Introducción a la Teoría de Códigos

M.A. García, L. Martínez, T. Ramírez

Facultad de Ciencia y Tecnología. UPV/EHU

Resumen Teórico
Tema 3: CÓDIGOS LINEALES

Mayo de 2017

Tema 3

Códigos Lineales

1 Definición y primeras propiedades

Tanto en este tema como en el siguiente, nos vamos a centrar en los códigos de bloque de longitud n sobre $\mathbb{F}_q$. Esto es, trabajaremos con $C \subseteq \mathbb{F}_q^n$. Para seguir la notación introducida en el tema anterior los elementos de $\mathbb{F}_q^n$ se denotarán por $x_1 \dots x_n$, siendo $x_i \in \mathbb{F}_q$ para $i = 1, \dots, n$, salvo que pueda inducir a error. En tal caso, esto es, cuando pueda inducir a error la notación anterior, se empleará la notación habitual de los elementos de $\mathbb{F}_q^n$: $(x_1, \dots, x_n)$.

Sabemos que $\mathbb{F}_q^n$ es un $\mathbb{F}_q$-espacio vectorial de dimensión n con la suma y la multiplicación por un escalar habitual. Es por ello que, de entre los subconjuntos de $\mathbb{F}_q^n$, nos fijaremos en aquellos que posean alguna estructura algebraica detrás. En concreto, en este tema nos centraremos en el estudio de los códigos conocidos como códigos lineales, que definimos a continuación:

Definición Sea $C \subseteq \mathbb{F}_q^n$. Se dice que C es un **código lineal**, si C es un subespacio vectorial de $\mathbb{F}_q^n$.

Observamos que si $C \subseteq \mathbb{F}_q^n$ es un código lineal, entonces por ser un subespacio vectorial de $\mathbb{F}_q^n$ se verifica:

1. $0 \dots 0 \in C$.
2. Si $\mathbf{x}, \mathbf{y} \in C$, entonces $\mathbf{x}+\mathbf{y}, \mathbf{x}\text{-}\mathbf{y} \in C$.
3. Existe una base $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ de C, siendo $k \leq n$. Además, dado $\mathbf{y} \in C$, existen unos únicos escalares $(\alpha_1, \dots, \alpha_k) \in \mathbb{F}_q^k$ tales que $\mathbf{y} = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i$.
4. Si la dimensión de C es k, entonces $|C| = q^k$.

Si $C \subseteq \mathbb{F}_q^n$ es código lineal de dimensión k, le llamaremos de forma breve (n, k)-código.

CC BY NC SA

Nos interesa ver si podemos calcular de forma sencilla la distancia mínima de un código lineal. Al estar trabajando en $\mathbb{F}_q^n$, lo primero que vemos es que podemos relacionar $d(\mathbf{x}, \mathbf{y})$ con el peso de otra palabra de $\mathbb{F}_q^n$. En concreto, se puede probar

Lema 1.1 *Sean $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}_q^n$. Entonces, $d(\boldsymbol{x}, \boldsymbol{y}) = w(\boldsymbol{x} - \boldsymbol{y})$.*

Este lema nos va a servir para tener otra forma de calcular la distancia mínima de un código lineal, tal y como se establece en la siguiente proposición:

Proposición 1.2 *Sea $C \subseteq \mathbb{F}_q^n$ un código lineal con distancia mínima d y peso mínimo w. Entonces, $d = w$.*

Este resultado va a facilitar enormemente el cálculo de la distancia mínima de un (n, k)-código C sobre $\mathbb{F}_q$, porque en lugar de tener que calcular las distancias de $\binom{q^k}{2}$ pares de palabras de C distintas y luego determinar su mínimo, solamente necesitaremos calcular los pesos de $q^k - 1$ palabras de C no nulas.

Ejemplo El código binario

$$\begin{aligned} C \ = \{&0000000, 0110100, 0011010, 0001101, \\ &1000110, 1001011, 1011100, 0010111, \\ &1010001, 1110010, 011100, 1111111, \\ &0101110, 1101000, 0100011, 1100101\} \end{aligned}$$

es un código lineal de dimensión 4,¡ ya que una base de este código lineal es

$$\{1000110, 0110100, 0011010, 0001101\}.$$

Aplicando la Proposición 1.2, deducimos que la distancia mínima de C es 3, ya que este valor es el peso mínimo de C.

2 Matriz generadora de un código lineal

Dado un (n, k)-código lineal C, sabemos que existe una base $\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$ de C. Estos elementos de la base nos sirven para determinar de forma única todos los elementos de C como combinación lineal de ellos. Si identificamos el elemento $\mathbf{x} = x_1 \ldots x_n \in \mathbb{F}_q^n$ con la matriz $(x_1 \ldots x_n) \in Mat_{1\times n}(\mathbb{F}_q)$, podemos dar la siguiente definición:

Definición Sea C un (n, k)-código lineal sobre $\mathbb{F}_q$. Se llama **matriz generadora** de C a una matriz $G \in \mathrm{Mat}_{k\times n}(\mathbb{F}_q)$ cuyas filas forman una base de C.

Ejemplo El (7,4)-código lineal

$$C = < 0110100, 0011010, 0001101, 1000110 >$$

CC BY NC SA

tiene por matriz generadora a

$$G = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

La ventaja de conocer una matriz generadora G de un (n,k)-código lineal sobre $\mathbb{F}_q$ es que a partir de ella se pueden obtener todas las palabras de C de forma sencilla: basta obtener todos los elementos de $\mathbb{F}_q^k$ y calcular los productos $(y_1 \dots y_k)G$, que serán precisamente las palabras del código C.

Para algunos (n,k)-códigos lineales sobre $\mathbb{F}_q$, de entre las matrices generadoras que podemos hallar vamos a destacar unas con las que va a ser más fácil trabajar:

Definición Sea C un (n,k)-código lineal con matriz generadora $G \in \mathrm{Mat}_{k\times n}(\mathbb{F}_q)$. Se dice que G está dada en **forma estándar** si $G = (I_k|B)$, donde $B \in \mathrm{Mat}_{k\times n-k}(\mathbb{F}_q)$.

Ejemplo El (7,4)-código lineal

$$C = < 0110100, 0011010, 0001101, 1000110 >$$

tiene por matriz generadora a

$$G_1 = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix},$$

que no está dada en forma estándar.

En cambio, si realizamos permutación cíclica de las filas de G_1, obtenemos otra matriz generadora de C,

$$G_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix},$$

que tampoco está dada en forma estándar. Si a G_2 le aplicamos transformaciones elementales por filas, obtenemos otra generadora G_3, que sí está dada en forma estándar:

$$G_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

La ventaja que presentan las matrices generadoras en forma estándar frente a las que no lo están es la siguiente: si C es un código lineal sobre $\mathbb{F}_q$ con matriz generadora G

CC BY NC SA

dada en forma estándar y si la palabra $\mathbf{c}$ de C tiene coordenadas $(y_1 \dots y_k)$ en la base que forman las filas de G, entonces $\mathbf{c} = y_1 \dots y_k c_{k+1} \dots c_n$, esto es, las coordenadas de C en la base formada por las filas de G son precisamente las k primeras letras de $\mathbf{c}$. Es por ello que nos planteamos la siguiente cuestión:

Dado un (n, k)-código lineal C sobre $\mathbb{F}_q$ ¿podemos construir siempre una matriz generadora que esté en forma estándar?

En un principio, si tenemos en cuenta que la matriz generadora G es de dimensión k y esto implica que G es equivalente a una matriz del tipo $(I_K|B)$, podríamos pensar que es factible obtener siempre una matriz generadora en forma estándar para C. Pero, por desgracia, la respuesta a la cuestión planteada es NO. Por ejemplo, si tomamos el código lineal $C =< 100001, 000100 > \subseteq \mathbb{F}_2^6$ cualquier matriz generadora que busquemos no está dada en forma estándar que que las palabras de C siempre llevan un 0 en la segunda posición.

Sin embargo, podemos plantearnos modificar la cuestión anterior y reformularla de la siguiente manera:

Dado un (n, k)-código lineal C sobre $\mathbb{F}_q$ ¿podemos construir un código lineal equivalente a C que admita una matriz generadora que esté en forma estándar?

Al tratar de contestar la cuestión anterior nos damos cuenta que lo primero que debemos hacer es fijarnos en qué operaciones podemos realizar en las palabras del código C para que el código equivalente resultante siga siendo lineal y posteriormente nos centraremos en cómo construir este código equivalente para que además admita matriz generadora estándar. Así, en primer lugar vamos a determinar qué operaciones elementales se pueden realizar para obtener códigos equivalentes al dado que no pierdan la linealidad. En concreto, si se realizan solamente permutaciones en las posiciones del código y multiplicar los símbolos de una posición fija por un escalar no nulo, es obvio que no perdemos la linealidad del código inicial. Estas restricciones que hay que imponer para que se mantenga el carácter de subespacio vectorial del código equivalente resultante se traducen en que no podemos hacer todo tipo de operaciones elementales en la matriz generadora G para transformarla en otra que esté dada en forma estándar y que genere un código lineal equivalente. En concreto, las operaciones elementales en G que nos llevan a matrices generadoras equivalentes a G y que generan un código equivalente a C son:

1. Permutación de filas
2. Multiplicación de una fila por un escalar no nulo
3. Sumar a una fila una combinación lineal de las restantes filas
4. Permutación de columnas
5. Multiplicar cualquier columna por un escalar no nulo

CC BY NC SA

Es decir, de las operaciones elementales que nos permiten obtener una matriz equivalente a G, solamente eliminamos la de sustituir una columna por ella misma más una combinación lineal de las restantes. Esto es lógico porque si pensamos en lo que significa realizar esta operación elemental en términos de las palabras del código, quiere decir que se está mezclando información de varias posiciones.

Observamos que si realizamos solamente las operaciones elementales indicadas en filas lo que estamos haciendo es cambiar una base de C por otra, luego el código obtenido es el mismo. Si además realizamos alguna de las dos operaciones elementales permitidas en columnas, entonces obtenemos otro código lineal que es equivalente al dado. En cualquier caso, si tenemos en cuenta que cualquier matriz generadora de C es de rango k, le podemos aplicar las operaciones elementales anteriores para transformar G en otra matriz equivalente a ella que esté dada en forma estándar. Esta matriz es matriz generadora de un código equivalente a C. En definitiva, teniendo en cuenta lo anterior se prueba

Proposición 2.1 *Sea C un (n,k)-código lineal. Entonces, existe un (n,k)-código lineal C' equivalente a C tal que tiene una matriz generadora dada en forma estándar.*

3 Matriz de control de un código lineal. Código dual de un código lineal

En el espacio vectorial $\mathbb{F}_q^n$ disponemos del producto escalar siguiente:

$$\forall \mathbf{x} = x_1 \dots x_n, \mathbf{y} = y_1 \dots y_n \in \mathbb{F}_q^n,\ \mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i$$

y dado $C \subseteq \mathbb{F}_q^n$ un (n,k)-código lineal sobre $\mathbb{F}_q$, definimos

$$C^{\perp} = \{\mathbf{x} \in \mathbb{F}_q^n \,|\, \mathbf{x} \cdot \mathbf{y} = 0, \forall \mathbf{y} \in C\}.$$

Empleando las propiedades de los subespacios vectoriales, se demuestra:

Proposición 3.1 *Si $C \subseteq \mathbb{F}_q^n$ es un (n,k)-código lineal, entonces $C^{\perp}$ es un código lineal de $\mathbb{F}_q^n$.*

Si $C \subseteq \mathbb{F}_q^n$ es un (n,k)-código lineal, entonces a $C^{\perp}$ se le conoce como el **código dual** de C. Se pueden caracterizar de forma sencilla los elementos de $C^{\perp}$, si conocemos una matriz generadora de C. En efecto, se puede demostrar el siguiente lema

Lema 3.2 *Sea $C \subseteq \mathbb{F}_q^n$ un (n,k)-código lineal con matriz generadora G y $C^{\perp}$ su código dual. Entonces, $\boldsymbol{x} \in C^{\perp}$ si y sólo si $\boldsymbol{x}G^t = 0$.*

BY NC SA

Esta caracterización de los elementos del código dual nos va a servir para calcular la dimensión de $C^\perp$:

Proposición 3.3 *Si $C \subseteq \mathbb{F}_q^n$ es un (n,k)-código lineal, entonces $C^\perp$ es un código lineal de dimensión $n-k$.*

En resumen, si $C \subseteq \mathbb{F}_q^n$ es un (n,k)-código lineal, sabemos que $C^\perp$ es otro código lineal de longitud n y dimensión $n-k$. Por otro lado, $C^\perp$ admite una matriz generadora $H \in \mathrm{Mat}_{(n-k)\times n}(\mathbb{F}_q)$, por ser $C^\perp$ un código lineal. A esta matriz H, generadora de $C^\perp$, se le llama **matriz de control (de paridad)** de C.

Además, de la definición y de las propiedades del código dual, se deduce que si C es un (n,k)-código lineal y $C^\perp$ su código dual, entonces $(C^\perp)^\perp = C$. Entonces, si H es la matriz de control de C, podemos caracterizar C, vía H, de la siguiente manera:

$$C = \{\mathbf{z} \in \mathbb{F}_q^n \,|\, \mathbf{z}H^t = 0\}.$$

Por otro lado, también se cumple que si G y H son matrices generadoras y de control de un (n,k)-código lineal, entonces $GH^t = 0$.

Esta última propiedad nos servirá para localizar una matriz de control cuando la matriz generadora de un código lineal está dada en forma estándar, tal y como enunciamos en el siguiente resultado:

Proposición 3.4 *Sea $C \subseteq \mathbb{F}_q^n$ un (n,k)-código lineal con matriz generadora $G = (I_k|B)$. Entonces, una matriz de control para C es $H = (-B^t|I_{n-k})$.*

Ejemplo Consideremos el código $C = \{(x_1,...,x_n) \in \mathbb{F}_q^n \mid x_1 + \cdots + x_n = 0\}$, que es un $(n, n-1)$-código lineal sobre $\mathbb{F}_q$. Si tomamos el conjunto

$$\{(1,0,0,\ldots,0,0,-1),(0,1,0,\ldots,0,-1),\ldots,(0,0,0,\ldots,0,1,-1)\},$$

resulta que es una base de C, por lo que la matriz

$$G = \begin{pmatrix} 1 & 0 & \ldots & 0 & -1 \\ 0 & 1 & \ldots & 0 & -1 \\ & & \ddots & & \\ 0 & 0 & \ldots & 1 & -1 \end{pmatrix}$$

es una matriz generadora de C que además está dada en forma estándar. Entonces, si aplicamos la Proposición 3.4, la matriz de control de C será

$$H = (1\ 1\ \ldots\ 1\ 1).$$

Esto implica que $C^\perp$ va a ser de dimensión 1 y, obviamente, la condición que debe cumplir un elemento $(x_1,...,x_n)$ de $\mathbb{F}_q^n$ para estar en C es

$$(x_1,...,x_n)H^t = 0 \ \Rightarrow\ x_1 + \cdots + x_n = 0.$$

CC BY NC SA

Finalmente, la matriz de control de un (n,k)-código lineal también nos sirve para determinar la distancia mínima de un código lineal:

Proposición 3.5 *Sea $C \subseteq \mathbb{F}_q^n$ un (n,k)-código lineal con matriz de control H. Entonces, la distancia mínima de C es d si y, solo si, cualesquiera $d-1$ columnas de H son linealmente independientes y existen d columnas de H linealmente dependientes*

De la proposición anterior podemos deducir que el rango de H es al menos $d-1$. Luego si $C \subseteq \mathbb{F}_q^n$ un (n,k)-código lineal con distancia mínima d, se tiene la siguiente desigualdad, conocida como **Cota de Singleton**:

$$d-1 \leq \mathrm{rg}(H) = n-k \;\Rightarrow\; d \leq n-k+1.$$

Ejemplo Si consideremos el código $C = \{(x_1,...,x_n) \in \mathbb{F}_q^n \mid x_1 + \cdots + x_n = 0\}$, que es un $(n,n-1)$-código lineal sobre $\mathbb{F}_q$, sabemos que $H = (1\ 1\ \ldots\ 1\ 1)$ es una matriz de control de C, luego por la Proposición 3.5 deducimos que la distancia mínima de C es 2. Observamos que en este caso se da la igualdad en la Cota de Singleton.

4 Codificación y decodificación de un código lineal

Como ya hemos indicado, si $G \in \mathrm{Mat}_{k\times n}(\mathbb{F}_q)$ es una matriz generadora del (n,k)-código lineal C, podemos obtener de forma sencilla las palabras de C. Basta considerar $(y_1 \ldots y_k) \in \mathbb{F}_q^k$ y calcular $(y_1 \ldots y_k)G$. Precisamente, en esta idea se basa el proceso de codificación de palabras. Así, si para construir nuestros mensajes (antes de codificar) vamos a utilizar un diccionario que consta de $\mathbb{F}_q^k$ palabras, podemos identificar cada una de ellas con un elemento $(y_1 \ldots y_k) \in \mathbb{F}_q^k$. A continuación, elegimos un (n,k)-código lineal C que en cuanto a su capacidad correctora sea adecuado al canal que vamos a emplear y determinamos una matriz generadora de C, que denotamos por G. Entonces, la codificación de la palabra $(y_1 \ldots y_k)$ usando el código elegido será $(y_1 \ldots y_k)G$ y esto será precisamente lo que se envíe a través del canal, cuando en el mensaje inicial aparezca la palabra $(y_1 \ldots y_k)$.

Ilustramos el proceso de codificación con el siguiente ejemplo:

Ejemplo Supongamos que queremos transmitir una fotografía desde Marte a la Tierra. Hemos mandado un equipo que detecta 8 colores básicos que son

$$L = \{\text{blanco, negro, rojo, amarillo, azul, verde, marrón, violeta}\}.$$

El equipo ha sacado la fotografía y ha dividido la misma en cuadrados muy pequeños y lo que quiere transmitir es el color que aparece en cada uno de ellos. Como las

CC BY NC SA

señales que se envían desde Marte solo constan de 0 y 1, se ha identificado cada uno de los colores con una terna de $\mathbb{F}_2^3$, según la siguiente biyección:

$$\begin{array}{rlcl} f: & L & \to & \mathbb{F}_2^3 \\ & \text{blanco} & \mapsto & (1,1,1) \\ & \text{negro} & \mapsto & (0,0,0) \\ & \text{rojo} & \mapsto & (1,0,0) \\ & \text{amarillo} & \mapsto & (0,1,0) \\ & \text{azul} & \mapsto & (0,0,1) \\ & \text{verde} & \mapsto & (1,1,0) \\ & \text{marrón} & \mapsto & (1,0,1) \\ & \text{violeta} & \mapsto & (0,1,1) \end{array}$$

Si mandaramos la información desde Marte utilizando únicamente ternas de $\mathbb{F}_2^3$, no seríamos capaz de detectar si se han producido errores en la transmisión. Esto podría ser un problema porque al estar Marte tan alejado de la Tierra, las señales llegan debilitadas y puede ser que interpretemos de forma errónea el valor que recibimos. Por ello, si únicamente transmitimos desde Marte las ternas asociadas a cada color no tendríamos forma de garantizar que el color que asociamos a un cuadrado determinado sea el verdadero. Pero este problema lo podemos solucionar de forma sencilla usando un código lineal binario de mayor longitud (o sea, que las tuplas usadas para cada color sean más largas) y que tenga dimensión 3, para que conste de 8 palabras. Por ejemplo, podemos usar el (7,3)-código lineal $C =< 0110100, 0011010, 0001101 >$, que es de longitud 7, dimensión 3 y distancia mínima 3. Empleando C, transmitiríamos $(y_1\ y_2\ y_3)G$, donde $G = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$. Así, el color blanco se transmitiría como la tupla 0100011, el color negro sería 0000000, el rojo 0110100, etc. Con esta codificación seríamos capaces de detectar hasta 2 errores y corregir 1, ya que el código C elegido tiene distancia mínima $d = 3$. Por tanto, si sabemos que lo más probable sea que se produzca a lo sumo un error en la transmisión de cada 7-tupla, la elección realizada sería adecuada. En cambio, si lo más probable es que se produzcan 2 o más errores, habría que elegir otro código.

Ahora, entre códigos lineales de la misma longitud y misma capacidad de corrección si uno de ellos tiene matriz generadora $G = (I_k|B)$ dada en forma estándar, su codificación será sencilla porque dado $\mathbf{y} = (y_1 \dots y_k) \in \mathbb{F}_q^k$, entonces $\mathbf{y}G = (y_1 \dots y_k|\mathbf{y}B)$ y los k primeros elementos son las coordenadas en la base de G de la palabra emitida $\mathbf{y}G$. Es por este motivo por lo que se prefiere, siempre que sea posible, el uso de matrices generadoras dadas en forma estándar.

Una vez recibido el mensaje el receptor debe verificar si la palabra recibida está o no en el código usado. Si lo está, por el principio de máxima verosimilitud, supondrá que el mensaje recibido es el que le quería enviar el emisor. Si la palabra recibida no pertenece al código usado, significa que se ha producido al menos un error en la transmisión, por lo que procederá a decodificarla, si es posible. El proceso de decodificación precisamente lo que hace es hallar la palabra que ha sido emitida por

BY NC SA

el emisor, siempre que sea posible. Hay dos métodos de decodificación generales que se emplean para los códigos lineales:

1. Método de decodificación basado en los líderes
2. Método de decodificación mediante síndromes

Ambos métodos, utilizados en códigos lineales, son equivalentes. Esto es, si una palabra recibida $\mathbf{z} \in \mathbb{F}_q^n$ admite una decodificación única y le corresponde como palabra emitida $\mathbf{c} \in C$, esta palabra $\mathbf{c}$ se va a obtener independientemente de cuál sea el método de decodificación que empleemos.

Describimos ahora ambos métodos de decodificación:

4.1 Método de decodificación basado en los líderes

Sea $C \in \mathbb{F}_q^n$ un código lineal. Se define la siguiente relación de equivalencia: $\forall x_1 \cdots x_n, y_1 \cdots y_n \in \mathbb{F}_q^n$

$$x_1 \cdots x_n \sim y_1 \cdots y_n \iff x_1 \cdots x_n - y_1 \cdots y_n \in C.$$

Observamos que

$$\begin{array}{rcl} [x_1 \cdots x_n] & = & \{y_1 \cdots y_n \in \mathbb{F}_q^n | x_1 \cdots x_n \sim y_1 \cdots y_n\} \\ & = & \{y_1 \cdots y_n \in \mathbb{F}_q^n | x_1 \cdots x_n - y_1 \cdots y_n \in C\} \end{array}.$$

Entonces, si recibimos la palabra $z_1 \cdots z_n \in \mathbb{F}_q^n$, calculamos $[z_1 \cdots z_n]$ y como el número de errores producidos en la transmisión lo suponemos mínimo por el principio de máxima verosimilitud, la decodificamos como $z_1 \cdots z_n - y_1 \cdots y_n$, donde $y_1 \cdots y_n \in [z_1 \cdots z_n]$ y es de peso mínimo. Si hay un único $y_1 \cdots y_n \in [z_1 \cdots z_n]$ de peso mínimo, llamaremos a $y_1 \cdots y_n$ **líder** de $[z_1 \cdots z_n]$ y este líder es precisamente el error cometido en la transmisión. Si hay varias $y_1 \cdots y_n \in [z_1 \cdots z_n]$ con el mismo peso mínimo diremos que $z_1 \cdots z_n$ no admite decodificación única.

4.2 Método de decodificación mediante síndromes

Sea $C \in \mathbb{F}_q^n$ un código lineal con matriz de control H. Dada $z_1 \cdots z_n \in \mathbb{F}_q^n$ se llama **síndrome de** $z_1 \cdots z_n \in \mathbb{F}_q^n$ a $S(z_1 \cdots z_n) = (z_1 \cdots z_n)H^t$. Observamos que las palabras del código C satisfacen que su síndrome es $\mathbf{0}$. Además, podemos definir la relación de equivalencia: $\forall x_1 \cdots x_n, y_1 \cdots y_n \in \mathbb{F}_q^n$

$$x_1 \cdots x_n \mathfrak{R} y_1 \cdots y_n \iff S(x_1 \cdots x_n) = S(y_1 \cdots y_n)$$

Entonces, si $x_1 \cdots x_n \mathfrak{R} y_1 \cdots y_n$, se tiene $S(x_1 \cdots x_n) = S(y_1 \cdots y_n)$ lo que implica $S(x_1 \cdots x_n - y_1 \cdots y_n) = \mathbf{0}$, esto es, $x_1 \cdots x_n - y_1 \cdots y_n \in C$. Esto nos proporciona el siguiente método de decodificación:

BY NC SA

Si recibimos la palabra $z_1 \cdots z_n \in \mathbb{F}_q^n$, determinamos $S(z_1 \cdots z_n)$ y decodificamos $z_1 \cdots z_n$ como $z_1 \cdots z_n - y_1 \cdots y_n$ siendo $y_1 \cdots y_n \in \mathbb{F}_q^n$ de peso mínimo tal que $S(z_1 \cdots z_n) = S(y_1 \cdots y_n)$.

Ejemplo Sea C el $(6, 3)$-código binario con matriz de control

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

Si se recibe la tupla $\mathbf{v} = 110101$, entonces $S(110101) = (0\ 0\ 1)$. Pero $S(100000) = (0\ 0\ 1)$, así que la tupla se corrige a

$$u = \mathbf{v} - \mathbf{e} = 010101.$$

En cambio, si recibimos $\mathbf{y} = 100001$ tenemos que $S(\mathbf{y}) = (1\ 1\ 1)$. Pero no hay ningún vector de peso 1 con este síndrome y sí al menos dos de peso 2: 100001 y 001100 con el mismo síndrome, por lo que $\mathbf{y}$ no tiene decodificación única, pues podríamos decodificarla como 000000 ó 101101.

5 Ejemplo de códigos lineales: Códigos de Hamming

En $\mathbb{F}_q^r$ consideramos los subespacios vectoriales de dimensión 1. Por Álgebra Lineal sabemos que hay $\frac{q^r-1}{q-1}$ subespacios. De cada uno de estos subespacios tomamos un vector no nulo (una base) y construimos la matriz H de orden $r \times \frac{q^r-1}{q-1}$ cuyas columnas son precisamente los vectores no nulos seleccionados. Entonces, esta matriz H tiene las siguientes propiedades:

1. Es de rango r.
2. Dos columnas cualesquiera distintas son siempre linealmente independientes y si tomamos dos columnas de H , $H^{(i)}$ y $H^{(j)}$, entonces existe en H una columna $H^{(k)} = H^{(i)} + H^{(j)}$, esto es, $H^{(i)}$, $H^{(j)}$ y $H^{(k)}$ son linealmente dependientes.

Esta matriz H nos sirve como matriz de control de un código lineal de longitud $\frac{q^r-1}{q-1}$, que tiene dimensión $\frac{q^r-1}{q-1} - r$ y, por la construcción de H, distancia mínima 3, llamado $(\frac{q^r-1}{q-1}, \frac{q^r-1}{q-1} - r)$**-código de Hamming**. En algunos textos, se les denomina también códigos de Hamming de parámetros r y q, donde q es el cardinal del cuerpo y r el número de filas de H. Si $q = 2$, observamos que los códigos de Hamming tienen longitud $2^r - 1$ y dimensión $2^r - 1 - r$.

En la definición dada no se ha especificado el orden de las columnas de H. Ello no es obstáculo ya que al cambiar el orden de las columnas de H, lo que se obtiene es otro código de Hamming equivalente al dado.

CC BY NC SA

Esta familia de códigos lineales es una de las más estudiadas y conocidas. Además de poder conocer de antemano su distancia mínima, los códigos de Hamming tienen otra propiedad que los hace especialmente interesantes: son códigos perfectos, puesto que alcanzan la cota de Hamming.

Ejemplo Si queremos construir un código de Hamming ternario con $r = 2$, éste será de longitud $\frac{3^2-1}{3-1} = 4$ y dimensión $\frac{3^2-1}{3-1} - 2 = 4 - 2 = 2$. Además, su matriz de control vendrá dada por

$$H = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{pmatrix}.$$

Es fácil ver que una matriz generadora de este código lineal vendrá dada por

$$G = \begin{pmatrix} 2 & 2 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix},$$

por lo que $C =< 2210, 1201 >$.

CC BY NC SA