Teoría de códigos y polinomios torcidos

Adrián Ranea

22 de mayo de 2016

Resumen

En este trabajo hacemos una introducción de la teoría de códigos algebraica para explicar los códigos θ-cíclicos que se basan en anillos de polinomios torcidos. Estos códigos son un buen ejemplo de aplicación práctica del álgebra no conmutativa.

Índice

1. Introducción

Los códigos correctores de errores se utilizan para detectar y corregir errores que ocurren cuando la información se transmite por una canal con ruido.

Por ejemplo, los CDs usan códigos correctores de errores para que un reproductor de CD pueda leer la información del CD incluso si esta se ha corrompido por ruido en forma de imperfecciones en el CD. Las fotografías envidadas desde el espacio exterior a la Tierra usan códigos correctores de errores para protegerse del ruido causado por la luz y otras interrupciones atmosféricas.

Para conseguir su objetivo, los códigos correctores de errores añaden redundancia al mensaje. En general, si un mensaje tiene longitud k, el mensaje codificado tendrá longitud $n > k$.

Veamos un ejemplo muy simple antes de empezar con el tratamiento matemático de la teoría de códigos.

Ejemplo 1.1 (El código de repetición)**.** Supongamos que queremos enviar un 1 que significa «si» y un 0 que representa «no». Si enviamos solo un bit, entonces existe la posibilidad de que el bit se corrompa por el ruido del canal y se reciba un mensaje no intencionado. Una solución simple es usar un *código de repetición*. En lugar de enviar un solo bit, enviamos 111 para comunicar «si» y 000 para decir «no». Si se produce un único error, el receptor puede detectar y corregir el error tomando la «palabra» más cercana. Por ejemplo, si el receptor recibe 001 y suponiendo que solo se ha producido un error, el receptor puede deducir que el mensaje original era 000 que significa «sí».

2. Preliminares

En este apartado introduciremos los conceptos básicos de la teoría de códigos y algunos resultados básicos. El objetivo de este sección y de la sección 3 es sentar las bases para poder estudiar los códigos cíclicos torcidos.

2.1. Terminología básica

Llamamos *alfabeto* al conjunto de símbolos que utiliza un código. Tradicionalmente, los alfabetos usados en teoría de códigos son los cuerpos finitos $\mathbb{F}_q$. Los más utilizados son las extensiones de cuerpos binarias $\mathbb{F}_{2^m}$.

Definición 2.1. Un *código de bloque* consiste en una función codificadora $E : \mathbb{F}_q^m \to \mathbb{F}_q^n$ y una función decodificadora $D : \mathbb{F}_q^n \to \mathbb{F}_q^m$. Los elementos de $Im(E) \subset \mathbb{F}_q^n$ se llaman palabra código y los elementos de $\mathbb{F}_q^n$ palabras.

A no ser que se indique lo contrario, consideraremos siempre los códigos como códigos de bloque y los alfabetos serán cuerpos finitos de q elementos.

2.2. Paramétros importantes

Al desarollar códigos, es necesario evaluarlos y decidir cuales son «buenos» códigos. Hay tres parámetros principales para describir códigos:

1. La *longitud de código, n.* En el ejemplo 1.1, la longitud era 3 ya que las palabra código $\{000, 111\}$ contenían 3 bits cada una.

2. El *número total de palabra código, M.* En el ejemplo 1.1, el numero total de palabra código es 2.

3. El tercer parámetro mide la *distancia* entre pares de palabras en un código. Para explicar bien este parámetro necesitamos las siguientes definiciones.

Definición 2.2. El *peso de Hamming* $w(\mathbf{c})$ de una palabra código $\mathbf{c}$ es el número de componentes no nulas de la palabra.

Ejemplo 2.1. $w(000) = 0,\ w(111) = 3$

Definición 2.3. La *distancia de Hamming* entre dos palabra código $d(\mathbf{x}, \mathbf{y})$ es el peso de Hamming del vector diferencia $\mathbf{x} - \mathbf{y}$, esto es, $w(\mathbf{x} - \mathbf{y})$.

En otras palabras, la distancia de Hamming entre dos palabra código $d(\mathbf{x}, \mathbf{y})$ es el número de posiciones en los que $\mathbf{x}$ e $\mathbf{y}$ difieren. De ahora en adelante, distancia siempre significará distancia de Hamming.

Definición 2.4. La *distancia mínima (de Hamming)* de un código C es la distancia mínima entre dos palabra código cualesquiera. Simbólicamente:

$$d(C) = \text{mín}\{d(\mathbf{x}, \mathbf{y}) : \mathbf{x} \neq \mathbf{y},\ \mathbf{x}, \mathbf{y} \in C\}$$

En el ejemplo 1.1, la distancia mínima era 3 ya que las dos palabra código diferían en las 3 posiciones.

De hecho, la distancia de Hamming es una *métrica* en el espacio de todas las n-tuplas de $\mathbb{F}_q$, esto es, d verifica las siguientes propiedades para cualesquiera $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}_q^n$:

- $d(\mathbf{x}, \mathbf{x}) = 0$
- $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$
- $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$

Llamaremos (n, M, d)-código a un código con distancia mínima d que consiste en M palabras código y todas ellas de longitud n. Uno de los objetivos principales de la teoría de códigos es desarollar códigos que consigan un equilibrio entre n pequeño (para transmisiones rápidas de mensajes), M grande (para permitir transmisiones de una amplia variedad de mensajes) y d grande (para detectar mucho errores).

2.3. Corrigiendo y detectando errores

Cuando hablamos del número de errores en una palabra recibida, estamos hablando de la distancia entre la palabra recibida y la palabra código originalmente transmitida. El trabajo del decodificador es decidir que palabra código fue transmitida. Muchos códigos usan el esquema de decodificación *vecino más cercano* que elige la palabra código que minimiza la distancia entre el vector recibido y los posibles vectores transmitidos. El esquema *vecino más cercano* para un código sobre $\mathbb{F}_q$ suele hacer las siguientes suposiciones sobre el canal de comunicación:

1. Cada símbolo transmitido tiene la misma probabilidad p de recibirse erróneamente.
2. Si un símbolo se recibe erróneamente, cada uno de los $q-1$ posibles errores es equiprobable.

Tales canales se llaman *canales simétricos* y de ahora en adelante todos los canales los supondremos simétricos.

Volviendo al ejemplo 1.1, es fácil ver que dicho código puede detectar hasta dos errores y corregir hasta un error. En general, se tiene el siguiente resultado:

Proposicion 2.1. 1. Un código C puede detectar hasta s errores en cualquier palabra si $d(C) \leq s+1$.

2. Un código C puede corregir hasta t errores en cualquier palabra si $d(C) \geq 2t+1$.

Demostración. 1. Supongamos que $d(C) \geq s+1$. Supongamos que una palabra código $\mathbf{c}$ se transmite y que s o menos errores ocurren durante la transmisión. Entonces, la palabra recibida no puede ser una palabra código, ya que todas las palabra código difieren de $\mathbf{c}$ en al menos $s+1$ posiciones. Luego los errores son detectados.

2. Supongamos que $d(C) \geq 2t+1$. Supongamos que una palabra código $\mathbf{x}$ se transmite y que la palabra recibida, $\mathbf{y}$, contiene t o menos errores. Entonces $d(\mathbf{x},\mathbf{y}) \leq t$. Sea $\mathbf{z}$ una palabra código distinta de $\mathbf{x}$. Entonces $d(\mathbf{z},\mathbf{y}) \geq t+1$, ya que en caso contrario $d(\mathbf{z},\mathbf{y}) \leq t$ implicaría que $d(\mathbf{x},\mathbf{z}) \leq d(\mathbf{x},\mathbf{y}) + d(\mathbf{z},\mathbf{y}) \leq 2t$ (por la desigualdad triangular de la distancia), lo cual es imposible ya que $d(C) \geq 2t+1$. Luego $\mathbf{x}$ es la palabra código más cercana a $\mathbf{y}$ y $\mathbf{y}$ es decodificada correctamente.

□

Podemos interpretar geométricamente la proposición 2.1. Como cada palabra código C esta a distancia $2t+1$ de cualquiera otra palabra código, podemos imaginar cada palabra código $\mathbf{x}$ rodeada por una esfera de radio t tales que dichas esferas son disjuntas. Cualquier vector recibido que esté en la esfera centrada en $\mathbf{x}$ será decodificada como $\mathbf{x}$.

Esto explica porque los vectores recibidos que contienen t o menos errores se pueden decodificar correctamente: aun están en la esfera correcta. De hecho usando las esferas de $\mathbb{F}_q^n$ se puede encontrar una cota para el número de palabras código M:

Proposicion 2.2. Un código corrector de t-errores sobre $\mathbb{F}_q$ de longitud n debe verificar

$$M \sum_{n=0}^{t} \binom{n}{i} (q-1)^i \leq q^n$$

Una demostración posible se encuentra en [3].

3. Teoría de códigos algebraica

La teoría de códigos algebraica es el área de la matemática discreta que se encarga de desarrollar códigos correctores de errores y procedimientos para codificar y descodificar en términos algebraicos. En este sección se introducirán los códigos lineales y los códigos cíclicos que no son más que códigos correctores de errores con estructura adicional.

3.1. Códigos lineales

Definición 3.1. Un *código lineal* de longitud n sobre $\mathbb{F}_q$ es un subespacio vectorial del espacio vectorial $\mathbb{F}_q^n$.

Los códigos lineales son muy utilizados por varios razones: son fáciles de encontrar, la codificación es rápida y fácil y la decodificación se ayuda mucho de la linealidad del código (teorema 3.1).

Si C es un código lineal y tiene dimensión k como subespacio vectorial de $\mathbb{F}_q^n$, diremos que C es un $[n,k]$-código lineal. Si C es un $[n,k]$ código lineal, entonces C tiene q^k palabra código. Esto significa que C puede usarse para comunicar q^k distintos mensajes. Podemos identificar el espacio de mensajes con el espacio vectorial $\mathbb{F}_q^k$. Así, las funciones de codificación/decodificación serían:

$$E : \mathbb{F}_q^k \to \mathbb{F}_q^n$$
$$D : \mathbb{F}_q^n \to \mathbb{F}_q^k$$

esto es, codificamos mensajes de longitud k para obtener palabras código de longitud n, con $n \geq k$ y viceversa.

Proposicion 3.1. Sea C un código lineal. Toda combinación lineal de palabras código de C es una palabra código de C.

Demostración. Trivial, ya que C es un subespacio vectorial de $\mathbb{F}_q^n$. □

Proposicion 3.2. La distancia mínima $d(C)$ de un código lineal C es igual a $w^*(C)$, el mínimo peso entre las palabras código no nulas.

Demostración. Sean $\mathbf{x}, \mathbf{y} \in C,\ \mathbf{x} \neq \mathbf{y}$, donde la distancia mínima se consigue, esto es, $d(C) = d(\mathbf{x}, \mathbf{y})$. Por la definición de la distancia de Hamming, $d(\mathbf{x}, \mathbf{y}) = w(\mathbf{x}, \mathbf{y})$. Como $\mathbf{x} - \mathbf{y}$ es una palabra código por la proposición 3.1 y por la definición de $w^*(C)$, tenemos $d(C) = w(x - y) \geq w^*(C)$.

La desigualdad opuesta se obtiene como sigue. Sea $\mathbf{c} \in C$ donde el peso mínimo se consigue, esto es, $w^*(C) = w(\mathbf{c})$. Como $w(\mathbf{c}) = d(\mathbf{c}, 0)$ ya que $\mathbf{0} \in C$, se obtiene $w^*(C) = d(\mathbf{c}, 0) \geq d(C)$. □

La proposición 3.2 facilita encontrar la distancia mínima de un código lineal ya que en lugar de buscar la distancia entre todos los posibles pares de palabras código basta buscar entre los pesos de cada palabra código.

Es posible dar una cota inferior de la distancia mínima de un código lineal. Puede encontrar la demostración del siguiente resultado en [4].

Lema 3.1 (límite de Singleton)**.** Si C es un $[n, k]$-código lineal entonces $d \leq n - k + 1$.

Sea C un $[n, k]$-código lineal que consigue la igualdad del lema 3.1. Entonces C es un $(n, q^k, n - k + 1)$-código y en consecuencia puede detectar hasta $n - k$ errores y corregir hasta $(n - k - 1)/2$ errores. A esta clase de códigos se les llama *códigos separables de distancia máxima* y un ejemplo no trivial es el código *Reed-Solomon* que se ha empleado en tecnologías tan conocidas como los CDs, DVDs, discos Blu-ray o los códigos QR, entre otras muchas aplicaciones.

3.1.1. Matriz generatriz

Como los códigos lineales son espacios vectoriales, un $[n, k]$-código lineal puede ser especificado por una base de k palabras código.

Definición 3.2. Una matriz G de tamaño $k \times n$ cuyas filas forman una base para un $[n, k]$-código lineal se llama *matriz generatriz* del código.

De hecho, una matriz G es una matriz generatriz para algún código lineal si y solo si las filas de G son linealmente independientes.

La matriz generatriz es una forma compacta de describir todas las palabras código de C. Es más, esta matriz puede usarse para codificar mensajes. Si C es un $[n, k]$-código lineal con matriz generatriz G, entonces la función de codificación puede escribirse como:

$$\begin{aligned} E :& \mathbb{F}_q^k \to \mathbb{F}_q^n \\ & \mathbf{x} \mapsto \mathbf{x}G \end{aligned}$$

3.1.2. Clases laterales y decodificación

En esta subsección describiremos el esquema *vecino más cercano* que usa la estructura de subgrupo aditivo del espacio vectorial de un código lineal. Veremos como las clases laterales pueden ser usadas para decodificar un código lineal.

Definición 3.3. Sea C un $[n, k]$-código lineal sobre $\mathbb{F}_q$ y $\mathbf{a}$ un vector de $\mathbb{F}_q^n$. El conjunto $\mathbf{a} + C = \{\mathbf{a} + \mathbf{x} \mid \mathbf{x} \in C\}$ se llama una *clase lateral* de C.

Este definición es un caso particular del concepto de clase lateral a izquierda (o derecha) de un subgrupo. En los espacios vectoriales, las clases laterales son los subespaciones afínes.

Proposicion 3.3. Si C es un $[n, k]$ código lineal sobre $\mathbb{F}_q$, entonces cada vector de $\mathbb{F}_q^n$ está en alguna clase lateral de C, cada clase lateral contiene q^k vectores y dos clases laterales son idénticas o disjuntas.

Demostración. Se puede probar bien usando la teoría de grupos o la teoría de espacios vectoriales. □

En nuestro caso, tomaremos como representantes de las clases laterales aquellos vectores de peso mínimo que estén en la clase lateral. Los representantes definidos de este modo son los vectores de error, esto es, si $\mathbf{x}$ fue enviado pero se recibió y, el *vector de error* es $\mathbf{e} = \mathbf{y} - \mathbf{x}$.

El algoritmo de decodificación con clases laterales para un $[n, k]$-código lineal funciona como sigue. Particionamos $\mathbb{F}_q^n$ en clases laterales de C. Hay $q^n/q^k = q^{n-k}$ clases laterales y cada una contiene q^k elementos. Para cada clase lateral, elegimos un representante (si hay varios de peso mínimo, se puede elegir uno arbitariamente). Entonces, si se recibe $\mathbf{y}$, se busca la clase lateral que contiene a $\mathbf{y}$. Esta clase lateral será de la forma $\mathbf{e} + C$ y podemos deducir que $\mathbf{y} - \mathbf{e}$ fue la palabra original enviada. Se está asumiendo que los vectores de errores de menor peso son los más probables.

3.2. Códigos cíclicos

Definición 3.4. Un código lineal C de longitud n es *cíclico* si siempre que $(c_0, c_1, \ldots, c_{n-1})$ sea una palabra código de C, entonces $(c_{n-1}, c_0, \ldots, c_{n-1})$ es también una palabra código de C.

Los códigos cíclicos se pueden implementar de forma eficiente usando dispositivos hardware llamados registros de desplazamiento. Esto es de gran interés en aplicaciones que involucran la fibra óptica, donde la velocidad de transmisión es muy elevada.

Con los códigos cíclicos, se suele utilizar una representación polinomial en vez de

vectorial. Consideremos la aplicación:

$$\phi : \mathbb{F}_q^n \to \mathbb{F}_q[x]/(x^n - 1)$$
$$(c_0, c_1, \ldots, c_{n-1}) \mapsto c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$$

Es fácil ver que ϕ es un isomorfismo de espacios vectoriales. Dado un código lineal C, $\phi(C)$ es un subespacio vectorial de $\mathbb{F}_q[x]/(x^n - 1)$ ya que C es un subespacio vectorial. Además, las siguientes afirmaciones son equivalentes:

1. $(c_0, c_1, \ldots, c_{n-1}) \in C \implies (c_{n-1}, c_0, \ldots, c_{n-2}) \in C$
2. $x(c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}) \in \phi(C)$

Demostración. La equivalencia se basta en la siguiente igualdad en $\mathbb{F}_q[x]/(x^n - 1)$:

$$\begin{aligned} x(c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}) &= xc_0 + xc_1 x + \cdots + xc_{n-1}x^{n-1} \\ &= c_0 x + c_1 x^2 + \cdots + c_{n-1}x^n \\ &= c_{n-1} + c_0 x + c_1 x^2 + \cdots + c_{n-2}x^{n-1} \end{aligned}$$

□

Es decir, un desplazamiento circular hacia la derecha en $\mathbb{F}_q^n$ es equivalente a multiplicar por x en $\mathbb{F}_q[x]/(x^n - 1)$. Sea C un código cíclico. Denotando $f(x) = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$, si $f(x) \in \phi(C)$, entonces $xf(x), x^2 f(x), x^3 f(x), \ldots, x^{n-1} f(x)$ también están en $\phi(C)$. Usando la linealidad de C se llega a que $\forall p(x) \in \mathbb{F}_q[x]/(x^n - 1),\ p(x)f(x) \in \phi(C)$. Hemos probado que $\phi(C)$ es un ideal de $\mathbb{F}_q[x]/(x^n - 1)$ si C es un codigo cíclico.

A los elementos de $\phi(C)$ los llamaremos *polinomios código*. De ahora en adelante, identificaremos C con $\phi(C)$ y usaremos los términos palabra código y polinomios código de forma intercambiable. De hecho los códigos cíclicos son justamente los ideales de $\mathbb{F}_q[x]/(x^n - 1)$ como indica el siguiente resultado:

Teorema 3.1. Los códigos cíclicos de longitud n sobre $\mathbb{F}_q$ son justamente los ideales del anillo $\mathbb{F}_q[x]/(x^n - 1)$

Demostración. Sabemos que todo código cíclico es un ideal. Basta ver el recíproco. Sea I un ideal de $\mathbb{F}_q[x]/(x^n-1)$. En particular I es un subgrupo aditivo, luego $\mathbb{F}_q[x]/(x^n-1)$ es un código lineal. Además por ser ideal, si $p(x) \in \mathbb{F}_q[x]/(x^n - 1)$, $xp(x) \in Fx$, luego I es cíclico. □

Sabemos que $\mathbb{F}_q[x]/(x^n - 1)$ es un dominio euclideo (en particular un dominio de ideales principales) e incluso conocemos la estructura de sus ideales. Con esto podemos afirmar:

Proposicion 3.4. Sea C un código cíclico de longitud n. Entonces C está generado por un divisor $g(x)$ de $(x^n - 1)$ en $\mathbb{F}_q[x]$. Tomando $g(x)$ mónico y de grado mínimo, $g(x)$ es único y lo llamaremos el *polinomio generador* de C.

El siguiente resultado facilita el listado de todos los códigos cíclicos de una longitud dada.

Proposicion 3.5. Hay una correspondencia uno a uno entre los divisores mónicos de $x^n - 1$ en $\mathbb{F}_q[x]$ y los códigos cíclicos sobre $\mathbb{F}_q$ de longitud n.

Demostración. Ya sabemos que dado un código cíclico, este tiene asociado un único divisor $g(x)$ de $(x^n - 1)$ mónico. Por otro lado, supongamos que $h(x)$ es un divisor mónico de $x^n - 1$ en $\mathbb{F}_q[x]$. Consideramos el código $C = (h(x))$. Evidentemente, $h(x)$ genera C. Supongamos que $h(x)$ no es el polinomio generador $g(x)$. En tal caso, $g(x)$ divide a $h(x)$. Pero como $g(x) \in C = (h(x))$, $h(x)$ divide a $g(x)$ módulo $x^n - 1$. Esto implica que $g(x) = h(x)$. □

Proposicion 3.6. Sea C un código cíclico con polinomio generador $g(x) = g_0 + g_1x + \cdots + g_rx^r$ de grado r. Entonces la dimensión de C es $n - r$ y una matriz generatriz para C es la siguiente matriz $(n - r) \times n$:

$$G = \begin{pmatrix} g_0 & g_1 & \cdots & g_r & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & g_r & 0 & \cdots & 0 \\ 0 & 0 & g_0 & g_1 & \cdots & g_r & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \cdots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & g_0 & g_1 & \cdots & g_r \end{pmatrix}$$

Demostración. En primer lugar, g_0 no puede ser 0. En otro caso, $(0, g_1, \ldots, g_r, 0 \ldots, 0) \in C$ implicaría que $(g_1, \ldots, g_r, 0 \ldots, 0, 0) \in C$, esto es, $g_1 + g_2x + \cdots + g_rx^r \in C$ lo cual contradice la minimalidad del grado r del polinomio generador. Las $n - r$ filas de la matriz G son linealmente independientes por la forma escalonada de la matriz con 0s debajo de los g_0s no nulos. Estas $n - r$ filas representan los polinomios código $g(x), xg(x), \ldots, x^{n-r-1}g(x)$. Basta ver que estos elementos forman un sistema de generadores. Dado $c(x) \in C$, existe $m(x) \in \mathbb{F}_q[x]/(x^n - 1)$ de grado $n - r$ tal que $c(x) = m(x)g(x)$. La siguiente cuenta

$$\begin{aligned} c(x) &= m(x)g(x) \\ &= (m_0 + m_1x + \cdots + m_{n-r-1}x^{n-r-1})g(x) \\ &= m_0g(x) + m_1xg(x) + \cdots m_{n-r-1}x^{n-r-1}g(x) \end{aligned}$$

muestra que todo polinomio código $c(x)$ de C puede escribirse como una combinación lineal de los polinomios código representados en los $n - r$ filas de G. Luego G es una matriz generatriz de C y la dimensión de C es $n - r$. □

Como en los códigos lineales, podemos usar la matriz generatriz para realizar la codificación. Sin embargo, para los códigos cíclicos podemos usar también el polinomio generador. Dado un mensaje $\mathbf{a} = (a_0, \ldots, a_{n-1})$, y su representación polinomial $a(x)$, mediante multiplicación de polinomios podemos realizar la codificación:

$$\begin{aligned} E : \mathbb{F}_q[x]/(x^n - 1) &\to \mathbb{F}_q[x]/(x^n - 1) \\ a(x) &\mapsto a(x)g(x) \end{aligned}$$

Ejemplo 3.1. Vamos a calcular todos los códigos cíclicos sobre $\mathbb{F}_2$ de longitud 4. La factorización de $x^4 + 1$ es $(x + 1)^4$, luego sus factores mónicos son:

$$\begin{aligned} g_0(x) &= 1 \\ g_1(x) &= x + 1 \\ g_2(x) &= (x + 1)^2 = x^2 + 1 \\ g_3(x) &= (x + 1)^3 = x^3 + x^2 + x + 1 \\ g_4(x) &= (x + 1)^4 = x^4 + 1 \end{aligned}$$

El primer y el último factor generan los códigos triviales $\mathbb{F}_2[x]/(x^4 + 1)$ y $\{0\}$ respectivamente. El resto de polinomios generan los códigos:

$$\begin{aligned} g_1(x) \to C_1 &= \left\{0, x + 1, x^2 + 1, x^2 + x, x^3 + 1, x^3 + x, x^3 + x^2, x^3 + x^2 + x + 1\right\} \\ g_2(x) \to C_2 &= \left\{0, x^2 + 1, x^3 + x, x^3 + x^2 + x + 1\right\} \\ g_3(x) \to C_3 &= \left\{0, x^3 + x^2 + x + 1\right\} \end{aligned}$$

Además, el número de palabras de estos tres códigos no triviales son: $M_{C_1} = 8$, $M_{C_2} = 4$, $M_{C_3} = 2$. Teniendo en cuenta los grados de los polinomios generadores, C_1 es un código $[4, 3]$-cíclico, C_2 es un código $[4, 2]$-cíclico y C_3 es un código $[4, 1]$-cíclico.

Para calcular el código cíclico que genera cada polinomio hemos usado el siguiente script en *python* junto con la librería *sympy*:

```
def generaCiclico(g):
    F2_x = [
        0, 1, x, x+1, x**2, x**2 + x, x**2 + x + 1, x**2 + 1, x**3,
        x**3 + x**2, x**3 + x**2 + x, x**3 + x**2 + x + 1,
        x**3 + x**2 + 1, x**3 + 1, x**3 + x, x**3 + x + 1
    ]
    f = x** 4 + 1
    C = set()
    for p in F2_x:
        palabra_cod = expand(rem(expand(g*p, modulus=2), f), modulus=2)
        C.add(palabra_cod)
    return C
```

4. Códigos cíclicos torcidos

Vamos a generalizar la noción de códiglos cíclicos al concepto de códigos θ-cíclicos.

Definición 4.1. Sea θ un automorfismo de $\mathbb{F}_q$. Un código θ-cíclico es un código lineal C_θ verificando:

$$(a_0, a_1, \ldots, a_{n-1}) \in C_\theta \implies (\theta(a_{n-1}), \theta(a_0), \ldots, \theta(a_{n-2}))$$

Para generalizar los códigos cíclicos (se corresponde al caso donde θ es la identidad) vamos a considerar el anillo de polinomios torcidos sobre $\mathbb{F}_q$, esto es, la extensión de Ore:

$$\mathbb{F}_q[x, \theta] = \{a_0 + a_1 x + \cdots + a_{n-1}x^{n-1},\ a_i \in \mathbb{F}_q,\ n \in \mathbb{N}\}$$

donde la regla de multiplicación es $xa = \theta(a)x$.

Nuestro objetivo es conseguir una representación polinomial al igual que hicimos con los códigos cíclicos. Veremos que los códigos θ-cíclicos son ideales principales del anillo $\mathbb{F}_q[x, \theta]/(x^n - 1)$ lo cual nos garantizará un método para codificar tan sencillo como para los códigos cíclicos. Tambien veremos que la clase de los códigos θ-cíclicos son una clase de códigos lineales que contiene a los códigos cíclicos y es mucho más grande.

4.0.1. Propiedades generales

Las propiedades de los códigos cíclicos están íntimamente relacionadas con las propiedades de $\mathbb{F}_q[x, \theta]$. El anillo $\mathbb{F}_q[x, \theta]$ es un anillo euclídeo por la izquierda y por la derecha (ya que existe un algoritmo de división euclídea por la izquierda y por la derecha) y su ideales por la izquierda y por la derecha son principales.

Denotamos por $K \subset \mathbb{F}_q$ al subcuerpo cuyos elementos se quedan fijos por θ.

Proposicion 4.1. Los elementos centrales de $\mathbb{F}_q[x, \theta]$ son de la forma $\sum_{i=0}^{m} c_i x^{i|\theta|}$ donde $c_i \in K$ y $|\theta|$ es el orden del automorfismo θ. Simbólicamente:

$$Z(\mathbb{F}_q[x, \theta]) = K[x^{|\theta|}]$$

Demostración. Supongamos $g(x) \in Z(\mathbb{F}_q[x, \theta])$. Si $g(x)$ es de grado m, $g(x)$ será de la forma:

$$g(x) = \sum_{i=0}^{m} g_i x^i$$

donde $g_m \neq 0$ pero el resto puede ser cero. Si multiplicamos $g(x)$ por un polinomio

constante arbitrario $a(x) = a_0$, usando la regla de multiplicación:

$$g(x)a(x) = \sum_{i=0}^{m} g_i \theta^i(a_0) x^i$$

$$a(x)g(x) = \sum_{i=0}^{m} a_0 g_i x^i$$

Como $g(x)$ es un elemento central, $a_0 = \theta^i(a_0)\ \forall a_0 \in \mathbb{F}_q$, luego i es 0 o es un múltiplo del orden del automorfismo θ, esto es, $g(x) \in \mathbb{F}_q[x^{|\theta|}]$. Tenemos que ver que los coeficientes de $g(x)$ están en K. Para ello multiplicamos $g(x)$ por el polinomio $b(x) = b_1 x, b_1 \in \mathbb{F}_q$:

$$g(x)b(x) = \sum_{i=0}^{m} g_i \theta^i(b_1) x^{i+1}$$

$$b(x)g(x) = \sum_{i=0}^{m} b_1 \theta(g_i) x^{i+1}$$

Como $g(x)$ es un elemento central y $b_1 = \theta^i(b_1)$ por lo anterior, deducimos que $g_i = \theta(g_i)$, esto es, $g_i \in K$. Hemos visto que si $g(x) \in Z(\mathbb{F}_q[x, \theta])$ entonces $g(x) \in K[x^{|\theta|}]$. Veamos ahora la otra inclusión. Sea $c\ x^{j|\theta|} \in K[x^{|\theta|}]$ para algún $j \in \mathbb{N}$ y $c \in K$. Vamos a multiplicar este monomio por un polinomio arbitrario $f(x) = \sum_{i=0}^{l} f_i x^i$ para algún $l \in \mathbb{N}$:

$$c\ x^{j|\theta|} \times f(x) = \sum_{i=0}^{l} c\ \theta^{j|\theta|}(f_i) x^{i+j|\theta|}$$

$$f(x) \times c\ x^{j|\theta|} = \sum_{i=0}^{l} f_i \theta^i(c) x^{i+j|\theta|}$$

Como ambas expresiones son las mismas, $c\ x^{j|\theta|} \in Z(\mathbb{F}_q[x, \theta])$. Por linealidad se extiente a cualquier polinomio de $K[x^{|\theta|}]$. □

Sea n un entero divisible por el orden $|\theta|$ de θ. Vimos en clase los siguientes resultados:

Proposicion 4.2. Los elementos centrales de $\mathbb{F}_q[x, \theta]$ son los generadores de los ideales biláteros de $\mathbb{F}_q[x, \theta]$ y $(x^n - 1) \subset \mathbb{F}_q[x, \theta]$ es un ideal bilátero.

Proposicion 4.3. El anillo $\mathbb{F}_q[x, \theta]/(x^n - 1)$ es un anillo de ideales principales por la izquierda cuyos ideales por la izquierda están generados por un divisor de $x^n - 1$ en $\mathbb{F}_q[x, \theta]$.

Veamos como dar una representación polinomial para un código θ-cíclico . Consideramos la aplicación:

$$\phi : \mathbb{F}_q^n \to \mathbb{F}_q[x, \theta]/(x^n - 1)$$
$$(c_0, c_1, \ldots, c_{n-1}) \mapsto c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$$

Es fácil ver que ϕ es un isomorfismo de $\mathbb{F}_q$-módulos. Dado un código lineal C, $\phi(C)$ es un submódulo de $\mathbb{F}_q[x,\theta]/(x^n-1)$ ya que C es un submódulo de $\mathbb{F}_q^n$ (por ser un subespacio vectorial). Las siguientes afirmaciones son equivalentes:

1. $(c_0, c_1, \ldots, c_{n-1}) \in C \implies (\theta(c_{n-1}), \theta(c_0), \ldots, \theta(c_{n-2})) \in C$
2. $x(c_0 + c_1x + \cdots + c_{n-1}x^{n-1}) \in \phi(C)$

Demostración. La equivalencia se basta en la siguiente igualdad en $\mathbb{F}_q[x,\theta]/(x^n-1)$:

$$\begin{aligned} x(c_0 + c_1x + \cdots + c_{n-1}x^{n-1}) &= xc_0 + xc_1x + \cdots + xc_{n-1}x^{n-1} \\ &= \theta(c_0)x + \theta(c_1)x^2 + \cdots + \theta(c_{n-1})x^n \\ &= \theta(c_{n-1}) + \theta(c_0)x + \theta(c_1)x^2 + \cdots + \theta(c_{n-2})x^{n-1} \end{aligned}$$

□

Es decir, un desplazamiento circular «torcido» hacia la derecha en $\mathbb{F}_q^n$ es equivalente a multiplicar por x en $\mathbb{F}_q[x,\theta]/(x^n-1)$. Sea C un código θ-cíclico . Denotando $f(x) = c_0 + c_1x + \cdots + c_{n-1}x^{n-1}$, si $f(x) \in \phi(C)$, entonces $xf(x), x^2f(x), x^3f(x), \ldots, x^{n-1}f(x)$ también están en $\phi(C)$. Usando la linealidad de C se llega a que $\forall p(x) \in \mathbb{F}_q[x,\theta]/(x^n-1)$, $p(x)f(x) \in \phi(C)$. Hemos probado que $\phi(C)$ es un ideal por la izquierda de $\mathbb{F}_q[x,\theta]/(x^n-1)$ si C es un código θ-cíclico .

A los elementos de $\phi(C)$ los llamaremos *polinomios código* y a $\phi(C)$ representación polinomial torcida de C. De ahora en adelante, identificaremos C con $\phi(C)$ y usaremos los términos palabra código y polinomios código de forma intercambiable. De hecho los códigos cíclicos son justamente los ideales por la izquierda de $\mathbb{F}_q[x]/(x^n-1)$ como indica el siguiente resultado:

Teorema 4.1. Los códigos θ-cíclicos de longitud n sobre $\mathbb{F}_q$ son justamente los ideales por la izquierda del anillo $\mathbb{F}_q[x,\theta]/(x^n-1)$.

Demostración. Sabemos que todo código θ-cíclico es un ideal por la izquierda. Basta ver el recíproco. Sea I un ideal por la izquierda de $\mathbb{F}_q[x,\theta]/(x^n-1)$. En particular $\mathbb{F}_q[x,\theta]/(x^n-1)$ es un subgrupo aditivo, luego I es un código lineal. Además por ser ideal por la izquierda, si $p(x) \in \mathbb{F}_q[x,\theta]/(x^n-1)$, $xp(x) \in \mathbb{F}_q[x,\theta]/(x^n-1)$, luego I es cíclico. □

Un factor por la derecha de grado $n-k$ de x^n-1 genera un $[n,k]$-código lineal. Si θ no es la identidad (corresponde a los códigos cíclicos), entonces $\mathbb{F}_q[x,\theta]$ no es en general un dominio de factorización única. En este caso hay muchos mas factores por la derecha que en el caso conmutativo, produciendo muchos códigos θ-cíclicos . Nótese que aunque la factorización no sea única, los grados de los polinomios torcidos irreducibles en la factorización de un elemento en $\mathbb{F}_q[x,\theta]$ si son únicos salvo permutación. Veamos un ejemplo que refleja esto.

Ejemplo 4.1. Vamos a buscar todos los [4, 2]-códigos θ-cíclicos con $\theta(a) = a^2$ sobre $\mathbb{F}_4$.

Sea α un generador del grupo multiplicativo de $\mathbb{F}_4$, esto es, un cero de $z^2+z+1 \in \mathbb{F}_2$ en $\overline{\mathbb{F}_2}$. Así, $\mathbb{F}_4 = \{0, 1, \alpha, \alpha^2\}$.

Calculamos los factores mónicos por la derecha de grado dos de $x^4 + 1 \in \mathbb{F}_4[x, \theta]$.

$$\begin{aligned} g_1(x) &= x^2 + 1 \\ g_2(x) &= x^2 + \alpha x + \alpha^2 \\ g_3(x) &= x^2 + \alpha^2 x + \alpha \\ g_4(x) &= x^2 + \alpha^2 x + \alpha^2 \\ g_5(x) &= x^2 + x + \alpha \\ g_6(x) &= x^2 + x + \alpha^2 \\ g_7(x) &= x^2 + \alpha x + \alpha \end{aligned}$$

Las factorizaciones correspondientes son:

$$\begin{aligned} x^4 + 1 &= (x^2 + 1)(x^2 + 1) \\ &= (x^2 + \alpha x + \alpha)(x^2 + \alpha x + \alpha^2) \\ &= (x^2 + \alpha^2 x + \alpha^2)(x^2 + \alpha^2 x + \alpha) \\ &= (x^2 + \alpha^2 x + \alpha)(x^2 + \alpha^2 x + \alpha^2) \\ &= (x^2 + x + \alpha^2)(x^2 + x + \alpha) \\ &= (x^2 + x + \alpha)(x^2 + x + \alpha^2) \\ &= (x^2 + \alpha x + \alpha^2)(x^2 + \alpha x + \alpha) \end{aligned}$$

Como $x^4+1 = (x+1)(x+1)(x+1)(x+1)$, los factores irreducibles de $x^4+1 \in \mathbb{F}_4[x, \theta]$ en cualquier descomposicón son todos de grado uno. Por lo tanto, ninguno de los $g_i(x)$ anteriores es irreducible.

El polinomio g_1 genera el único [4, 2]-código θ-cíclico sobre $\mathbb{F}_4$. Los otros polinomios generan un código lineal no cíclico $[4, 2]$ (los seis códigos son equivalentes).

5. Referencias

[1] D. Boucher, W. Geiselmann, F. Ulmert, *Skew-cyclic codes*, 2006.

[2] D. Boucher, F. Ulmert, *Coding with skew polynomial ring*, 2007.

[3] Sarah A. Spence, *Introduction to Algebraic Coding Theory*, Cornell University, 2008.

[4] Singleton, R.C., *Maximum distance q-nary codes*, IEEE Trans. Inf. Theory 10, 1964.

Referencias web

http://mathworld.wolfram.com/Error-CorrectingCode.html

https://en.wikipedia.org/wiki/Coding_theory