as Polynomials

- When working with cyclic codes, it is convenient to represent codewords as polynomials.
- A codeword vector $\mathbf{c} = (c_0, c_1, \ldots, c_{n-1})$ of length n is associated with the **code polynomial**:

$$c(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1}$$

Code Polynomials and Cyclic Codes (1)

Now consider the code polynomials $\{0,\ 1+x,\ x+x^2,\ 1+x^2\}$ corresponding to the code $C = \{000,\ 110,\ 011,\ 101\}$. Notice that the highest power appearing among the code polynomials is x^2, since a higher power term would indicate a codeword with length greater than 3.

Code Polynomials and Cyclic Codes (2)

Similarly to working with integers modulo a specific prime integer p, we can work with polynomials modulo a specific polynomial, for example $p(x) = x^3 - 1$.
In this situation, we think of $x^3 - 1$ being equivalent to 0, or in other words, x^3 is equivalent to 1. Let's see what this means with an example.

Example (1)

Example

Consider the code polynomial $c(x) = 1 + x^2$ corresponding to the codeword $c = 101$ from $C = \{000, 110, 011, 101\}$.
Let's multiply $c(x)$ by x to obtain $c(x) \cdot x = c_0(x) = x + x^3$.
However, if we are working modulo $p(x) = x^3 - 1$, then x^3 is equivalent to 1.
So, $c_0(x)$ is equivalent to $x + 1$ modulo $p(x)$.

Example (2)

Example

Rearranging this new polynomial with terms from lowest to highest powers gives $1 + x$, which is the code polynomial for the codeword $110 \in C$.
Notice that 110 is the right cyclic shift of 101.
Let's try this again. Start with the code polynomial $d(x) = x + x^2$ corresponding to the codeword 011 in C.
Multiplying $d(x)$ by x gives $d_0(x) = x^2 + x^3$. Taking this result modulo $x^3 - 1$ gives $x^2 + 1$, which corresponds to the codeword $101 \in C$ and is the right cyclic shift of 011. Interesting!

Example - Cyclic Code Shift (1)

Previous example suggests a true fact about cyclic codes.
In a cyclic code C of length n, the product $x \cdot c(x)$ modulo $x^n - 1$ produces another code polynomial in C, namely the right cyclic shift of $c(x)$.
More precisely, when working with code polynomials of degree less than n corresponding to codewords of length n, by working modulo $x^n - 1$, we can achieve a right cyclic shift of a codeword by multiplying the associated code polynomial by x.

Example - Cyclic Code Shift (2)

Consider the code polynomial $c(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$. Multiplying $c(x)$ by x modulo $x^n - 1$ gives:

$$c_0(x) = c_0 x + c_1 x^2 + \cdots + c_{n-1} x^n \equiv c_0 x + c_1 x^2 + \cdots + c_{n-1} \mod x^n - 1.$$

The codeword associated with $c_0(x)$ is $(c_{n-1}, c_0, \ldots, c_{n-2})$, which is clearly the right cyclic shift of the codeword associated with $c(x)$.

Theorem - Characterization of Cyclic Codes

Theorem

A linear code of length n over $GF(q)$ is cyclic if and only if C satisfies the following two conditions:

1. *If $a(x)$ and $b(x)$ are code polynomials in C, then $a(x) - b(x) \in C$.*
2. *If $a(x)$ is a code polynomial in C and $r(x)$ is any polynomial of degree less than n, then $r(x)a(x) \in C$.*

References I

W. C. Huffman and V. Pless,
Fundamentals of Error-Correcting Codes,
Cambridge University Press, 2010.

H. Bordbar, *On the construction of nonbinary LCD quadratic residue and double quadratic residue codes*, Finite Fields and Their Applications, 103 (2025) 102591. DOI: `https://doi.org/10.1016/j.ffa.2025.102591`.

H. Bordbar, *Ordered Algebraic Structure in a Linear Code*, Advances in Mathematics of Communications, 0-0 (2024). DOI: `10.3934/amc.2024052`

References II

- H. Bordbar, *BCK Codes*, Advances in Computational Intelligence, 2 (1), 4 (2022). DOI: `https://doi.org/10.1007/s43674-021-00018-4`.
- H. Bordbar, *The Structure of the Block Code Generated by a BL-Algebra*, Mathematics 10 (5), 692 (2022). DOI: `https://doi.org/10.3390/math10050692`.

Thank you for your attention!