Códigos lineales

José Luis Tábara

jltabara@gmail.com

Índice

1. Distancia de Hamming

Todos los cuerpos que aparezcan en estas notas serán finitos. Recordemos que el número de elementos de todo cuerpo finito es siempre una potencia de un primo, siendo ese primo la característica del cuerpo. Además, dada una potencia de un primo, $q = p^n$, existe, salvo isomorfismos, un único cuerpo que posee q elementos. Es costumbre denotar a dicho cuerpo por $\mathbb{F}_q$ o también mediante $GF(q)$ (que proviene del inglés *Galois Field*). Reservaremos la letra q para indicar el cardinal del cuerpo.

Del mismo modo, todos los espacios vectoriales con los que trabajaremos serán de dimensión finita (que denotaremos por n). Como el cuerpo es finito, todos los espacios vectoriales tendrán un número finito de elementos, siendo este número q^n, con las notaciones adoptadas.

En lugar de tratar con espacios vectoriales abstractos, en teoría de códigos se trabaja con el espacio vectorial canónico k^n. Para denotar los vectores de estos espacios vectoriales, siempre que no cause confusión, omitiremos los paréntesis y las comas que separan las componentes. Denotaremos con las letras x, y, z a los elementos del espacio vectorial.

En todo espacio vectorial de la forma k^n se puede definir una función distancia, que convierte a dicho conjunto en un espacio métrico.

Definición 1.1 *Dados dos vectores $x, y \in k^n$, llamamos* distancia (de Hamming) *entre ellos, y denotamos $d(x, y)$, al número de coordenadas donde ambos vectores no coinciden. En fórmulas*

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

Ejemplos.

- Si $x = 1001$ e $y = 0111$ tenemos que $d(x, y) = 3$, pues ambos vectores se diferencian en las tres primeras coordenadas.
- En k^n la distancia entre dos vectores distintos es siempre menor o igual que n y mayor o igual que 1.

- Dado el vector nulo, los vectores que están a distancia 1 de él deben diferenciarse únicamente en una coordenada. Hay $q-1$ vectores con la primera coordenada distinta y otros tantos para cada coordenada. En definitiva, existen $n(q-1)$ vectores de k^n que distan la unidad del elemento neutro. El mismo razonamiento es válido si se utiliza un vector no nulo.

- En cualquier conjunto, en particular en k, se puede introducir la distancia discreta, donde todos los elementos distan entre si la unidad. La distancia de Hamming en k^n es el producto directo de estas distancias. Esto conduce a otra definición, por supuesto equivalente, de distancia de Hamming, dada por la fórmula

$$d(x_1x_2\cdots x_n, y_1y_2\cdots y_n) = d(x_1,y_1) + d(x_2,y_2) + \cdots + d(x_n,y_n)$$

Sea E un espacio vectorial y $\phi : E \to k^n$ un isomorfismo. Gracias a este isomorfismo, podemos introducir una distancia de Hamming en el espacio vectorial E. Pero si $\phi' : E \to k^n$ es otro isomorfismo, la distancia de Hamming introducida varía. El concepto de distancia de Hamming no es invariante por isomorfismos. No existe un concepto de distancia de Hamming en un espacio vectorial abstracto.

La definición que hemos introducido es una distancia en el sentido topológico del término. En particular cumple las siguientes propiedades:

- Es simétrica: $d(x,y) = d(y,x)$.
- Es positiva: $d(x,y) \geq 0$.
- Solo se anula sobre vectores iguales: $d(x,y) = 0$ si y solo si $x = y$.
- Cumple la desigualdad triangular: $d(x,y) \leq d(x,z) + d(z,y)$.

Las tres primeras propiedades son prácticamente evidentes. Para demostrar la última vamos a dar otra interpretación de la distancia. Dados dos

vectores x e y la distancia entre ellos es el mínimo número de coordenadas que debemos cambiar para transformar x en y. Ahora transformamos x primeramente en z, lo que implica $d(x,z)$ cambios de coordenadas y posteriormente transformamos z en y, aplicando $d(z,y)$ cambios. Naturalmente $d(x,y)$ debe ser menor o igual que la suma de estos dos números.

Sabemos que en muchos espacios vectoriales la distancia, que es una función de dos variables, se puede construir a partir de una norma, que es una función de una única variable. En este caso tenemos algo análogo a la norma.

Definición 1.2 *El* peso *de un vector $x \in k^n$ es el número de coordenadas no nulas que posee dicho vector. Se denota por $\omega(x)$*

$$\omega(x) = |\{i \text{ tales que } x_i \neq 0\}|$$

Ejemplos.

- El único vector cuyo peso es cero es el vector nulo. En k^n todo vector no nulo tiene un peso comprendido entre 1 y n.
- El peso se puede definir en función de la distancia

$$\omega(x) = d(x,0)$$

- El peso tiene propiedades similares a la norma. En particular cumple que $\omega(x+y) \leq \omega(x) + \omega(y)$, que es la traducción de la desigualdad triangular.
- En el caso binario los vectores de peso 1 son los vectores de la base canónica. Existen $\binom{n}{2}$ vectores de peso 2 y en general $\binom{n}{i}$ vectores de peso i.

La relación entre distancia y peso viene dada en el

Lema 1.1 *Dados dos vectores $x, y \in k^n$ se tiene que*

$$d(x,y) = \omega(x-y)$$

Demostración.

Dados x e y tenemos que la coordenada i-ésima de $x - y$ es nula si y solo si $x_i = y_i$. $\square$

En el caso de que la característica del cuerpo sea 2 la resta y la suma coinciden, pudiendose calcular las distancias con el

Corolario 1.2 *Si* $char(k) = 2$ *entonces*

$$d(x, y) = \omega(x + y)$$

En k^n, además de la estructura vectorial, podemos considerar la multiplicación de vectores componente a componente. Denotaremos dicho producto (no es una notación estandard) mediante $x * y$. En el caso binario la componente i-ésima del producto es la unidad si y solo si las componentes i-ésimas de cada vector son la unidad.

Corolario 1.3 *En el caso binario*

$$d(x, y) = \omega(x) + \omega(y) - 2\,\omega(x * y)$$

Demostración.

La distancia es igual a $\omega(x + y)$, que es la cantidad de unos de la suma de dos vectores. Fijémonos en una coordenada i. Si en dicha coordenada únicamente uno de los vectores tiene un 1, entonces aparece un 1 en la suma. Sin embargo, si en dicha coordenada los dos vectores tienen un 1, en la suma aparece un 0. Pero precisamente los dos vectores tienen un 1 en una coordenada cuando $x*y$ tiene esa coordenada no nula. Por lo tanto el número de unos de $x + y$ coincide con la suma de los unos de x más la suma de los unos de y menos el doble del número de unos de $x * y$. $\square$

Ejemplo.

- Sea $x = 10011$ e $y = 00111$. Entonces $x + y = 10100$ y $x * y = 00011$.

Tenemos que

$$\omega(x+y) = 2 = \omega(x) + \omega(y) - \omega(x * y) = 3 + 3 - 2 \cdot 2$$

El espacio k^n junto con esta métrica es un espacio topológico. Sin embargo, al estar todos los vectores a distancia mayor o igual a la unidad, dicho espacio es siempre discreto y topológicamente no tiene interés.

Como en cualquier espacio métrico, tiene sentido considerar las bolas. En nuestro caso las bolas abiertas y cerradas coinciden, por lo que siempre consideraremos bolas cerradas. Denotaremos por $B_x(r)$ a la bola cerrada de centro x y de radio r:

$$B_x(r) = \{y \in k^n \text{ tales que } d(x,y) \leq r\}$$

En este espacio métrico todas las bolas tienen un número finito de elementos. Vamos a contar exactamente cuantos tienen.

Lema 1.4 *El número de elementos de la bola de radio r es*

$$|\, B_x(r)| = 1 + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \cdots + \binom{n}{r}(q-1)^r$$

Demostración.

La bola de radio cero tiene claramente un único elemento. La bola de radio unidad contiene al centro de la bola y a todos los elementos cuya distancia es exactamente la unidad. Como hay n coordenadas, existen n maneras distintas de elegir una coordenada. En cualquiera de dichos lugares podemos colocar cualquier elemento del cuerpo, con la excepción del que tiene el vector x en dicha posición. Hemos visto que la bola de radio 1 tiene $1 + n(q-1)$ elementos. Para la bola de radio dos debemos sumar a éstos los elementos que están a distancia dos. Esto es, elementos que se diferencia de x en exactamente dos posiciones. Primero elegimos las dos posiciones. Para ello formamos el número combinatorio $\binom{n}{2}$ y después rellenamos dichas posiciones con los $(q-1)$ elementos posibles. Se concluye por inducción. $\square$

En teoría de códigos el cuerpo más utilizado es el de dos elementos. En este caso tenemos el

Corolario 1.5 *Si $q = 2$ entonces*

$$|\, B_x(r)| = 1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{r} = \sum_{i=0}^{r} \binom{n}{i}$$

Ejemplos.

- Sea $q = 2$. La bola de radio 2 en k^{90}

$$|\, B_x(2)| = 1 + 90 + \frac{90 \cdot 89}{2} = 4096$$

- Si $q = 3$, la bola de radio 2 en k^{11}

$$|\, B_x(2)| = 1 + 11 \cdot 2 + \frac{11 \cdot 10}{2}\, 2^2 = 1 + 22 + 220 = 243$$

- Si $q = 2$ la bola de radio n necesariamente contiene a todos los elementos del espacio vectorial. Se obtiene la fórmula de sumación

$$\sum_{i=0}^{n} \binom{n}{i} = 2^n$$

2. Códigos

Si queremos transmitir un mensaje a través de un canal, debemos escribir nuestro mensaje en una forma subceptible de ser enviada por dicho canal. Para ello debemos traducir nuestro mensaje a un alfabeto $\mathcal{A}$ que pueda ser transmitido a través del canal. Si el canal es digital, solamente puede transmitir un número finito de símbolos, lo que nos impone que el alfabeto sea finito. Cada mensaje a transmitir estará entonces formado por un conjunto de palabras, siendo cada palabra una sucesión de elementos del alfabeto $\mathcal{A}$. En principio, las palabras escritas en el nuevo alfabeto pueden tener distintas longitudes, pero para la corrección de errores es conveniente que todas tengan la misma longitud. Para aprovechar toda la potencia de las matemáticas, introducimos una estructura algebraica en el alfabeto considerado. El conjunto de todas las palabras escritas en el nuevo alfabeto es lo que consideramos un código. Formalizamos todas estas intuiciones en la

Definición 2.1 *Un código es un subconjunto* $\mathcal{C} \subset k^n$.

Si el cuerpo tiene q elementos decimos que $\mathcal{C}$ es un código q-ario. En particular, llamamos códigos binarios y ternarios a los que tienen $q = 2$ y $q = 3$ respectivamente.

El número de elementos de un código es siempre finito. En general denotaremos por M al cardinal del código. Si un código tiene M vectores y cada vector tiene longitud n, decimos que se trata de un código de tipo (n, M). Para hacer más fluido el lenguaje, y por analogía con el lenguaje escrito, llamaremos palabras (*codewords* en inglés) a los elementos del código.

En todo código se puede considerar la distancia heredada de los espacios vectoriales que los contienen.

Definición 2.2 *Llamamos* distancia mínima *del código a la menor de las distancias entre dos palabras distintas del código. Lo denotamos por* $d_{\mathcal{C}}$ *o simplemente por* d. *En fórmulas*

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

En vista de nuestra definición, existen dos palabras $x, x' \in \mathcal{C}$ tales que $d(x, x') = d_{\mathcal{C}}$. Además, para cualquier par de palabras del código se tiene que $d(x, y) \geq d_{\mathcal{C}}$. Si la distancia mínima de un código es d diremos que es un código de tipo (n, M, d).

Ejemplos.

- Sea $\mathcal{C} = \{0000, 1111\}$. Tenemos que $n = 4$ pues esa es la longitud de las palabras del código. $M = 2$ por tener dos elementos. Además es claro que $d_{\mathcal{C}} = 4$. Hemos construido un ejemplo de $(4, 2, 4)$-código.

- Códigos triviales. Sea $\mathcal{C} = k^n$. Este código tiene como parámetros $M = q^n$ y $d = 1$. Consideremos también el código de repetición, formado por los múltiplos del vector $11\cdots1$. En este caso $M = q$ y la distancia es $d = n$.

- Consideremos el código trivial

$$\mathcal{C} = \{00, 01, 10, 11\}$$

  A cada palabra de este código le vamos a añadir un dígito. Le añadiremos un 0 si el peso de la palabra es par y un 1 si el peso de la palabra es impar. Obtenemos de esta forma un nuevo código

$$\mathcal{C}' = \{000, 011, 101, 110\}$$

  que es de tipo $(3, 4, 2)$ (comprobar que la distancia mínima es 2). Se dice que $\mathcal{C}'$ se ha construido añadiendo un dígito de paridad al código $\mathcal{C}$.

- Códigos con bit de paridad. Sigamos con $q = 2$ y generalicemos el ejemplo anterior. Dado $x \in k^n$ construimos $\hat{x} \in k^{n+1}$ añadiendo un 0 si $\omega(x)$ es par y un 1 si $\omega(x)$ es impar. Esto establece una biyección de k^n con un subconjunto $\mathcal{C}$ de k^{n+1} que será nuestro código. Se tiene que $\mathcal{C}$ es de tipo $(n+1, 2^n)$. Calculemos ahora su distancia mínima. Por construcción $\omega(\hat{x})$ es siempre par. Aplicando el corolario 1.3

$$d(\hat{x}, \hat{y}) = \omega(\hat{x} + \hat{y}) = \omega(\hat{x}) + \omega(\hat{y}) - 2\omega(\hat{x} \cdot \hat{y})$$

y la distancia entre palabras del código es siempre par. La distancia mínima es dos. Esta misma construcción es válida para cualquier código binario y se obtiene siempre un código con distancia mínima par.

- Códigos extendidos. Sea $\mathcal{C}$ un código sobre un cuerpo con q arbitrario. A cada elemento del código se le añade un nuevo dígito, de tal forma que al sumar todas las coordenadas, incluyendo el nuevo dígito, la suma sea cero. Como ejemplo, si $q = 7$ y el vector es $x = 1324$, lo debemos transformar en $\hat{x} = 13244$, puesto que $1+3+2+4+4 = 0$, realizando las operaciones módulo 7. En general se debe seguir la siguiente regla para construir el código extendido

$$x = (x_1, x_2, \ldots, x_n) \quad \longrightarrow \quad \hat{x} = (x_1, x_2, \ldots, x_n, -\textstyle\sum_{i=1}^{n} x_i)$$

  Los códigos con bit de paridad son los códigos extendidos en el caso particular $q = 2$.

Los códigos se emplean normalmente para transmitir información a través de canales. Supongamos que el emisor envía una palabra $x \in \mathcal{C}$. Si el receptor recibe esa misma palabra, no existe ningún problema y puede interpretar correctamente la palabra emitida. Pero si el canal deforma la palabra x, llegando al receptor la palabra y, decimos que se han producido errores en la comunicación. Indicaremos este hecho con la notación $x \rightsquigarrow y$.

Definición 2.3 *Sea $x \rightsquigarrow y$. Llamamos* error *al vector $y - x$.*

Si en la transmisión de una palabra se comete solamente un error, la palabra y tiene una coordenada distinta de la palabra x. Si se cometen dos errores variarán dos coordenadas y en general tenemos la

Definición 2.4 *Dado $x \rightsquigarrow y$, el* número de errores *cometido es $d(x, y)$.*

Ejemplos.

- El receptor y el emisor se ponen de acuerdo en la utilización del código $\mathcal{C} = \{000, 111\}$. Si el emisor envía $x = 000$ y al atravesar el canal se transforma en $y = 010$, el receptor percibe que ha habido algún error, puesto que lo que ha recibido no es una palabra del código. En general, si en la transmisión se producen 1 ó 2 errores en cualquier palabra que se transmita, el receptor detecta que la palabra recibida no es del código. Nuestro código es capaz de detectar dos errores.

- Supongamos que ahora el código es $\mathcal{C} = k^n$. Si en la transmisión se produce un error, el receptor no lo puede detectar, puesto que lo que recibe es también una palabra del código. Este código no detecta ningún error.

- Sea $\mathcal{C}$ el código binario obtenido añadiendo un dígito de paridad al código trivial. Si sumamos todos los 1 de cualquier palabra del código se obtiene un número par. Si en la transmisión se produce un solo error, la palabra recibida tiene un número impar de 1, lo que nos permite detectar el error.

Teorema 2.1 *Sea $\mathcal{C}$ un código de distancia mínima d. Entonces $\mathcal{C}$ detecta $d-1$ errores.*

Demostración.

Dado la transmisión $x \rightsquigarrow y$, al cometer $d-1$ o menos errores tenemos que $d(x, y) < d$. La palabra y no puede pertenecer al código puesto que contradice la definición de distancia mínima. □

Volvamos a nuestro código $\mathcal{C} = \{000, 111\}$. Si se recibe 010 sabemos que existe un error. Además parece que lo más probable es que el emisor haya enviado la palabra 000 y no la 111. Ello es debido a que en el primer caso se ha producido un error en la transmisión, mientras que en el segundo caso serían necesarios dos errores. Nosotros consideraremos siempre canales donde es poco probable que se produzcan errores. Pero supongamos que el código

es $\mathcal{C} = \{00, 11\}$ y que se recibe 01. Ahora el receptor no tiene ningún motivo para decidir entre una palabra u otra. En este caso lo más normal es que el receptor decida no decodificar el mensaje.

Definición 2.5 *Dado un código $\mathcal{C} \subset k^n$, una* regla de decisión *es una función*

$$f : k^n \to \mathcal{C} \cup \{?\}$$

Como el canal tiene ruido, se puede recibir cualquier palabra $y \in k^n$. Para decodificarla el receptor emplea una regla de decisión, asignándole o bien la palabra $f(y) \in \mathcal{C}$ o bien el símbolo ? que significa que no "sabe" decodificar la palabra.

Decimos que una regla de decisión es capaz de corregir e errores si decodifica correctamente todas las palabras que tienen un número de errores menor o igual que e. La formulación precisa de este concepto viene recogida en la

Definición 2.6 *Una regla de decisión f corrige e errores si para toda transmisión $x \rightsquigarrow y$ tal que $d(x, y) \leq e$ se tiene que $f(y) = x$.*

Ejemplos.

- La regla que detecta errores. Construimos f de la siguiente forma

$$f(y) = \begin{cases} y \text{ si } y \in \mathcal{C} \\ ? \text{ si } y \notin \mathcal{C} \end{cases}$$

  Si el receptor aplica esta regla de decisión detecta cualquier error en la transmisión, pero no puede corregir ningún error. En efecto, si $x \rightsquigarrow y$ y $d(x, y) \geq 1$ entonces $f(y) =?$.

- La regla del vecino más cercano. Si existe una única palabra del código $x \in \mathcal{C}$ a distancia mínima de y, entonces $f(y) = x$. Si existen dos o más palabras del código que están a la misma distancia de y entonces $f(y) =?$. Veremos posteriormente que esta regla de decisión permite corregir errores.

- La regla de las esferas. Consideramos las esferas de radio 0 centradas en las palabras del código. Claramente son disjuntas. Ahora consideramos las esferas de radio unidad. Si éstas son disjuntas seguimos aumentando el radio. Este proceso de aumentar el radio manteniendo disjuntas las esferas debe detenerse en algún momento. El mayor radio que permite que las esferas sean disjuntas, se llama radio de empaquetamiento del código. Sea t cualquier valor menor o igual que el radio de empaquetamiento. Si $y \in B_x(t)$ entonces $f(y) = x$. Si y no pertenece a ninguna de las esferas, entonces $f(y) =?$. En particular la primera de las reglas que hemos mencionado es un caso particular de este ejemplo con $t = 0$. Veremos que esta regla permite corregir errores si el radio es mayor o igual que uno.

- Regla de las esferas II. Supongamos nuevamente que las esferas centradas en palabras del código y de radio t son disjuntas. Debido a esta hipótesis, si tomamos un $y \in k^n$ arbitrario, la esfera centrada en y y de radio t contiene, como mucho, una palabra del código. Para obtener $f(y)$ calculamos la esfera de radio t centrada en y. Si dicha esfera contiene a una palabra x del código, entonces $f(y) = x$. Si no contiene a ninguna palabra entonces $f(y) =?$. En realidad las dos reglas de las esferas producen la misma regla de decisión.

Teorema 2.2 *Sea $\mathcal{C}$ un código con distancia mínima d. Si $d \geq 2t + 1$ entonces la regla del vecino más cercano permite corregir t errores.*

Demostración.

Supongamos que $x \rightsquigarrow y$ y que $d(x, y) \leq t$. Veamos que todas las demás palabras del código están a una distancia superior. Sea x' otra palabra del código. Aplicando la desigualdad triangular

$$d(x', y) \geq d(x, x') - d(x, y)$$

Como x y x' pertenecen al código la distancia entre ellas debe ser al menos d. Entonces

$$d(x', y) \geq d - t \geq (2t + 1) - t = t + 1$$

Por lo tanto $f(y) = x$ y se corrigen los errores. $\square$

Ejemplos.

- El código $\mathcal{C} = \{000, 111\}$ permite detectar 2 errores y corregir un error.
- Sea el código $\mathcal{C} = \{000000, 100000, 111111\}$. Si tenemos

$$111111 \rightsquigarrow 001111$$

entonces la regla del vecino más cercano nos dice que

$$f(001111) = 111111$$

En este caso se corrigen dos errores. Pero este código no corrige ningún error, puesto que la propiedad debe ser cierta para todas las palabras y todos los posibles errores. Si tenemos

$$000000 \rightsquigarrow 100000$$

observamos que ni tan siquiera detecta un error.

Teorema 2.3 *Sea $\mathcal{C}$ un código de distancia mínima $d \geq 2t + 1$. El método de las esferas de radio t permite detectar t errores.*

Demostración.

Veamos que las esferas de radio t centradas en las palabras del código son disjuntas. Supongamos que y pertenece a dos de estas esferas. Aplicando la desigualdad triangular

$$d(x, x') \leq d(x, y) + d(y, x') \leq t + t = 2t$$

lo que contradice la definición de distancia mínima. Hemos visto que las esferas son disjuntas y tiene sentido la regla de decisión.

Veamos que corrige e errores. Sea $x \rightsquigarrow y$ con $d(x, y) \leq t$. Entonces $y \in B_x(t)$ y $f(y) = x$. $\square$

Si d es impar, entonces $d = 2t + 1$ y permite corregir t errores. Si d es par, entonces $d = 2t + 2$ y permite corregir t errores. Existe una fórmula que nos proporciona el número t de errores que puede corregir cualquier código, independientemente de si d es par o impar

$$t = \left\lfloor \frac{d-1}{2} \right\rfloor$$

donde $\lfloor\ \rfloor$ denota la parte entera de un número.

3. Cotas del tamaño de los códigos

Los parámetros fundamentales de todo código son la longitud, el tamaño y la distancia mínima. El valor de alguno de estos parámetros normalmente limita el posible valor de otros. Por ejemplo, si n es la longitud del código, no es posible que d sea mayor que n. De la misma forma no es posible que M supere el valor de q^n. Estas limitaciones en los parámetros, dadas en forma de desigualdades, se denominan cotas. Examinaremos ahora algunas de las cotas más importantes.

Hemos demostrado en el teorema anterior que si $d \geq 2t+1$ las esferas de radio t centradas en las palabras del código son disjuntas dos a dos. Por lo tanto la suma de todos los elementos de las esferas debe ser menor o igual que el número de elementos del espacio vectorial que los contiene.

Corolario 3.1 (Cota de Hamming) *Si $d \geq 2t+1$, se tiene la desigualdad*

$$|\,\mathcal{C}| \cdot |\,B_x(t)| \leq q^n$$

Si escribimos esta desigualdad en función de los parámetros obtenemos

$$|\,C| \cdot \left(\sum_{i=0}^{t} \binom{n}{i} (q-1)^i \right) \leq q^n$$

Esta desigualdad también se denomina cota de empaquetamiento de esferas (*sphere-packing bound* en inglés). En general la desigualdad es estricta. Aquellos códigos donde se da la igualdad son una clase especial de códigos.

Definición 3.1 *Decimos que un código es* perfecto *si se alcanza la cota de Hamming.*

Si el código es de tipo $(n, M, 2t+1)$ o de tipo $(n, M, 2t+2)$, los códigos perfectos son aquellos que cumplen la llamada ecuación de Hamming

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

Ejemplos.

- El código $\mathcal{C} = k^n$ es perfecto, puesto que

$$q^n \cdot 1 = q^n$$

- El código de repetición binario es el formado los vectores $00\cdots0$ y $11\cdots1$. Es de tipo $(n, 2, n)$ pues dos vectores se diferencian siempre en todas las coordenadas. Si $n = 2t + 1$ se tiene

$$2 \cdot \left(\sum_{i=0}^{t} \binom{n}{i} \right) = 2 \cdot \frac{2^n}{2} = 2^n$$

Salvo los códigos triviales, no existen muchos ejemplos de códigos perfectos. Si $d = 3$ existe una colección numerable de soluciones de la ecuación de Hamming. Para cada una de estas soluciones se puede construir un código con esos parámetros. Son los famosos códigos de Hamming, que analizaremos en una sección posterior.

Pero si d es mayor que 3 solamente existen tres soluciones. Para dos de ellas se pueden construir códigos asociados, pero para la otra no existe ningún código con dichos parámetros. Estas dos nuevas soluciones dan lugar a los códigos de Golay. El primero es un código binario, que se denota $\mathcal{G}_{23}$ y es de tipo $[23, 12, 7]$. El segundo, denotado $\mathcal{G}_{11}$, es un código ternario de tipo $[11, 6, 5]$. Analizaremos su construcción en la sección de los códigos cíclicos.

Los códigos perfectos presentan algunas características que los hacen interesantes. Veamos algunas:

- Todo elemento de k^n está en alguna bola centrada en alguna palabra del código. La regla de decisión del vecino más cercano siempre proporciona un elemento del código y nunca le ocurre que $f(y) =?$.

- El conjunto de códigos de longitud n y distancia d está ordenado por inclusión. Dado un código perfecto $\mathcal{C} \in k^n$ de distancia mínima d no

puede existir ningún otro código $\mathcal{C}'$ tal que $\mathcal{C} \subset \mathcal{C}' \subset k^n$ que tenga la misma distancia mínima. Los códigos perfectos son elementos maximales respecto a esta relación de orden. Como estamos trabajando con un número finito de elementos, la existencia de elementos maximales está garantizada.

- Dados n y d es interesante encontrar códigos con la M lo más grande posible. Denotaremos mediante $A_q(n, d)$ a dicho valor de M. Este problema se denomina a veces como el problema principal de la teoría de códigos. El valor de M está acotado superiormente por tener que cumplir la desigualdad de Hamming. Los códigos perfectos son soluciones a este problema.

Ya hemos visto una cota superior al valor de M en un código maximal. Veremos ahora una cota mínima.

Proposición 3.2 (Cota de Gilbert-Varshamov) *Si $\mathcal{C} \subset k^n$ es un código maximal de distancia mínima d, entonces*

$$M \geq \frac{q^n}{|\, B_x(d-1)|}$$

Demostración.

Si $\mathcal{C}$ es maximal entonces el conjunto de bolas de radio $d-1$ y centradas en las palabras del código debe recubrir todo el espacio. Si existiese un vector y que no estuviese en ninguna de esas bolas, entonces el código $\mathcal{C}' = \mathcal{C} \cup \{y\}$ tendría un elemento más y la misma distancia mínima. El número de elementos de estas bolas debe ser mayor que el número de elementos de todo el espacio, lo que da lugar a

$$M \cdot |\, B_x(d-1)| \geq q^n$$

y queda demostrado. □

Este teorema garantiza que siempre existen códigos que superan dicha cota. Además nos indica un método que podemos seguir para su construcción:

si tenemos un código que no cumple la cota, le podemos añadir un nuevo elemento. Sin embargo este método es efectivo en la práctica solamente para valores pequeños de n.

A pesar de que existen muchas otras cotas, vamos a estudiar únicamente otra más.

Proposición 3.3 (Cota de Singleton) *Sea $\mathcal{C}$ un código de tipo (n, M, d). Entonces*

$$M \leq q^{n-d+1}$$

Demostración.

Consideremos la proyección $\phi : k^n \to k^{n-d+1}$, consistente en "borrar" las $d-1$ últimas coordenadas. La aplicación ϕ restringida al código es inyectiva, pues si $\phi(x) = \phi(y)$, estos dos vectores tendrían, como mucho, distintas las últimas $d-1$ coordenadas. El número de elementos del código debe ser menor o igual que el número de elementos del espacio de llegada. $\square$

4. Códigos lineales. Introducción

El conjunto k^n se puede dotar de distintas estructuras algebraicas. Dependiendo de la estructura elegida, los subconjuntos que conservan dicha estructura presentan características especiales que los hacen más sencillos de estudiar que los subconjuntos generales. La primera estructura que se puede introducir en k^n es la de grupo. Los subconjuntos de k^n que sean subgrupos forman una clase especial de códigos. La otra estructura natural que se suele asociar con k^n es la de espacio vectorial, y es la que nos interesa en esta sección.

Definición 4.1 *Un* código lineal *es un subespacio vectorial* $\mathcal{C} \subset k^n$.

Si el subespacio es de dimensión m, y la distancia mínima es d, diremos que es de tipo $[n, m, d]$ (con corchetes). Es claro que este código es también de tipo (n, q^m, d). Sin embargo el recíproco no es cierto: existen códigos de tipo (n, q^m, d) que no son lineales.

En el estudio de los códigos lineales la utilidad del álgebra lineal es evidente. Repasemos ahora algunos resultados.

- Dada una colección de vectores $v_1, v_2, \ldots, v_n$, el subespacio vectorial generado por ellos se denota $\langle v_1, v_2, \ldots, v_n \rangle$ y es el menor subespacio que contiene a todos los vectores. Si los vectores son linealmente dependientes, podemos ir quitando uno a uno vectores que dependan linealmente del resto. Tras realizar este proceso el subespacio vectorial generado es el mismo, pero ahora los vectores son linealmente independientes. Hemos extraido una base del subespacio vectorial.

- Sea $\varphi : E \to E'$ una aplicación lineal. La imagen y el núcleo de esta aplicación son subespacios vectoriales. Además se cumple que

$$\dim(\mathrm{Im}(\varphi)) + \dim(\mathrm{Ker}(\varphi)) = \dim(E)$$

  En el caso de los espacios vectoriales canónicos las aplicaciones lineales vendrán dadas por matrices.

- Si α es una forma lineal ($\alpha \in E^*$), el conjunto

$$\{v \in E \text{ tales que } \alpha(v) = 0\}$$

es un subespacio vectorial. En general, dado un subespacio vectorial del dual $F \subset E^*$, su incidente

$$F^0 = \{v \in E \text{ tales que } \alpha(v) = 0 \text{ para todo } \alpha \in F\}$$

es un subespacio vectorial. Se cumple la fórmula de las dimensiones

$$\dim(F) + \dim(F^0) = \dim(E)$$

- Si tenemos varios subespacios vectoriales, su intersección y su suma siguen siendo subespacios vectoriales. Tenemos también una fórmula que liga las dimensiones

$$\dim(F + F') = \dim(F) + \dim(F') - \dim(F \cap F')$$

Pasemos ya al caso concreto de los códigos.

Ejemplos.

- Existen algunos códigos lineales que por su simplicidad se denominan triviales. El primero es $\mathcal{C} = k^n$. Es de tipo $[n, n, 1]$. No tiene interés pues no permite ni detectar un error. El otro es el código de repetición, que está formado por todos los múltiplos del vector $11\ldots1$. Este es de tipo $[n, 1, n]$, pues dos de sus elementos siempre se diferencian en todas las coordenadas. El código nulo es también trivial.

- En el caso de que $q = 2$ las combinaciones lineales son simplemente sumas. Un subconjunto $\mathcal{C}$ es un código lineal si para todo $x, y \in \mathcal{C}$ se tiene que $x + y \in \mathcal{C}$. En este caso los códigos lineales coinciden con los subgrupos de k^n.

- Dado $q = 2$ y el espacio vectorial k^3. Los elementos que cumplen

$$x_1 + x_2 + x_3 = 0$$

forman un subespacio vectorial de dimensión 2. Está formado por 4 elementos, que son precisamente los elementos de peso par.

- Sea $\mathcal{C} \subset k^n$ un código lineal de dimensión m. Consideremos el hiperplano F que tiene como ecuación implícita $x_1 + x_2 + \cdots + x_n = 0$. Aplicando la fórmula de las dimensiones, tenemos que la intersección $\mathcal{C} \cap F$ es un nuevo código, cuya dimensión es o bien m (esto ocurre si $\mathcal{C} \subset F$) o bien $m - 1$. La distancia mínima del nuevo código es siempre mayor o igual que la del código de partida. El ejemplo anterior es un caso particular.

- Sea F_i el hiperplano de ecuación $x_i = 0$. Análogamente la intersección con un código es un nuevo código. La dimensión y la distancia mínima del nuevo código se analiza como en el ejemplo anterior.

- Otra manera cómoda de indicar un código lineal es dar una base, pues a través de ella se pueden construir todas las palabras del código. Sea $q = 3$ y sea $\mathcal{C}$ el subespacio vectorial de k^4 generado por los vectores $u = 1011$ y $v = 0112$. Los nueve elementos de este código son:

$$\mathcal{C} = \{0000, 1011, 0112, 2022, 1221, 1120, 2210, 1202, 2101\}$$

Algunos autores prefieren definir los códigos de otra manera. Consideran un espacio vectorial k^m e identifican sus elementos con las palabras sin codificar. Posteriormente inyectan dicho conjunto en un espacio vectorial k^n mediante una aplicación $c : k^m \to k^m$. La imagen de cada elemento es la palabra codificada y la aplicación c se llama la regla de codificación. El problema que tiene esta definición es que todos los códigos se ven forzados a tener q^n elementos. Sin embargo en el caso lineal no existe tal problema y se puede optar por la siguiente

Definición 4.2 *Un* código lineal *de tipo* $[n, m]$ *es una aplicación lineal inyectiva* $\varphi : k^m \to k^n$.

Para nosotros el código es precisamente la imagen de la aplicación lineal. Si en un subespacio vectorial $\mathcal{C}$ tomamos una base, se puede construir una aplicación lineal $\varphi : k^m \to k^n$ cuya imagen sea $\mathcal{C}$, sin más que asociar a cada vector de la base canónica el vector correspondiente de la base. Hemos probado que ambas definiciones son equivalentes. Si cambiamos la base, cambiamos la aplicación lineal. Estamos empleando el mismo código pero hemos modificado la regla de codificación.

Los elementos de k^m son las palabras que nos interesa transmitir. Para ello las codificamos, prevenimos errores en la transmisión y finalmente las decodificamos. De las n coordenadas que tiene una palabra del código podemos pensar que solamente m de ellas transportan información y que el resto de las coordenadas sirven únicamente para corregir los errores.

Definición 4.3 *Dado un código lineal de tipo* $[n, m]$*, llamamos* tasa de transmisión *del código al cociente (la* R *proviene del inglés* rate*.)*

$$R = \frac{m}{n}$$

Es claro que esta definición no tiene sentido para códigos arbitrarios. Reformulemos la cuestión. El código $\mathcal{C}$ tiene q^m palabras y el espacio que lo contiene tiene q^n palabras. Podemos entonces definir la tasa de transmisión del código mediante la regla

$$R = \frac{m}{n} = \frac{\log_q(q^m)}{\log_q(q^n)} = \frac{\log_q(|\,\mathcal{C}|)}{n}$$

que ya se puede aplicar a códigos arbitrarios. La codimensión del código (el número $n - m$) se llama redundancia del código.

En general en un código interesa que la tasa de transmisión sea lo mayor posible, pues de esta manera se aprovecharía al máximo el ancho de banda del canal. Ello va en detrimento de la distancia mínima que debe ser en-

tonces pequeña. Dependiendo del tipo canal debemos favorecer una u otra característica.

Definición 4.4 *El* peso mínimo $\omega_{\mathcal{C}}$ *de un código lineal es el menor de los pesos no nulos.*

$$\omega_{\mathcal{C}} = \min_{x \neq 0} \omega(x)$$

En el caso de los códigos lineales la distancia mínima se puede calcular en función del peso sus elementos

Proposición 4.1 *Si* $\mathcal{C}$ *es lineal, entonces* $d_{\mathcal{C}} = \omega_{\mathcal{C}}$.

Demostración.

Como $\omega(x) = d(x, 0)$, necesariamente el peso mínimo es mayor que la distancia mínima.

Sean x e y dos palabras tal que $d(x, y) = d_{\mathcal{C}}$. Como

$$d(x, y) = \omega(x - y)$$

y la palabra $x - y$ pertenece al código por éste lineal, entonces se tiene que la distancia mínima es mayor que el peso mínimo. □

En un código general de tamaño M debemos calcular $\binom{M}{2}$ distancias para calcular la distancia mínima. En el caso lineal basta con calcular M pesos.

Imaginemos una permutación de las coordenadas en un espacio vectorial canónico. Dicha permutación se puede entender como una aplicación lineal. A nivel matricial dicha aplicación lineal está dada por una matriz que tiene un 1 en cada fila y en cada columna. Si le aplicamos una permutación de coordenadas a los elementos de un código lineal $\mathcal{C}$ obtenemos otro código $\mathcal{C}'$ que también es lineal. Ello conduce a la

Definición 4.5 *Dos códigos lineales* $\mathcal{C}$ *y* $\mathcal{C}'$ *son* equivalentes *si uno se puede transformar en el otro mediante una permutación de las coordenadas.*

Debido a la estructura de grupo del conjunto de permutaciones, es claro que estamos ante una relación de equivalencia. La razón de que a efectos prácticos podamos utilizar un código o su equivalente se deben al siguiente resultado.

Proposición 4.2 *Si dos códigos son equivalentes entonces tienen los mismos parámetros.*

Demostración.

Es claro que tienen la misma longitud. Como los dos tipos de aplicaciones lineales consideradas son biyectivas, también tienen la misma dimensión. Efectuar permutaciones a los símbolos no varía el peso de las palabras. □

Ejemplos.

- Sea el código lineal ternario

$$\mathcal{C} = \{0000, 1011, 0112, 2022, 1221, 1120, 2210, 1202, 2101\}$$

  Si permutamos entre si las dos primeras coordenadas obtenemos

$$\mathcal{C}' = \{0000, 0111, 1012, 0222, 2121, 1120, 2210, 2102, 1201\}$$

- El código anterior $\mathcal{C}$ tiene como base (cada fila es un vector de la base)

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix}$$

  Si permutamos las dos primera columnas de esta matriz se obtiene una base del código $\mathcal{C}'$. Este es un resultado general: para permutar todos los elementos de un código lineal, basta efectuar la permutación a los elementos de una base. Si escribimos en forma de matriz los elementos de la base, simplemente estamos permutando columnas.

En general, si aplicamos una permutación $\sigma \in S_n$ a un código $\mathcal{C}$ (lo que denotaremos como $\sigma(\mathcal{C})$) se obtiene un código distinto $\mathcal{C}'$. Pero para alguna de las permutaciones es posible que el código no cambie.

Definición 4.6 *Decimos que un código* $\mathcal{C} \in k^n$ *es* invariante *por una permutación* $\sigma \in S_n$ *si* $\sigma(\mathcal{C}) = \mathcal{C}$.

El conjunto de permutaciones que dejan invariante un código es un subgrupo del grupo de permutaciones. Dicho grupo se denota $\mathrm{Aut}(\mathcal{C})$ y sus elementos se dice que son los automorfismos o simetrías del código.
Ejemplos.

- Si $\mathcal{C} = k^n$, cualquier permutación de coordenadas vuelve a dar el mismo código. El grupo de automorfismos coincide con el de permutaciones. El código de repetición también tiene el mismo grupo de simetría.

- Sea $\mathcal{C} = \{0000, 1010, 0101, 1111\}$. Si efectuamos la permutación cíclica

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 2 & 3 & 4 & 1 \end{pmatrix}$$

  el resultado es nuevamente el mismo código. El grupo de simetría de este código contiene al subgrupo de las permutaciones cíclicas. Estudiaremos con más profundidad este tipo de códigos a lo largo de estas notas.

5. Códigos de Hamming

Antes de realizar el estudio general de los códigos lineales, vamos a estudiar un tipo particular de ellos, pues muchas de las ideas que empleamos en esta sección nos serán útiles con posterioridad. Empezaremos con un ejemplo concreto, seguiremos con los códigos binarios y terminaremos con los códigos sobre cualquier cuerpo.

Consideremos los números binarios de tres cifras no nulos. Con ellos formamos la siguiente matriz.

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

Denotaremos por h_i al vector columna que ocupa la posición i-ésima. Hemos construido la matriz de tal forma que h_i es precisamente el número i escrito en binario. También debemos observar que en nuestra matriz están todos los vectores (en forma de columna) no nulos de k^3.

Esta matriz actua sobre vectores de $2^3 - 1 = 7$ componentes. Consideremos el subconjunto

$$Ham(3) = \{x \in k^7 \text{ tales que } Hx = 0\}$$

Este conjunto es un subespacio vectorial. Como el rango de la matriz es tres (los elementos de la base canónica pertenecen a la matriz), el subespacio vectorial tiene dimensión $(2^3 - 1) - 3 = 4$. Este código, llamado de Hamming, es de tipo $[7, 4]$.

Veremos ahora que todos los elementos de este subespacio tienen peso mayor o igual que 3, demostrando que no existen elementos de peso 1 y 2. Si x tiene peso 1, entonces es un vector de la base canónica, que denotaremos por e_i. Al actuar H sobre un vector e_i se obtiene h_i, que por supuesto es no nulo. Si $\omega(x) = 1$ entonces $Hx \neq 0$ y no pertenece al subespacio. Si $\omega(x) = 2$, al actuar H sobre él se obtiene la suma de dos columnas de la matriz. Como estas son linealmente independientes, necesariamente $Hx \neq 0$. Sin embargo si

existen vectores de peso 3. Veamos la razón. Al considerar todos los números binarios no nulos, estamos considerando también todos los vectores no nulos de k^3. Si tomamos dos de estos vectores, que son linealmente independientes, su suma necesariamente es otra columna de la matriz, pues cualquier vector no nulo está en dicha matriz. Por ejemplo, al sumar la columna primera y segunda da lugar a la tercera. Ello implica que el vector que tiene exactamente tres unos en las tres primeras posiciones si pertenece al subespacio. Existen vectores de peso tres. Hemos demostrado que $Ham(3)$ es de tipo $[7, 4, 3]$.

Generalizemos este resultado. Dado un número natural r, consideramos todos los números binarios no nulos de r cifras. En total son $n = 2^r - 1$. Los colocamos en forma de matriz siguiendo el orden creciente. El número total de filas de la matriz H formada a partir de ellos es r. Esta matriz es de rango r, pues la base canónica está incluida dentro de la matriz. Debido a ello el conjunto

$$Ham(r) = \{x \in k^n \text{ tales que } Hx = 0\}$$

es un subespacio vectorial de dimensión $n-r$. El mismo argumento que en el caso de $r = 3$ prueba que el peso mínimo es 3. Por lo tanto hemos construido un código lineal de tipo $[2^r - 1, 2^r - 1 - r, 3]$.

Ejemplos.

- La matriz asociada a $Ham(2)$ es

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

  Este código tiene únicamente dos elementos: $\{000, 111\}$.

- El código $Ham(3)$ tiene 16 elementos. Los siguientes elementos forman una base de dicho subespacio vectorial (cada fila de la matriz es un

vector de la base)

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Proposición 5.1 *$Ham(r)$ es un código perfecto.*

Demostración.

Como $d = 3$, sabemos que $t = 1$. Tenemos también que $n = 2^r - 1$, $M = 2^{n-r}$. Operando

$$2^{n-r}(1+n) = 2^{n-r}(1+2^r-1) = 2^{n-r+r} = 2^n$$

y se cumple la igualdad. □

Como los códigos de Hamming tiene $d = 3$ permiten detectar dos errores. Si $x \rightsquigarrow y$ tiene dos errores entonces $y = x + e_i + e_j$. Si hacemos actuar la matriz H sobre el vector y tenemos

$$Hy = H(x + e_i + e_j) = 0 + h_i + h_j \neq 0$$

y sabemos que ha habido algún error.

Si se produce un solo error entonces $y = x + e_i$. Haciendo la misma operación $Hy = h_i$. El resultado de esta operación nos informa (en binario) de la posición donde ha tenido lugar el error, y podemos corregirlo.

Si a cada código de Hamming se le añade un dígito de paridad aumentamos en una unidad la longitud del código. Veamos que también aumenta en una unidad la distancia mínima. Al añadir un dígito de paridad, todos los vectores tienen peso par, por lo que la distancia entre ellos también es par. Al añadir un dígito la distancia mínima puede o bien quedarse igual o bien aumentar en una unidad, por lo que necesariamente debe ser cuatro. Este nuevo código, llamado código de Hamming extendido, se denota por $EHam(r)$ y permite detectar 3 errores.

Pasemos ya al caso de un cuerpo cualquiera. Hemos visto que la existencia de palabras de peso unidad está asociada a columnas nulas en la matriz H. La existencia de palabras de peso 2 está asociada a la existencia de dos columnas linealmente dependientes. Si queremos que la distancia mínima sea 3 debemos evitar estas situaciones. Ahora ya no podemos poner todos los vectores no nulos como en el caso binario, pues entonces tendríamos columnas linealmente dependientes. Entre todos los vectores que son proporcionales (y que forman una recta) debemos elegir uno para situarlo en la matriz. De esta forma nunca tendríamos dos columnas proporcionales. Pero si ponemos un vector por cada recta, necesariamente existen conjuntos de tres columnas linealmente dependientes, lo que implica la existencia de vectores de peso 3. Por lo tanto formamos la matriz H colocando un vector por cada recta de k^r. En el caso binario la colocación de estos vectores era natural y se realizaba en orden creciente, siempre que pensaramos que cada vector es un número en base binaria. Aquí tenemos el problema de elegir un punto de cada recta y despues colocarlos en la matriz. Seguiremos pensando que un vector es un número natural escrito en base q y emplearemos el siguiente procedimiento:

- En cada recta elegimos el vector que, como número en base q, sea el más pequeño. Esto es equivalente a que su primera cifra significativa sea un 1.
- Colocamos estos vectores en orden creciente.

Denotaremos por $Ham(r, q)$ al código lineal que acabamos de construir. En vista de nuestra construcción necesariamente $d = 3$. Calculemos ahora los otros parámetros. Cada vector no nulo de k^r pertenece a una recta y cada recta tiene $q-1$ elementos no nulos. Los elementos no nulos de dos rectas son claramente disjuntos (dos rectas vectoriales se cortan en el cero). Vectores no nulos existen $q^r - 1$. Luego la matriz tendrá r filas y $(q^r - 1)/(q - 1)$ columnas. De nuevo el rango de esta matriz es el máximo posible, r. Por lo tanto tenemos

$$n = \frac{q^r - 1}{q - 1} \qquad m = n - r$$

Ejemplos.

- La matriz del código ternario $Ham(2,3)$ es

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

- La matriz asociada al código ternario $Ham(3,3)$ es

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

  Su longitud es

$$n = \frac{3^3 - 1}{3 - 1} = \frac{26}{2} = 13$$

  y su dimensión es $m = 13 - 3 = 10$.

- En la matriz H hemos colocado un vector columna por cada elemento del espacio proyectivo asociado a k^r. Si colocamos las columnas en un orden distinto obtenemos un código equivalente. Todos los códigos equivalentes también se denominan códigos de Hamming.

Un cálculo rutinario prueba la siguiente

Proposición 5.2 *El código de Hamming $Ham(r,q)$ es perfecto.*

La decodificación en el caso no binario es similar. Supongamos que en la transmisión $x \rightsquigarrow y$ ha habido un error. Entonces $y = x + \lambda e_i$. Si hacemos actuar H sobre y obtenemos $Hy = \lambda h_i$ que es un múltiplo de la columna i-ésima. El número por el que se multiplica a dicha columna nos informa del error cometido y la columna del lugar donde ha tenido lugar el error.

Imaginemos que una de las matrices que hemos construido "borramos" una columna y hacemos un análisis similar. El mismo razonamiento prueba que no existen vectores de peso 1 y 2. Lo que ya no está tan claro es que existan vectores de peso 3. Si colocamos menos vectores columna la distancia mínima permanece igual a tres o incluso puede aumentar. Decimos que

estos nuevos códigos se obtienen recortando una coordenada. Naturalmente podemos recortar más de una coordenada.

Ejemplo.

- Si recortamos la última coordenada al código de Hamming $[7, 4, 3]$ se obtiene la matriz
$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$
Este código es de tipo $[6, 3, 3]$.

- Sea $q = 7$. El código $Ham(2, 7)$ tiene por matriz
$$H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$
Si recortamos las dos primeras coordenadas, se obtiene la nueva matriz
$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

De todos los ejemplos y construcciones particulares que hemos trabajado, se pueden extraer algunas ideas que son fundamentales en el estudio general de los códigos lineales.

- El código lineal queda perfectamente definido conociendo una matriz H de rango máximo. La longitud del código es el número de columnas de esta matriz y la codimensión del código es el número de filas.

- La distancia mínima parece tener que ver con la dependencia o independencia lineal de las columnas.

- En una transmisión $x \rightsquigarrow y$, si $Hy \neq 0$ sabemos que se han producido errores.

- Parece que el vector Hy tiene relación con el lugar donde se ha producido el error y con el error cometido.

6. Matrices generadoras

Una de las formas más sencillas de especificar un subespacio vectorial es dar una base. Con ánimo de operar con los vectores, colocaremos los vectores como filas de una matriz.

Definición 6.1 *Llamamos* matriz generadora *a toda matriz cuyos vectores filas sean linealmente independientes. Decimos que una matriz G genera un código $\mathcal{C}$ si los vectores fila de la matriz forman una base del subespacio vectorial $\mathcal{C}$.*

El número de columnas de G indica la longitud del código y el número de filas nos informa de la dimensión del código.

Si efectuamos la multiplicación matricial xG, siendo x un vector arbitrario del tamaño adecuado, el resultado es una combinación lineal de las filas de la matriz. Todos los elementos del código se pueden escribir de manera única en la forma xG. Entonces, si G es una matriz generadora del código $\mathcal{C}$ de dimensión m tenemos

$$\mathcal{C} = \{xG \text{ con } x \in k^m\}$$

Consideremos una aplicación lineal $\varphi : k^m \to k^n$ cuya imagen es el código $\mathcal{C}$. La matriz M de esta aplicación lineal está formada por las imágenes de los vectores de la base puestos en forma de columna. A nivel matricial, si queremos calcular la imagen de un vector $x \in k^m$ debemos efectuar la multiplicación Mx (el vector x se escribe en forma de columna). Este es el convenio habitual que se sigue en todos los libros de álgebra lineal. Sin embargo en teoría de códigos se sigue el otro convenio: multiplicar el vector por la izquierda ($\varphi(x) = xM$). Si seguimos este convenio los elementos de la base se deben escribir en forma de filas y la matriz G es precisamente la matriz de la aplicación lineal. Todo esto es cuestión de convenio, pero debemos tener cuidado al traducir los resultados.

Ejemplos.

- La matriz generadora del código de repetición es

$$G = (1, 1, \ldots, 1)$$

- Ya hemos visto que una matriz generadora del código $Ham(3)$ es

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

- Si al código trivial k^n le añadimos un dígito de paridad, la matriz generadora es

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

  Para construir la matriz generadora de un código extendido, se añade una nueva columna. Los elementos de dicha columna se obtienen sumando el resto de los elementos de dicha fila y cambiando el signo al resultado.

- A veces también se llama matriz generadora de un código a aquella cuyos vectores fila generan el código, aunque estos vectores no sean linealmente independientes. Si a esta matriz le quitamos los vectores linealmente dependientes, se obtiene una matriz que cumple nuestra definición.

La codificación de una palabra consiste en la realización de la multiplicación matricial xG. Si ahora tenemos una palabra codificada y y queremos saber cual es la palabra sin codificar, debemos resolver el sistema $xG = y$. En este cometido será de suma importancia el método de Gauss.

Ejemplos.

- Sea $q = 11$. Dada G la matriz

$$G = \begin{pmatrix} 1 & 0 & 2 & 3 & 5 \\ 2 & 4 & 2 & 3 & 2 \\ 2 & 8 & 9 & 3 & 4 \end{pmatrix}$$

para codificar el vector $x = (a, b, c)$ realizamos la multiplicación xG y obtenemos

$$(a + 2b + 2c, 4b + 8c, 2a + 2b + 9c, 3a + 3b + 3c, 5a + 2b + 4c)$$

Como vemos, a simple vista, la relación que liga a una palabra con su codificada es compleja. Si nos dan una palabra codificada y queremos obtener la palabra sin codificar nos supondrá cierto esfuerzo.

- Sea otra vez $q = 11$, pero ahora consideramos la matriz

$$G = \begin{pmatrix} 1 & 0 & 0 & 2 & 5 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 3 & 2 \end{pmatrix}$$

Haciendo la misma operación obtenemos

$$xG = (a, b, c, 2a + b + 3c, 5a + 3b + 2c)$$

Ahora la decodificación es trivial.

Como cualquier subespacio vectorial puede tener distintas bases, la matriz generadora G no es única. Por ejemplo, si tomamos una matriz G y permutamos sus filas, la nueva matriz también genera el mismo código. Lo mismo le ocurre si a una fila la multiplicamos por un escalar no nulo o bien sumamos combinaciones lineales de filas. Se demuestra en álgebra lineal que aplicando estas operaciones elementales se pueden trasformar entre si cualquier par de bases de un subespacio.

Proposición 6.1 *Dos matrices G y G' generan el mismo código si y solo se pueden transformar una en la otra aplicando las siguientes operaciones:*

- *Permutar filas.*
- *Multiplicar una fila por un escalar no nulo.*
- *Sumar combinaciones lineales de filas a otra fila.*

Ejemplo.

Tomemos la matriz generadora

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

Calculemos otra matriz generadora del mismo código.

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{matrix} F_3' = F_3 + F_2 \\ F_2' = F_2 + F_1 \end{matrix} \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{matrix} F_1' = F_1 + F_3 \\ F_2' = F_2 + F_3 \end{matrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Si utilizamos la nueva matriz para codificar, la decodificación será más sencilla.

Definición 6.2 *Decimos que un código admite una* codificación sistemática *si posee alguna matriz generadora de la forma* $G = (\mathrm{Id}\,|A)$. *También se dice que* G *está escrita en* forma estandard.

Si $\mathcal{C}$ admite una codificación sistemática, la codificación de cualquier palabra x es de la forma (x, a). Dicho de otro modo: la palabra codificada "tiene como prefijo" a la palabra sin codificar. Desgraciadamente no todo código admite una codificación sistemática. Analicemos la razón. Si la matriz G de tamaño $n \times r$ tiene las r primeras columnas linealmente dependientes, ya podemos realizar con las filas todo tipo de operaciones, que seguirán siendo dependientes. Es imposible escribirla en la forma $(\mathrm{Id}\,|A)$. Pero como el rango es r, necesariamente existen r columnas linealmente independientes. Si las colocamos al principio, si que podemos reducir G a la forma estandard

operando con las filas. Pero esta nueva operación implica la aparición de los códigos equivalentes. El método de Gauss nos permite demostrar la

Proposición 6.2 *Todo código lineal* $\mathcal{C}$ *posee un código equivalente que posee matriz en forma estandard.*

7. Código dual y matriz de paridad

En todo espacio vectorial se pueden introducir formas bilineales. En particular, en los espacios canónicos k^n, se puede introducir una forma bilineal canónica, que es la generalización de la métrica euclídea del espacio $\mathbb{R}^n$. Debemos señalar que solamente algunas de las propiedades de la métrica euclídea se trasladan a nuestro caso.

Definición 7.1 *Llamamos* producto interior *(o* producto escalar*) de x e y a*

$$\langle x, y \rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

Si se cumple que $\langle x, y \rangle = 0$ decimos que x e y son ortogonales.

Dados dos vectores de la base se tiene $\langle e_i, e_j \rangle = \delta_{ij}$ y la base es ortonormal. La matriz del producto escalar respecto a la base canónica es la matriz identidad.

Ejemplos.

- En el caso binario, los vectores 1111 e 0101 son ortogonales. Cualquier vector x ortogonal 0101 debe cumplir $\langle 0101, x \rangle = 0$. Si escribimos esta condición en coordenadas obtenemos

  $$x_2 + x_4 = 0$$

  que son las ecuaciones implícitas de un subespacio. Observamos que 1111 pertenece a este subespacio.

- Si e_i es un vector de la base, entonces $\langle x, e_i \rangle = x_i$. Si un vector x es ortogonal a todos los vectores, en particular debe ser ortogonal a todos los elementos de la base, de donde se deduce que $x_i = 0$ para todo i y x es el vector nulo. Las formas bilineales que cumplen que solamente el cero es ortogonal a todo vector se llaman no degeneradas.

- En el caso euclídeo sabemos que si $\langle x, x \rangle = 0$ entonces necesariamente x es nulo. Pero esta propiedad no es necesariamente cierta en otros

contextos. Llamamos vectores isótropos a los que son ortogonales a si mismos. Por ejemplo, si $q = 2$ y $x = 1010$ tenemos que $\langle x, x \rangle = 0$.

- Toda forma bilineal induce una aplicación lineal, llamada polaridad entre el espacio y dual

$$\begin{array}{rccc} \phi: & E & \to & E^* \\ & x & \to & \phi(x) = \alpha_x \end{array}$$

donde $\alpha_x(y) = \langle x, y \rangle$. Si la forma es no degenerada la aplicación es inyectiva. En el caso de dimensión finita, por razones dimensionales, también es epiyectiva. Como el producto escalar que hemos introducido en k^n es canónico, se tiene que la polaridad es un isomorfismo canónico de k^n y su dual.

Definición 7.2 *Dado un subconjunto $V \subset k^n$ llamamos* ortogonal de V, *y denotamos $V^\perp$, a*

$$V^\perp = \{x \in k^n \text{ tales que } \langle x, y \rangle = 0 \text{ para todo } y \in V\}$$

El conjunto $V^\perp$ es siempre un subespacio (incluso si V no es): si tenemos $x, x' \in V^\perp$ e $y \in V$

$$\langle x + \lambda x', y \rangle = \langle x, y \rangle + \lambda \langle x', y \rangle = 0$$

debido a la estructura bilineal del producto escalar.

Si tomamos un vector fijo y, los vectores ortogonales a y cumplen una ecuación lineal, que se obtiene realizando el producto escalar, $\langle y, x \rangle = 0$, Si queremos que sea ortogonal a un conjunto V, debe ser ortogonal a cada uno de sus elementos. Esto da lugar a un sistema de ecuaciones lineales (una ecuación por cada elemento de V). La solución de ese sistema lineal es precisamente $V^\perp$.

En teoría de códigos, el código $\mathcal{C}^\perp$ se denomina dual y no ortogonal. Una posible razón de esta denominación es

Proposición 7.1 *Si $\mathcal{C}$ es un subespacio entonces $\mathcal{C}^\perp = \phi^{-1}(\mathcal{C}^0)$, donde ϕ denota la polaridad asociada a la métrica.*

Demostración.

Si $x \in \mathcal{C}^\perp$ entonces $\alpha_x(y) = \langle x, y \rangle = 0$ para todo $y \in \mathcal{C}$. Esto implica que $\phi(x) \in \mathcal{C}^0$.

Si $x \in \phi^{-1}(\mathcal{C}^0)$, entonces $\phi(x) = \alpha_x \in \mathcal{C}^0$. Por lo tanto $\alpha_x(y) = 0 = \langle x, y \rangle$ para todo $y \in \mathcal{C}$ y x es ortogonal al código. □

De la fórmula de dimensiones se extrae el siguiente

Corolario 7.2 *Si $\mathcal{C}$ es un código lineal*

$$\dim(\mathcal{C}) + \dim(\mathcal{C}^\perp) = \dim(k^n)$$

En el espacio euclídeo, dado cualquier subespacio vectorial su ortogonal se encuentra siempre en posición de suma directa con él. Sin embargo, en nuestro caso, perfectamente puede ocurrir que $\mathcal{C} \cap \mathcal{C}^\perp \neq 0$.

Ejemplos.

- Dado $q = 2$ y $n = 3$ sea $\mathcal{C}$ el código de repetición. Los vectores del ortogonal a $\mathcal{C}$ son los vectores ortogonales al vector 111. Si escribimos la condición $\langle 111, x \rangle = 0$, nos conduce a la ecuación $x_1 + x_2 + x_3 = 0$. Luego los vectores ortogonales a $\mathcal{C}$ son los de peso par. Este es un resultado general: el ortogonal del código de repetición está formado por los vectores de peso par.

- Sea $\mathcal{C} = \{0000, 1100, 0011, 1111\}$. Se tiene que $\mathcal{C}^\perp = \mathcal{C}$. Los códigos que cumplen $\mathcal{C}^\perp = \mathcal{C}$ se llaman autoduales.

Corolario 7.3 *Si $\mathcal{C}^\perp$ es un código lineal $(\mathcal{C}^\perp)^\perp = \mathcal{C}$.*

Demostración.

La inclusión $(\mathcal{C}^\perp)^\perp \subset \mathcal{C}$ es clara. Como ambos subespacios tienen la misma dimensión, tenemos la igualdad. □

Traduzcamos todos estos resultados a nivel matricial. La clave de dicha traducción se encuentra en la siguiente observación.

Proposición 7.4 *$x \in \mathcal{C}^\perp$ si y solo si x es ortogonal a una base de $\mathcal{C}$.*

Demostración.

Es claro que si $x \in \mathcal{C}^\perp$, entonces es ortogonal a todos los vectores de $\mathcal{C}$ y en particular es ortogonal a los elementos de una base.

Si $\{v_i\}$ es una base de $\mathcal{C}$ y x es ortogonal a todos esos elementos, también será ortogonal a cualquier combinación lineal de ellos (aplicando la linealidad) y por lo tanto x es ortogonal a todos los elementos del subespacio. □

Ejemplos.

- Consideremos el código de Hamming extendido $EHam(3)$ de tipo $[8, 4]$. Una matriz generadora se obtiene añadiendo un dígito de paridad a cada fila de la matriz generadora del código $[7, 4]$.

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

  El código dual tiene dimensión 4. Se comprueba que $\langle\, v_1, v_i \,\rangle = 0$, siendo v_i la fila i-ésima. Como v_1 es ortogonal a los vectores de la base, tenemos que v_1 pertenece al ortogonal. Lo mismo se puede hacer con v_2, v_3 y v_4. El dual está generado por los mismos vectores. Este código es autodual.

- Si hacemos un desarrollo similar con el código ternario de Hamming $Ham(2, 3)$ tenemos que también es autodual.

- Sea $\mathcal{C}$ el código de los vectores de peso par. Si x es ortogonal a él, debe ser ortogonal a una base. Tomando la base más sencilla, se obtiene que todas las coordenadas deben ser iguales a la última. El dual del código de paridad es el código de repetición.

En forma matricial, un vector x es ortogonal a una base (escrita como una matriz generadora G) si y solo si $Gx = 0$ (el vector se debe escribir en forma de columna). Dado un código $\mathcal{C}$, con matriz generadora G, su ortogonal es

$$\mathcal{C}^{\perp} = \{x \in k^n \text{ tales que } Gx = 0\}$$

Hemos demostrado que $\mathcal{C}^{\perp}$ es el núcleo de una matriz generadora. Esto vuelve a demostrar la fórmula de las dimensiones y nos da un método práctico para el cálculo del ortogonal.

Definición 7.3 *Dado un código $\mathcal{C}$ llamamos* matriz de paridad *a una matriz generadora de su código dual. La denotaremos con la letra H.*

Como tenemos la igualdad $(\mathcal{C}^{\perp})^{\perp} = \mathcal{C}$ resulta que $\mathcal{C}$ es núcleo de H

$$\mathcal{C} = \{x \in k^n \text{ tales que } Hx = 0\}$$

que es una nueva forma de decir que los elementos de $\mathcal{C}$ son perpendiculares a una base del ortogonal.

Ejemplos.

- Si escribimos en coordenadas la condición $Hx = 0$, Obtenemos las ecuaciones implícitas del subespacio $\mathcal{C}$.
- La matriz H que hemos utilizado para construir los códigos de Hamming es una matriz de paridad.
- Sea $\mathcal{C}$ un código y $\mathcal{C}'$ el código extendido. Si H es una matriz de paridad del código $\mathcal{C}$, entonces una matriz de paridad del código extendido es

$$H' = \left(\begin{array}{ccccc|c} 1 & 1 & \dots & 1 & 1 & 1 \\ \hline & & & & & 0 \\ & & & H & & \vdots \\ & & & & & 0 \end{array}\right)$$

Si x es un elemento de k^m sabemos que xG es un elemento del código. Si a este elemento le aplicamos H, debe obtenerse el vector nulo. Pero para poder aplicar H primeramente debemos escribir dicho vector en forma de columna, transponiendolo. Como $= (xG)^t = G^t x^t$ tenemos que $H(G^t x^t) = 0$. Como esto es cierto para cualquier x, se obtiene

$$HG^t = 0$$

Recíprocamente, si G es una matriz generadora y H es una matriz del tamaño y rango adecuados que cumple $HG^t = 0$, entonces H es una matriz de paridad del código. En general, dada G encontrar una matriz H que cumpla la condición anterior, nos conduce a un sistema de ecuaciones lineales. Pero si la matriz G está en forma estandard, el cálculo de H es sencillo.

Proposición 7.5 *Si G está escrita en forma estandard (Id_m, A), entonces la matriz de paridad es $H = (-A^t, \mathrm{Id}_{n-m})$*

Demostración.

La matriz H es del tamaño y del rango adecuados. Veamos que cumple la ecuación (omitimos los subíndices, pues son claros por el contexto)

$$HG^t = (-A^t, \mathrm{Id})(\mathrm{Id}, A)^t = (-A^t, \mathrm{Id})(\mathrm{Id}, A^t) = -A^t + A^t = 0$$

utilizando la multiplicación matricial por bloques. □

Definición 7.4 *Una matriz de paridad H es* estandard *si $H = (A, \mathrm{Id})$.*

Ejemplos.

- La matriz de paridad de $Ham(3)$ en forma estandar es

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

  La matriz generadora es

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

- Si $q = 7$ entonces

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 3 & 5 \end{pmatrix} \Rightarrow H = \begin{pmatrix} -1 & -3 & 1 & 0 \\ -1 & -5 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 4 & 1 & 0 \\ 6 & 2 & 0 & 1 \end{pmatrix}$$

8. Aplicaciones de la matriz de paridad

Una de las principales ventajas que tiene la matriz de paridad sobre la matriz generadora es que permite calcular la distancia mínima. Este resultado se basa en la siguiente observación: si denotamos por h_i la columna i-ésima de la matriz de paridad H, entonces

$$Hx = x_1h_1 + x_2h_2 + \cdots + x_nh_n$$

Sea d la distancia mínima. Existe entonces un vector x de peso d. Como $Hx = 0$ esto da lugar a una combinación lineal de d columnas igualadas a 0. Existe un conjunto de d columnas linealmente dependientes. Sin embargo, no existen menos de d columnas dependientes, pues ello implicaría la existencia de una combinación lineal de ellas igualadas a cero. Pero entonces tendríamos un vector de peso menor que d que pertenece al código. Hemos probado la

Proposición 8.1 *Dada una matriz de paridad H, la distancia mínima del código es el menor número de columnas linealmente dependientes.*

Sabemos que en las matrices de paridad de los códigos de Hamming no existe ningún par de columnas dependientes. Luego la distancia es mayor que 2. Sin embargo si existen trios de columnas dependientes, que era precisamente lo que implicaba que existieran vectores de peso 3.

Como la matriz de paridad es de rango $r = n - m$ podemos estar seguros que $r+1 = (n-m)+1$ columnas siempre son dependientes. Luego la distancia tiene que ser menor o igual que dicho número. Esto da origen al

Corolario 8.2 (Cota de Singleton) *Sea $\mathcal{C}$ un código de tipo $[n, m, d]$. Entonces $d \leq n - m + 1$.*

Esta desigualdad se puede escribir también en la forma $m \leq n - d + 1$, que tomando exponenciales da lugar a

$$q^m \leq q^{n-d+1} \Rightarrow |\,\mathcal{C}| \leq q^{n-d+1}$$

que es la cota de Singleton que ya habíamos estudiado (proposición 3.3). Esta nueva demostración solamente es válida para códigos lineales.

Fijados los valores de n y m este corolario da una cota superior a los posibles valores de d. Aquellos códigos para los que esta cota se transforme en igualdad se denominan códigos con distancia de separación máxima. También de denominan códigos MSD, por sus iniciales en inglés.

Proposición 8.3 *Si $\mathcal{C}$ es un código MSD entonces $\mathcal{C}^{\perp}$ también es MSD.*

Demostración.

Sea H la matriz de paridad de $\mathcal{C}$ de dimensión m. Como el código es MSD, cualquier conjunto de $r = n-m$ columnas de H es siempre linealmente independiente.

Como H es la matriz generadora del dual, todo elemento del dual se puede escribir en la forma

$$uH = (\langle h_1, u \rangle, \langle h_2, u \rangle, \ldots, \langle h_n, u \rangle)$$

donde h_i denota la i-ésima columna de H y u es un vector de longitud r. El número máximo de elementos nulos de este vector es $r-1$. Veamos que un número mayor nos lleva a contradicción. Si hay r coordenadas nulas, entonces existen r columnas que cumplen la condición $\langle x, u \rangle = 0$. Las r columnas pertenecen al ortogonal de u, cuya dimensión es $r-1$. Pero entonces dichas columnas no pueden ser linealmente independientes.

El número de coordenadas no nulas (que es lo mismo que el peso) es entonces mayor o igual que

$$n - (r-1) = n - (n-m-1) = m+1$$

La distancia mínima del dual es mayor o igual que $m+1$. Este valor alcanza la cota de Singleton, que es entonces un código MSD. □

La construcción de ejemplos explícitos de códigos MSD con parámetros dados n y m es sencilla. Sin embargo tiene un pequeño problema: el cuerpo

debe ser lo suficientemente grande. La idea es la siguiente. Si tenemos la matriz H de tamaño $r \times n$ y rango r, sabemos que existen r columnas linealmente independientes. La comprobación de esta hecho se puede realizar calculando el determinante de la submatriz formada por dichas columnas y viendo que es no nulo. Pero puede ocurrir que al tomar otra submatriz de tamaño r el determinante sea nulo y la distancia ya no puede ser $r+1$. Debemos construir matrices que todas las submatrices de orden máximo tengan determinante no nulo. Ello se consigue con los determinantes de Vandermonde. Recordemos este concepto.

Definición 8.1 *Dado un cuerpo k y $a_1, a_2, \ldots, a_r$ elementos distintos del cuerpo, llamamos* determinante de Vandermonde *a*

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ a_1 & a_2 & \ldots & a_r \\ a_1^2 & a_2^2 & \ldots & a_r^2 \\ \ldots & \ldots & \ldots & \ldots \\ a_1^{r-1} & a_2^{r-1} & \ldots & a_r^{r-1} \end{vmatrix}$$

Recordemos el resultado fundamental relativo a estos determinantes.

Proposición 8.4 *El determinante de Vandermonde es no nulo.*

Demostración.

Haremos la demostración por inducción. A cada fila le restamos la anterior multiplicada por a_1. Se obtiene el determinante

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ 0 & a_2 - a_1 & \ldots & a_r - a_1 \\ 0 & a_2(a_2 - a_1) & \ldots & a_r(a_r - a_1) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & a_2^{r-2}(a_2 - a_1) & \ldots & a_r^{r-2}(a_r - a_1) \end{vmatrix}$$

Sacamos factor común y desarrollamos por la primera columna

$$(a_2 - a_1)\dots(a_r - a_1)\begin{vmatrix} 1 & 1 & \dots & 1 \\ a_2 & a_3 & \dots & a_r \\ a_2^2 & a_3^2 & \dots & a_r^2 \\ \dots & \dots & \dots & \dots \\ a_2^{r-2} & a_3^{r-2} & \dots & a_r^{r-2} \end{vmatrix}$$

y aparece un nuevo determinante de Vandermonde, pero de tamaño menor. Inductivamente es no nulo. Con un poco más de esfuerzo se demuestra que el valor del deteminante es igual a $\prod_{i<j}(a_i - a_j)$. □

Si queremos construir un código de longitud n debemos buscar n elementos no nulos y distintos entre si. Esto implica que $n < q$. Si queremos que la distancia sea d, tenemos que construir determinantes de Vandermonde de tamaño $d - 1$.

Proposición 8.5 *Sean $a_1, a_2, \dots, a_n$ elementos distintos del cuerpo y sea $d < n$. El código que tiene por matriz de paridad*

$$H = \begin{pmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_n \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^{d-2} & a_2^{d-2} & \dots & a_n^{d-2} \end{pmatrix}$$

es MSD, con parámetros $[n, n - (d-1), d]$.

Demostración.

La matriz es de tamaño $(d-1)\times n$ y de rango máximo. De esto se deducen los dos primeros parámetros. Si tomamos un conjunto cualquiera de $d-1$ columnas, su determinante es de Vandermonde y no nulo. Luego la distancia es mayor que $d-1$. Como el rango de la matriz es $d-1$ entonces d columnas deben ser linealmente dependientes y la distancia es d. □

Ejemplo.

- Sea $q = 7$. El código que tiene por matriz de paridad

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1^2 & 2^2 & 3^2 & 4^2 & 5^2 & 6^2 \end{pmatrix}$$

  tiene $n = 6$, $m = 3$ y $d = 4$. Es un código MSD. Si borramos alguna de las columnas, sigue siendo MSD.

- Añadamos al código anterior una nueva columna

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 4 & 5 & 6 & 0 \\ 1^2 & 2^2 & 3^2 & 4^2 & 5^2 & 6^2 & 1 \end{pmatrix}$$

  Si en los determinantes aparecen únicamente elementos de las primeras 6 columnas, tenemos un determinante de Vandermonde. Si aparece la última, desarrollamos el determinante por dicha columna y obtenemos un determinante de Vandermonde pero de tamaño 2. Nuevamente es no nulo. Este código tiene como parámetros $[7, 4, 4]$.

Para decodificar códigos lineales se emplea la regla de decisión del vecino más cercano. Si n es relativamente grande, el método estandard para construir la regla de decisión es intratable en la práctica. Con ayuda de la matriz de paridad expondremos otro método que da lugar a la misma regla de decisión, pero que es computacionalmente más sencillo.

Definición 8.2 *Sea $\mathcal{C} \in k^n$ un código y H una matriz de paridad. Dado $x \in k^n$, llamamos* síndrome *de x, y denotamos por $s(x)$, al vector*

$$s(x) = Hx$$

El síndrome de un elemento es nulo si y solo si el elemento pertenece al código. Si tenemos una transmisión $x \rightsquigarrow y = x + e$, aplicando la linealidad,

tenemos que $s(y) = s(e)$. Luego el error cometido tiene el mismo síndrome que la palabra recibida.

Proposición 8.6 *Existe una correspondencia biunívoca entre el conjunto de clases de equivalencia de k^n módulo $\mathcal{C}$ y el conjunto de síndromes.*

Demostración.

Sea F una clase de equivalencia. Dos elementos x, y pertenecen a la misma clase, si y solo si $x - y \in \mathcal{C}$. Esto implica que $s(x) = s(y)$. A cada clase de equivalencia se le puede asignar un único síndrome.

Recíprocamente, si $s(y)$ es un síndrome, le podemos hacer corresponder la clase de equivalencia que contiene a y y concluimos que la correspondencia es biunívoca. □

La clase de equivalencia asociada al vector y le denotaremos $\mathcal{C}_y$ (no es una notación estandard). Hemos demostrado

$$\mathcal{C}_y = \mathcal{C}_{y'} \text{ si y solo si } s(y) = s(y')$$

A nivel teórico el síndrome presenta un problema serio. Como un código puede tener varias matrices de paridad, si cambiamos la matriz de paridad, cambiamos también el síndrome. Por ello es aconsejable trabajar a nivel de clases de equivalencia, que están perfectamente definidas y no depende de la matriz de paridad elegida. Sin embargo, a nivel operativo, es más sencillo trabajar con síndromes, habiendo elegido previamente una matriz de paridad.

El error cometido en la transmisión $x \rightsquigarrow y$ debe buscarse únicamente en la clase $\mathcal{C}_y$. Si queremos que el error sea lo menor posible, el peso del error debe ser lo menor posible. Esto nos conduce a la siguiente

Definición 8.3 *Dada una clase lateral $\mathcal{C}_y$, llamamos* líder *de la clase al vector $e \in \mathcal{C}_y$ de peso mínimo, siempre que este vector sea único.*

Es claro que en cada clase debe existir un vector de peso mínimo. Si existen varios vectores de peso mínimo, decimos que esa clase no tiene líder.

Sin embargo el siguiente lema garantiza la existencia de líder en los casos que nos interesa.

Lema 8.7 *Sea $\mathcal{C}$ un código de distancia mínima $d \geq 2t+1$. Si $d(x, \mathcal{C}) \leq t$ (esto significa que existe un $x \in \mathcal{C}$ tal que $d(x, y) \leq t$), entonces la clase $\mathcal{C}_y$ tiene líder.*

Demostración.

Sea x el único elemento del código que cumple $d(x, y) \leq t$. Construimos el vector $e = y - x$ que necesariamente pertenece a $\mathcal{C}_y$. Entonces el peso de e es menor que t. Si e' es otro vector de la clase, de peso menor o igual que t, se cumple

$$\omega(e - e') \leq \omega(e) + \omega(e') \leq 2t < d$$

lo que contradice la definición de distancia mínima, puesto el vector $e - e'$ pertenece al código. □

La regla de decisión del líder para decodificar sigue los siguientes pasos:

1. Dada la transmisión $x \rightsquigarrow y$, calculamos el síndrome de y. Si es nulo, entonces $f(y) = x$.

2. Si el síndrome es no nulo y su clase asociada tiene líder e, entonces $f(y) = y - e$.

3. Si el síndrome es no nulo y la clase asociada no tiene líder, entonces $f(y) =?$.

El lema anterior nos dice que si $d(x, y) \leq t$ la clase asociada tiene líder y por lo tanto esta regla de decisión es capaz de corregir hasta t errores. Es fácil ver que esta regla coinde con la del vecino más cercano, a pesar de que el método de construcción sea distinto.

9. Códigos de Reed-Muller binarios

Aunque nuestro principal interés se centra en los códigos binarios, empezaremos desde una perspectiva un poco más general. La construcción de los códigos de Reed-Muller se basa en la relación existente entre el anillo de funciones de r variables y el anillo de polinomios en las mismas variables. Repasemos algunos resultados que vamos a necesitar. Denotaremos por x_i a las variables del anillo de polinomios y por r el número de variables independientes.

- Un monomio es un producto de variables $x_1^{i_1} x_2^{i_2} \cdots x_r^{i_r}$ con $i_j \geq 0$. El grado de este monomio es $i_1 + i_2 + \cdots + i_r$. El conjunto de todos los monomios es linealmente independiente y todo polinomio es una combinación lineal finita de monomios. El grado de un polinomio es el mayor de los grados de sus monomios constituyentes.

- Si sustituimos cada variable por un escalar y realizamos las operaciones, cada polinomio $f \in k[x_1, \ldots, x_r]$ induce una función

$$\begin{array}{rccc} \hat{f}: & k^r & \to & k \\ & (a_1, \ldots a_r) & \to & f(a_1, \ldots, a_r) \end{array}$$

  La aplicación

$$\begin{array}{rccc} \phi: & k[x_1, \ldots, x_r] & \to & Funciones(k^r, k) \\ & f & \to & \hat{f} \end{array}$$

  es un morfismo de anillos, estando dotado el último anillo de las operaciones punto a punto. Decimos que $\hat{f}$ es la función polinómica asociada a f.

- Si $k = \mathbb{R}$ la aplicación ϕ es inyectiva. Distintos polinomios inducen funciones distintas. En este caso es correcto identificar a los polinomios con un subanillo del anillo de funciones. Es conocida la existencia de funciones no polinómicas (por ejemplo $e^{(x_1 + \cdots + x_r)}$). En este caso la aplicación ϕ no es epiyectiva.

- Si k es finito, el conjunto $Funciones(k^r, k)$ es finito. Como el anillo de polinomios es infinito, la aplicación ϕ no puede ser inyectiva. En este caso no es correcto identificar el polinomio con la función polinómica inducida, puesto que distintos polinomios dan lugar a la misma función.

Estudiaremos con detenimiento el caso de los polinomios en una variable. Las ideas y resultados presentados se generalizan sin dificultad mediante argumentos inductivos.

Si $|k| = q$ sabemos que $a^q = a$ para todo $a \in k$. Ello implica que, como funciones, x^q y x son iguales. El polinomio $x^q - x$ está contenido en el núcleo de ϕ. La implicación inversa también es cierta. Si $\phi(f) = 0$, entonces $f(a) = 0$ para todo elemento del cuerpo. Entonces f es divisible entre $x - a$. Como $x^q - x = \prod(x - a)$, f es múltiplo de $x^q - x$. Hemos demostrado la

Proposición 9.1 *El núcleo de ϕ es el ideal generado por $x^q - x$. El anillo cociente obtenido lo denotaremos por A.*

De todos los polinomios que inducen la misma función existe un único representante de grado mínimo. Como de costumbre trabajaremos siempre con dichos representantes. Dado $f \in k[x]$ denotamos por f^* al polinomio de grado mínimo que induce la misma función. Decimos que f^* es el polinomio reducido de f. Podemos considerar que A está formado por el conjunto de polinomios reducidos. Para construir el polinomio reducido sustituimos x^q por x y sustituciones similares en las potencias mayores que q. De esta forma todo polinomio reducido es de grado menor que q.

Ejemplos.

- Sea $q = 5$ y $f = 4x^4 + 3x^5 + 2x^8$. El polinomio reducido es

$$f^* = 4x^4 + 3x + 2x^4$$

donde hemos sustituido x^5 por x y también x^8 por x^4.

- Si $q = 2$ es muy sencillo reducir cualquier polinomio. Basta con "borrar" todos los exponentes mayores que la unidad.

$$f = 1 + x^2 + x^3 + x^6 \quad \Rightarrow \quad f^* = 1 + x + x + x$$

Corolario 9.2 *El anillo $A = k[x]/\langle x^q - x\rangle$ de los polinomios reducidos se identifica con el conjunto de polinomios de grado menor que q. Cada polinomio reducido induce una función polinómica. Si dos polinomios reducidos inducen la misma función, dichos polinomios son iguales.*

Si k es finito, a toda función de k en k se le puede aplicar la construcción del polinomio interpolador de Lagrange. Toda función en una variable es una función polinómica.

Corolario 9.3 *En una variable, el anillo de polinomios reducidos A es canónicamente isomorfo al anillo de funciones.*

A pesar de que los anillos son isomorfos, trabajaremos con el anillo A debido a que presenta una estructura graduada, inducida por la graduación del anillo de polinomios.

Todos estos resultados se generalizan a varias variables. El anillo de polinomios reducidos es ahora

$$A = k[x_1, \ldots, x_r]/\langle (x_1^q - x_1), \ldots, (x_r^q - x_r)\rangle$$

Para reducir un polinomio de varias variables basta reducir cada una de las variables. Cualquier polinomio reducido tiene grado menor o igual que $r(q-1)$, pero ahora ya no es cierto que cualquier polinomio de grado menor sea un polinomio reducido, pues a pesar de que el grado total del polinomio sea menor, una de las variables puede tener grado mayor que q.

Lema 9.4 *La aplicación que asocia a cada polinomio reducido su función polinómica es inyectiva.*

Demostración.

Realizaremos la demostración por inducción sobre el número de varibles. El caso de una variable ya está demostrado. Sea f^* un polinomio reducido de r variables. Podemos escribir f^* en la forma

$$f^*(x_1, \ldots, x_r) = \\ g_0(x_1, \ldots, x_{r-1}) + g_1(x_1, \ldots, x_{r-1})x_r + \cdots + g_m(x_1, \ldots, x_{r-1})x_r^m$$

donde los polinomios g_i son polinomios reducidos, pero en una variable menos. Si asignamos valores arbitrarios a las primeras $r-1$ coordenadas, obtenemos un polinomio de una sola variable

$$f^*(a_1, \ldots, a_{r-1}, x_r) = \\ g_0(a_1, \ldots, a_{r-1}) + g_1(a_1, \ldots, a_{r-1})x_r + \cdots + g_m(a_1, \ldots, a_{r-1})x_r^m$$

Como este polinomio se anula al asignar a x_r cualquier valor, tenemos que el polinomio es nulo. De esta forma, los coeficientes $g_i(a_1, \ldots, a_{r-1})$ son siempre nulos, para cualquier valor de las variables. Como los g_i son polinomios reducidos en una variable menos, la hipótesis de inducción nos garantiza que deben ser nulos y por lo tanto el polinomio inicial f^* es también nulo. □

Lema 9.5 *La aplicación anterior también es epiyectiva.*

Demostración.

La aplicación es claramente lineal. Dado un vector $v = (a_1, \ldots, a_r) \in k^r$, construimos el polinomio

$$e_v = \prod_{i=1}^{r} [1 - (x_i - a_i)^{q-1}]$$

Por construcción este polinomio es reducido, verifica $e_v(v) = 1$ y se anula sobre el resto de vectores. Si f es una función arbitraria, el polinomio

$$\sum_{v \in k^r} f(v) e_v$$

tiene como función asociada f. $\square$

El anillo de los polinomios reducidos se identifica con el anillo de funciones en el número adecuado de variables. Nuevamente el anillo de los polinomios reducidos tiene la ventaja de presentar una estructura graduada.

Para conocer una función de k^r en k es necesario conocer la imagen de todos los elementos del dominio. Como el dominio es finito, sus elementos se pueden ordenar. Existen muchas ordenaciones posibles de los elementos de k^r, pero nosotros utilizaremos la dada por la escritura de números naturales en base q. De esta forma, el vector i-ésimo, que denotaremos por v_i, es el vector de r componentes, que entendido como número natural escrito en base q, es justamente i. Por ejemplo, en el caso ternario y con $r = 3$ tenemos la ordenación

$$v_0 = 000, v_1 = 001, v_2 = 002, v_3 = 010, \ldots, v_{26} = 222$$

Como k^r tiene $n = q^r$ elementos, fijada una ordenación, cada función f de r variables da lugar a un vector

$$(f(v_0), f(v_1), f(v_2), \ldots, f(v_{n-1}))$$

que está contenido en el espacio vectorial k^n. Es el vector característico asociado a la función f. Recíprocamente, cada vector de k^n es el vector característico de una única función. Si consideramos en k^n la multiplicación componente a componente, es fácil comprobar que esta asociación conserva las operaciones. El anillo de funciones de r variables y k^n (con $n = q^r$) son isomorfos. Estos anillos también son isomorfos al anillo de polinomios reducidos. Debemos notar que estos isomorfismos no son canónicos, sino que dependen de la ordenación elegida. Si variamos la ordenación variamos los isomorfismos.

Ejemplos.

- Situémonos en el caso binario y tres variables independientes. El espacio

k^3 tiene 8 elementos, cuya ordenación, siguiendo nuestro criterio, es

$$000, 001, 010, 011, 100, 101, 110, 111$$

Consideremos el polinomio reducido $f = x_1x_2 + x_3$. Si calculamos las imágenes de estos vectores obtenemos el vector $(0, 1, 0, 1, 0, 1, 1, 0) \in k^8$.

- Con los datos del ejemplo anterior, los vectores característicos de los polinomios x_1 y x_2 son los vectores $(0, 0, 0, 0, 1, 1, 1, 1)$ y $(0, 0, 1, 1, 0, 0, 1, 1)$ respectivamente. Para calcular el vector asociado al producto x_1x_2 basta con multiplicar, coordenada a coordenada, los dos vectores anteriores, obteniendo el vector $(0, 0, 0, 0, 0, 0, 1, 1)$. Conociendo los vectores característicos asociados a las variables, se puede calcular por este procedimiento el vector asociado a cualquier monomio. Esto es consecuencia de que la variables generan (como anillo) el anillo de los polinomios y que la aplicación conserva la multiplicación.

Hemos visto que, fijada una ordenación, se tiene un isomorfismo de A con k^n, siendo $n = q^r$. Cada subespacio de A da lugar entonces a un código, cuya longitud es n. Para construir estos subespacios es para lo que emplearemos la estructura graduada de A. Denotaremos por A_m el conjunto de polinomios reducidos de grado menor o igual que m.

Definición 9.1 *Llamamos código de Reed-Muller de tipo $RM_q(m, r)$ a la imagen del subespacio A_m.*

Si utilizamos otra ordenación distinta, obtenemos un código equivalente, que también se denomina código de Reed-Muller.

La longitud de los códigos de Reed-Muller es q^r. Además, debido nuevamente a la estructura graduada, se tiene la siguiente cadena de inclusiones

$$RM_q(0, r) \subset RM_q(1, r) \subset \cdots \subset RM_q(r(q-1), r)$$

El primer elemento de esta cadena es la imagen de los polinomios de grado cero. Esto es, la imagen de las constantes. Es simplemente el código de repetición. Como todo polinomio reducido es de grado menor que $r(q-1)$, el

último eslabón de la cadena contiene a todos los polinomios reducidos, y por los isomorfismos anteriores es el código trivial k^n. Es por ello que siempre se considera que m es menor o igual que $r(q-1)$. En particular, en el caso binario, siempre se tiene que $m \leq r$.

Ejemplo.

- Fijemos $r = 3$. El único monomio de grado nulo es la unidad. Si pasamos a grado 1, además de este monomio tenemos x_1, x_2, x_3. Para encontrar las matrices generadoras de los códigos de Reed-Muller con $r = 3$ debemos calcular los vectores característicos de todos los monomios reducidos. Ello se encuentra recogido en la siguiente tabla

1	1	1	1	1	1	1	1	1
x_1	0	0	0	0	1	1	1	1
x_2	0	0	1	1	0	0	1	1
x_3	0	1	0	1	0	1	0	1
x_1x_2	0	0	0	0	0	0	1	1
x_1x_3	0	0	0	0	0	1	0	1
x_2x_3	0	0	0	1	0	0	0	1
$x_1x_2x_3$	0	0	0	0	0	0	0	1

  Aquí nos aparecen las matrices generadoras de todos los códigos con $r = 3$, separadas por lineas horizontales.

Pasemos ya estudiar únicamente el caso de los códigos binarios. Ya no escribiremos el 2 en la notación de los códigos.

Proposición 9.6 *La dimensión de $RM(m, r)$ es*

$$\sum_{i=0}^{m} \binom{r}{i}$$

Demostración.

El número de monomios de grado i coincide con las posibles elecciones de i variables de un total de r variables. □

Proposición 9.7 *Si un monomio tiene grado α entonces su vector característico tiene peso $2^{r-\alpha}$, siendo r el número de variables.*

Demostración.

Podemos suponer que el monomio es $x_1 x_2 \dots x_\alpha$. Para que la coordenada asociada a un vector sea no nula, es necesario y suficiente que las primeras α coordenadas del vector sean no nulas, independientemente del resto de las coordenadas. Si fijamos las α primeras coordenadas, existen en total $2^{r-\alpha}$ vectores que cumplen dicha condición. □

Proposición 9.8 *La distancia mínima de $RM(m,r)$ es 2^{r-m}.*

Demostración.

Como en $RM(m,r)$ existe un monomio de grado m, su peso es 2^{r-m}. La distancia mínima tiene que ser mayor o igual que dicho dato. Demostraremos este hecho por inducción.

Sea $f(x_1,\dots,x_r) \in RM(m,r)$. Sacando factor común x_1 en este polinomio, lo podemos escribir como

$$f(x_1,\dots,x_r) = g(x_2,\dots,x_r) + x_1 h(x_1,\dots,x_r)$$

donde $g \in RM(m,r-1)$ y $h \in RM(m-1,r)$. Dividimos el vector característico de f en dos partes iguales. Los vectores asociados a la primera parte tienen $x_1 = 0$ y los de la segunda parte tienen $x_1 = 1$. Por ello el vector característico de f se puede escribir en la forma $(u, u+v)$ siendo u el vector asociado a g y v el asociado a h. El peso total de este vector es siempre mayor que

$$\text{mín}(2 \cdot \omega(u), \omega(v))$$

que por hipótesis de inducción es superior a

$$\text{mín}(2 \cdot 2^{r-m-1}, 2^{r-m}) = 2^{r-m}$$

que es lo que queríamos demostrar. □

Observamos, que salvo el monomio de grado máximo, el resto de monomios tiene como peso una potencia de 2. En particular todos esos pesos son pares. Como existe un único monomio de grado máximo, el código $RM(r-1,r)$ tiene codimensión 1 y todos sus vectores tienen peso par. Obtenemos

Corolario 9.9 *$RM(r-1,r)$ es el código de los vectores de peso par.*

Resulta que el dual de un código de Reed-Muller es nuevamente un código del mismo tipo.

Proposición 9.10 *El dual de $RM(m,r)$ es $RM(r-m-1,r)$.*

Demostración.

Si calculamos la dimensión de ambos códigos, obtenemos el mismo resultado. Por ello basta comprobar que todo vector del primer código es perpendicular a todo vector del segundo. Sea f del primer código y g un elemento del segundo. Su producto fg es un polinomio cuyo grado es menor que r. Si es preciso se reduce, pero sigue cumpliendo que su grado es menor que r. Pero hemos visto en el corolario anterior que su peso es par. Pero entonces su producto escalar es nulo. □

Terminar

10. Códigos cíclicos I

Dado un cuerpo k, sea $k[x]$ su anillo de polinomios en una variable. Denotaremos por R_n al anillo cociente

$$R_n = k[x]/(x^n - 1)$$

Como conjunto cociente que es, todo elemento de R_n tiene varios representantes, pero utilizando el algoritmo de la división observamos que hay un único representante cuyo grado es menor que n. En efecto, dado un elemento $p(x)$ de grado arbitrario, efectuamos la división euclídea

$$p(x) = c(x)(x^n - 1) + r(x) \qquad \Rightarrow \qquad p(x) - r(x) = c(x)(x^n - 1)$$

y $r(x) \equiv p(x)$ es un polinomio de grado menor que n. Además el $r(x)$ encontrado es único. Nosotros siempre trabajaremos con estos representantes y entenderemos que un elemento de R_n es un polinomio de grado menor que n. Para realizar la multiplicación en este anillo, se multiplican los dos polinomios que representan a la clase. Si el resultado es de grado menor que n, esa es la multiplicación. Si al hacer el producto el grado es mayor que n, efectuamos la división euclídea de este producto entre el polinomio $x^n - 1$ y nos quedamos con el resto. Para realizar esta división basta sustituir x^n por 1, lo que implica también la sustitución de x^{n+1} por x y así sucesivamente. De esta forma obtenemos siempre un polinomio de grado menor que n.

Además de la estructura de anillo, R_n posee estructura de espacio vectorial sobre k. Si escribimos un elemento de R_n (siempre ordenado de menor a mayor grado)

$$a(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1}$$

lo podemos entender como el vector $a = (a_0, a_1, a_2, \ldots, a_{n-1}) \in k^n$. Esta correspondencia es lineal y biyectiva. Tenemos que R_n es canónicamente isomorfo, como espacio vectorial, a k^n. Mediante este isomorfismo conseguimos introducir una estructura de anillo en k^n. Como hemos dotado a k^n de una

estructura de anillo, los códigos adaptados a esta estructura son los ideales.

Definición 10.1 *Un código cíclico es un ideal de R_n.*

Para realizar el estudio de los códigos cíclicos recordaremos algunos resultados de la teoría de anillos (conmutativos y con unidad), en especial aquellos relacionados con el anillo de polinomios.

- Dado un anillo A y un elemento $a \in A$, el conjunto de sus múltiplos

  $$\langle a \rangle = \{ab \text{ con } b \in A\}$$

  es un ideal del anillo. Decimos que $\langle a \rangle$ es el ideal principal generado por a.

- Dado un ideal $I \subset k[x]$, sea $g(x) \in I$ de grado mínimo. Si $p(x)$ es otro polinomio del ideal, aplicando la división euclídea

  $$p(x) = c(x)g(x) + r(x)$$

  Como $r(x)$ es combinación de elementos del ideal, él mismo pertenece al ideal. Como es grado menor que $g(x)$, necesariamente es nulo. Resulta que $p(x)$ es un múltiplo de $g(x)$. El ideal I es principal. Si $g'(x)$ es otro generador, entonces $g(x) = c(x) \cdot g'(x)$. Analizando los grados, $c(x)$ es un polinomio de grado nulo, esto es, una constante. Entre todos los generadores del ideal existe un único polinomio mónico (cuyo coeficiente de mayor grado es uno). El anillo de polinomios es un dominio de ideales principales.

- La suma y la intersección de ideales es nuevamente un ideal. En un dominio de ideales principales, el generador de la intersección es el mínimo común múltiplo de los generadores. El generador de la suma es el máximo común divisor. Estos resultados se derivan de

  $$\langle g_1 \rangle \subset \langle g_2 \rangle \quad \Leftrightarrow \quad g_2 \text{ divide a } g_1$$

- Si a genera un ideal y u es una unidad del anillo, entonces au genera el mismo ideal. En general todos los ideales principales poseen varios generadores, pero muchas veces, imponiendo alguna condición más, logramos que dicho generador sea único. Por ejemplo en el anillo de polinomios, entre todos los generadores de un ideal, existe un único polinomio mónico.

- En el anillo de polinomios el concepto de ideal primo coincide con el de ideal maximal. Además, un ideal es primo si y solo si su polinomio generador es irreducible. Si hacemos cociente por un ideal obtenemos un cuerpo si el ideal es maximal. Como $x^n - 1$ tiene como solución el elemento neutro, se obtiene la factorización

$$x^n - 1 = (x - 1) \cdot p(x)$$

  Entonces R_n no puede ser nunca un cuerpo.

- Dado un cuerpo primo, sea $p(x)$ un polinomio irreducible de grado n (la demostración de la existencia de este polinomio no es trivial). El anillo cociente es entonces un cuerpo. Todos los cuerpos finitos admiten esta construcción.

Veamos la relación de esta definición con el concepto habitual de código cíclico. Demos una nueva

Definición 10.2 *Un código $\mathcal{C} \subset k^n$ es* cíclico *si:*

- $\mathcal{C}$ *es un código lineal.*

- *Para todo elemento $x = x_1x_2 \dots x_n$ del código, su permutación cíclica, $\sigma(x) = x_nx_1x_2 \dots x_{n-1}$ es también un elemento del código.*

En la definición hemos rotado todos los elementos una unidad hacia la derecha. Por aplicación repetida de esta propiedad, vemos que el código es invariante por la rotación de r elementos a la derecha. Como una rotación de $n - 1$ elementos hacia la derecha es igual que una rotación de un elemento a

la izquierda, el código también es invariante por rotaciones a la izquierda. En definitiva, un código cíclico es invariante por cualquier tipo de rotación de sus elementos. El grupo de los automorfismos del código contiene al subrupo generado por σ, que es un grupo cíclico de orden n.

Dada cualquier permutación τ, en especial una rotación, se induce una aplicación lineal y biyectiva del espacio. Para comprobar que $\tau(\mathcal{C}) \subset \mathcal{C}$, basta tomar una base $\{v_1, v_2, \ldots, v_m\}$ y comprobar que $\tau(v_i) \in \mathcal{C}$. En nuestro lenguaje, basta con tomar una matriz generadora y comprobar que la rotación de cualquiera de sus filas es una combinación lineal de esas mismas filas. Esto nos da un procedimiento para obtener códigos cíclicos.

Proposición 10.1 *Sea $v \in k^n$ un vector no nulo. Consideramos los n vectores $\sigma^i(v)$ (con $i = 0, 1, \ldots, n-1$) obtenidos de él por rotaciones. El subespacio generado por estos vectores es un código cíclico.*

Demostración.

Como los vectores $\sigma^i(v)$ generan el espacio, un subconjunto de ellos forman una base. Es claro que toda rotación de cualquier elemento de esta base es nuevamente un elemento de la forma $\sigma^i(v)$, que está contenido en el subespacio. □

El subespacio así construido es el menor código cíclico que contiene a v. Por razones obvias, decimos que v es un generador del código. Ello no implica que no existan otros vectores que también generen el mismo código.

Cuando se estudian los códigos cíclicos, es costumbre denotar a las palabras con una letra distinta de la x, para no confundirlas con la variable independiente. Además, la numeración de los subíndices empieza en 0.

Dado un elemento $a = a_0 a_1 \ldots a_{n-1}$ del código, al multiplicarlo por la variable x, obtenemos de nuevo un elemento del código, por ser este un ideal. Realicemos la multiplicación en notación polinomial

$$x \cdot (a_0 + a_1 x + a_2 x^2 + \ldots a_{n-1} x^{n-1}) = a_0 x + a_1 x^2 + a_2 x^3 + \cdots + a_{n-1} x^n$$

Pero como $x^n = 1$, resulta que este vector es $a_{n-1} a_0 a_1 \ldots a_{n-2}$. La multi-

plicación por x se traduce en una permutación cíclica de los elementos del código. Si $\mathcal{C}$ es un ideal, entonces por supuesto es un subespacio vectorial y es invariante por permutaciones cíclicas.

Recíprocamente, sea $\mathcal{C}$ un subespacio invariante por permutaciones cíclicas. Esto implica que $x \cdot a$ pertenece al código si $a \in \mathcal{C}$. Por el mismo motivo $x \cdot (x \cdot a) = x^2 \cdot a$ está en el código y lo mismo con el resto de las potencias. Pero si $p(x)$ es un polinomio arbitrario, tenemos que $p(x) \cdot a$ es una combinación lineal de los elementos anteriores y por lo tanto pertenece al código (que, recordémoslo, es un subespacio vectorial). Hemos demostrado que ambas definiciones de código cíclico son equivalentes.

Ejemplos.

- El cero y el total son siempre ideales. El código de repetición también es cíclico. El código formado por todos los vectores tales que la suma de sus elementos es cero (el dual del de repetición), también es cíclico. A nivel de códigos cíclicos, todos estos son triviales.

- El código con matriz generatriz

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

  es cíclico. Observemos que la segunda y tercera fila se obtienen por permutaciones cíclicas de la primera. Las siguientes rotaciones del primer vector dependen linealmente de las anteriores.

- El código equivalente a $Ham(3)$ que tiene por matriz de paridad

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

  es cíclico (calcular los 16 elementos y comprobarlo). Posteriormente demostraremos que todo código de Hamming binario es equivalente a

uno cíclico.

- La condición de cíclico no se conserva por equivalencia. Un código puede ser cíclico y otro equivalente a él no. El código $Ham(3)$ con su matriz de paridad ordenada de menor a mayor no es cíclico.

- Dado un código arbitrario $\mathcal{C}$ y una permutación arbitraria τ sabemos que $\tau(\mathcal{C}^{\perp}) = (\tau(\mathcal{C}))^{\perp}$. Si $\mathcal{C}$ es invariante por permutaciones cíclicas, entonces su dual también. Si un código es cíclico, su dual también.

Veamos ahora que R_n es un dominio de ideales principales y que todo ideal tiene un generador particular.

Proposición 10.2 *Sea $\mathcal{C} \subset R_n$ un código cíclico. Entonces $\mathcal{C}$ es principal y existe un único generador $g(x)$ que cumple:*

- *Es de grado mínimo.*

- *Es mónico.*

- *Es un divisor de $(x^n - 1)$ (es decir $x^n - 1 = g(x) \cdot h(x)$).*

Demostración.

Sea $g(x)$ un polinomio de grado mínimo del ideal. Multiplicándolo por una constante podemos suponer que es mónico. Si otro polinomio $p(x) \in \mathcal{C}$, dividiendo

$$p(x) = c(x) \cdot g(x) + r(x)$$

y el resto es nulo. El polinomio $p(x)$ es un múltiplo de $g(x)$, que genera el ideal. Si existiesen dos polinomios mónicos de grado mínimo, su resta, que es de grado menor, pertenecería al ideal, en contradicción con la elección de $g(x)$.

Dividimos también $x^n - 1$ entre el generador

$$x^n - 1 = c(x) \cdot g(x) + r(x) \quad \Rightarrow \quad 0 \equiv c(x) \cdot g(x) + r(x)$$

y nuevamente el resto es nulo. El generador es un divisor de $x^n - 1$. $\square$

A este polinomio, que es único, lo llamaremos polinomio generador del código. Ello no implica que este ideal no pueda tener otros generadores de distinto grado, debido a que el anillo tiene unidades no triviales, como muestra el siguiente

Lema 10.3 *Sea $u(x)$ un polinomio (de grado menor que n) primo con x^n-1. Entonces $u(x)$ es una unidad.*

Demostración.

Como $u(x)$ es primo con $x^n - 1$, el lema de Bezout implica la existencia de polinomios que cumplen

$$1 = a(x) \cdot u(x) + b(x) \cdot (x^n - 1)$$

Tomando clases módulo $x^n - 1$, se obtiene que

$$1 = a(x) \cdot u(x)$$

y $u(x)$ es una unidad. □

Aunque posiblemente el generador más interesante del código cíclico es el polinomio mónico de menor grado, existen otros generadores, con otras características, que también son importantes. En una sección posterior analizaremos uno especialmente importante: el generador idempotente.

Volviendo al generador mónico de menor grado, en el anillo cociente, la tercera propiedad enunciada se traduce en $g \cdot h = 0$ y tanto g como h son divisores de cero. Obsérvese que si sumamos los grados de g y de h se obtiene n y que además no puede existir ningún polinomio h_1 de grado menor que h que cumpla $g \cdot h_1 = 0$.

El razonamiento a la inversa de la proposición anterior también es correcto. Si partimos de un polinomio mónico $g(x)$ que divida a $x^n - 1$ podemos construir el ideal $\mathcal{C}$ generado por $g(x)$. Resulta que $g(x)$ es el polinomio generador del código. No pueden existir elementos de menor grado, pues si

$p(x) \in \langle g(x) \rangle$ fuese de grado menor, entonces $p(x) = c(x) \cdot g(x)$ y multiplicando por $h(x)$ (siendo $h(x)$ el polinomio que cumple $g(x) \cdot h(x) = x^n - 1$) obtenemos

$$p(x) \cdot h(x) = c(x) \cdot g(x) \cdot h(x) = c(x) \cdot 0 = 0$$

pero esto es imposible pues el grado de $p(x) \cdot h(x)$ es menor que n.

Corolario 10.4 *Existe una correspondencia biunívoca entre códigos cíclicos contenidos en R_n y polinomios mónicos que dividen a $x^n - 1$.*

En la siguiente sección analizaremos el problema general de la descomposición del polinomio $x^n - 1$ sobre cualquier cuerpo finito. De momento nos contentamos con los siguientes

Ejemplos.

- Dado $q = 2$, se tiene la factorización

  $$x^7 - 1 = (x + 1)(x^3 + x^2 + 1)(x^3 + x + 1)$$

  Los factores cúbicos son irreducibles, puesto que no tienen raices. Hemos encontrado la descomposición prima de $x^7 - 1$. Todo divisor se obtiene por producto de las componentes primas. Contando a la unidad y a él mismo, este polinomio tiene $2^3 = 8$ divisores.

- Siempre se tiene la descomposición

  $$x^n - 1 = (x - 1) \cdot p(x)$$

  y el polinomio tiene siempre, al menos, cuatro divisores. Estos dan lugar a los códigos cíclicos triviales. La existencia de otros códigos cíclicos depende de las descomposiciones no triviales del polinomio $x^n - 1$.

- Si $q = 2$ tenemos la descomposición

  $$x^{23} - 1 = (x - 1)(x^{11} + x^{10} + x^6 + x^5 + x^4 + x^2 + 1)$$
  $$(x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1) = (x - 1) \cdot g_1(x) \cdot g_2(x)$$

El código generado por $g_1(x)$ es equivalente al código de Golay $\mathcal{G}_{23}$. El código ternario de Golay se construye del mismo modo utilizando la descomposición

$$x^{11} - 1 = (x-1)(x^5 + x^4 - x^3 + x^2 - 1)(x^5 - x^3 + x^2 - x - 1)$$

Para una demostración véase *Hill*.

- Sea $x^n - 1 = p_1 \cdot p_2 \cdots p_r$ la descomposición en factores irreducibles sobre el cuerpo k. Cada polinomio p_i genera un ideal $\mathcal{C}_i$, que resulta ser un ideal maximal, debido al caracter irreducible de su generador. Como el mínimo común múltiplo de p_i y p_j es el producto de ambos, el código asociado al divisor $p_i \cdot p_j$ es la intersección $\mathcal{C}_i \cap \mathcal{C}_j$. En general, cualquier código cíclico de longitud n sobre el cuerpo k se obtiene por intersección de la familia de códigos maximales $\{\mathcal{C}_i\}$.

  De la misma forma a cada polinomio

  $$\hat{p}_i = \frac{x^n - 1}{p_i} = p_1 \cdots p_{i-1} \cdot p_{i+1} \cdots p_r$$

  se le asocia un código $\hat{\mathcal{C}}_i$, que no contiene a ningún ideal. Es lo que se llama un ideal minimal. El estudio de estos ideales se pospone hasta una sección posterior.

Si el generador tiene grado r, hemos visto que todos los polinomios del ideal tienen grado igual o superior. Entonces el código tiene, como mucho, dimensión $n - r$. Veamos que efectivamente esa es su dimensión.

Corolario 10.5 *Dado el polinomio generador $g(x)$ de grado r, todo elemento del código se puede escribir, de modo único, en la forma $p(x) \cdot g(x)$, donde $p(x)$ es un polinomio de grado menor que $n - r$.*

Demostración.

El argumento ya repetido de la división euclídea prueba la existencia y unicidad. □

Este resultado nos permite construir la función de codificación

$$\begin{array}{rccc} \phi: & R_{n-r} & \to & R_n \\ & p(x) & \to & p(x)\cdot g(x) \end{array}$$

que es lineal e inyectiva. La matriz de esta aplicación es una matriz generadora del código.

Corolario 10.6 *Si $g(x) = g_0 + g_1 x + \cdots + g_r x^r$ es el polinomio generador, una matriz generadora es de la forma*

$$G_c = \begin{pmatrix} g_0 & g_1 & g_2 & \dots & g_r & 0 & 0 & \dots & 0 \\ 0 & g_0 & g_1 & \dots & g_{r-1} & g_r & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & g_0 & g_1 & g_2 & \dots & \dots & g_r \end{pmatrix}$$

Dicho de otra forma, los elementos $\{g(x), x\cdot g(x), x^2\cdot g(x), \dots, x^{n-r-1}\cdot g(x)\}$ forman una base del código.

Esta matriz es la matriz cíclica asociada al generador. Se construye situando el generador en la primera fila y permutando cíclicamente este vector para obtener el resto de las filas.

Ejemplos.

- Recordemos la descomposición de $x^7 - 1$. El código generado por el divisor $x^3 + x + 1$ tiene por matriz generadora

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

- En cualquier cuerpo la descomposición trivial de $x^n - 1$ es

$$x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \cdots + x + 1)$$

La matriz generadora asociada al último divisor es

$$G = \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \end{pmatrix}$$

y da lugar el código de repetición.

La matriz obtenida por este método no es estandard. Pero sabemos que $g(x) \cdot h(x) = x^n - 1$ y sustituyendo x por 0, obtenemos que $g_0 \neq 0$. Con este resultado ya es fácil obtener la forma estandard de la matriz, empleando para ello solamente operaciones con filas.

Corolario 10.7 *Todo código cíclico admite una matriz generadora en forma estandard.*

En los códigos cíclicos se puede escribir rápidamente una matriz de paridad, aunque la matriz generadora no esté en forma estandard. Pero para realizar esta construcción debemos introducir un concepto nuevo.

Definición 10.3 *Dado el generador $g(x)$ del código, existe un único polinomio $h(x)$ (mónico también) que cumple $g(x) \cdot h(x) = x^n - 1$. El polinomio $h(x)$ se llama* polinomio anulador *(*check polynomial *en inglés) del código.*

En el espacio cociente, la definición anterior se transforma en $g \cdot h = 0$, de ahí el nombre de anulador. Entre todos los polinomios que cumplen la relación anterior, el anulador es el de menor grado. También es costumbre definir directamente el anulador como

$$h(x) = \frac{x^n - 1}{g(x)}$$

Como $h(x)$ es mónico y también divide $x^n - 1$, es el generador de su propio código cíclico.

Proposición 10.8 *Dado un polinomio generador g, sea h su anulador. Entonces*

$$a \in \langle g \rangle \quad \Leftrightarrow \quad a \cdot h = 0$$

Demostración.

Si $a \in \langle g \rangle$, entonces $a = b \cdot g$. Multiplicando

$$a \cdot h = b \cdot g \cdot h = b \cdot (g \cdot h) = 0$$

Si $a \cdot h = 0$, dividimos

$$a = c \cdot g + r \text{ con el grado de } r \text{ menor que el de } g$$

Multiplicamos por h

$$0 = a \cdot h = c \cdot g \cdot h + r \cdot h \Rightarrow r \cdot h = 0$$

Pero por cuestión de grados (el grado de $r \cdot h$ es menor que n), necesariamente $r = 0$ y c es múltiplo del generador. $\square$

Proposición 10.9 *Sea g un generador y $h = h_0 + h_1 x + \cdots + h_m x^m$ su polinomio anulador. Entonces*

$$H = \begin{pmatrix} h_m & h_{m-1} & h_{m-2} & \dots & h_0 & 0 & 0 & \dots & 0 \\ 0 & h_m & h_{m-1} & \dots & h_1 & h_0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & h_m & h_{m-1} & h_{m-2} & \dots & \dots & h_0 \end{pmatrix}$$

es una matriz de paridad (notar que el orden de los coeficientes es inverso).

Demostración.

Como la suma de los grados de h y de g es n, la matriz H tiene el tamaño adecuado para una matriz de paridad. Además como $h_m = 1$, el rango es máximo. Para comprobar que es una matriz de paridad, basta ver que es una matriz generadora del código dual. Pero, debido a que presenta el tamaño y rango adecuados, nos basta con comprobar que cada fila es ortogonal a todo elemento de $\mathcal{C}$.

Si $a \in \mathcal{C}$ entonces $a \cdot h = 0$. En particular debe ser nulo el coeficiente de x^m. Dicho coeficiente es

$$a_0 h_m + a_1 h_{m-1} + \cdots + a_m h_0$$

que es precisamente el producto escalar del vector a por la primera fila de la matriz. Si hacemos el mismo cálculo con la siguiente potencia, obtenemos el producto escalar con la segunda fila y así sucesivamente, $\square$

Esta matriz tampoco está en forma estandard, pero como h_0 no es nulo, podemos transformarla rápidamente en estandard. Como esta matriz es la generadora del código dual (que también es cíclico), un generador del dual es el polinomio $h(x)$ pero escrito "al revés". Debemos observar que este nuevo polinomio puede no ser mónico. Introduzcamos más notación.

Definición 10.4 *Dado un polinomio $p(x) = p_0 + p_1 x + \cdots + p_r x^r$, llamamos* polinomio recíproco, *y denotamos $p^*(x)$ a*

$$p^*(x) = p_r + p_{r-1} x + p_{r-2} x^2 + \cdots + p_1 x^{r-1} + p_0 x^r$$

Es claro que si $p_0 \neq 0$ entonces el polinomio recíproco tiene el mismo grado que el polinomio de partida. Una simple comprobación prueba la identidad

$$p^*(x) = x^r p(x^{-1})$$

Corolario 10.10 *Si g es un generador y h su polinomio anulador, entonces el dual está generado por h^* (eventualmente transformado en mónico).*

Ejemplos.

- Sea $q = 2$. Recordando la descomposición de $x^7 - 1$, si $g(x) = x^3 + x + 1$, su polinomio anulador es

$$h(x) = (x+1)(x^3 + x^2 + 1) = x^4 + x^2 + x + 1$$

Su polinomio recíproco es

$$h^* = x^4 + x^3 + x^2 + 1$$

y la matriz de paridad es

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

- Sea $q = 3$. Tenemos que $x^4 - 1 = (x^2 + 2)(x^2 + 1)$. Entonces

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

- El código generado por $h(x)$ y el generado por $h^*(x)$ son equivalentes. Basta para ello utilizar la permutación

$$\tau = \begin{pmatrix} 1 & 2 & 3 & \dots & n-1 & n \\ n & n-1 & n-2 & \dots & 2 & 1 \end{pmatrix}$$

11. Cuerpos finitos y polinomios

Hemos visto que la existencia de códigos cíclicos depende de la factorización de $x^n - 1$ en el cuerpo en cuestión. Este problema es el que ahora nos proponemos analizar.

Sea $k = \mathbb{F}_q$ y el polinomio $x^n - 1$. Nuestra intención es encontrar la mínima extensión $k \hookrightarrow K$ tal que $x^n - 1$ descomponga en factores lineales. La teoría de Galois nos garantizan la existencia y unicidad de dicho cuerpo y nos informan de su estructura. Nosotros haremos un estudio más constructivo, pero será necesario introducir un mínimo de conceptos y resultados de la teoría de cuerpos.

Teorema 11.1 (Del elemento primitivo) *El grupo multiplicativo k^* es un grupo cíclico.*

Demostración.

El grupo multiplicativo es abeliano y finito. El teorema de estructura nos da la descomposición

$$k^* = C_{m_1} \times C_{m_2} \times \cdots \times C_{m_k}$$

donde m_i es el orden del grupo cíclico C_{m_i}. Además se tiene que m_i divide a m_{i+1}. Por lo tanto m_i divide a m_k para todo i. Como cada elemento de $a \in C_{m_i}$ cumple $a^{m_i} = 1$, también verifica que $a^{m_k} = 1$. De este resultado se obtiene que todo elemento de k^* cumple la ecuación

$$x^{m_k} - 1 = 0$$

Como estamos en un cuerpo, este polinomio no puede tener más de m_k raices. Ello implica que $m_k = q$ y el resto de los grupos cíclicos son triviales. □

Definición 11.1 *Dado un cuerpo k finito, llamamos* elemento primitivo *del cuerpo a todo generador del grupo cíclico k^*.*

Ejemplos.

- Sea $k = \mathbb{F}_8$ y sea α una solución del polinomio $1 + x + x^3$. Utilizando que $\alpha^3 = \alpha + 1$ obtenemos:

$$\begin{aligned}
&\alpha^0 = 1\\
&\alpha^1 = \alpha\\
&\alpha^2 = \alpha^2\\
&\alpha^3 = \alpha + 1\\
&\alpha^4 = \alpha^2 + \alpha\\
&\alpha^5 = \alpha^3 + \alpha^2 = \alpha^2 + \alpha + 1\\
&\alpha^6 = \alpha^3 + \alpha^2 + \alpha = \alpha + 1 + \alpha^2 + \alpha = \alpha^2 + 1\\
&\alpha^7 = \alpha^3 + \alpha = \alpha + 1 + \alpha = 1
\end{aligned}$$

  Hemos demostrado que α es un elemento primitivo de $\mathbb{F}_8$.

- El elemento primitivo generalmente no es único (un grupo cíclico puede tener varios generadores). Si β es la raíz del polinomio $1 + x^2 + x^3$, también es un elemento primitivo de $\mathbb{F}_8$.

A partir de ahora supondremos que el grado del polinomio $x^n - 1$ y la característica del cuerpo donde está definido son primos entre si, para poder aplicar el siguiente

Lema 11.2 *Si n y p son primos entre sí, el polinomio $x^n - 1$ no posee raices múltiples.*

Demostración.

La derivada es nx^{n-1}. Como p y n son primos, esta derivada es no nula. Su única solución es 0, que no es solución de $x^n - 1$. □

Lema 11.3 *Existe m tal que n divide a $q^m - 1$.*

Demostración.

Sea $q = p^r$. Como n y p son primos, tenemos que $p \in \mathbb{Z}_n$ es una unidad del anillo. Por lo tanto $q = p^r$ también es una unidad. Si m es el orden del elemento en el grupo $\mathbb{Z}_n^*$ tenemos que $q^m = 1$ en dicho anillo. Pero esto es equivalente a que n divida a $q^m - 1$. $\square$

A partir de ahora denotaremos por m al menor entero positivo que cumpla el lema anterior. Escribiremos también $ns = q^m - 1$.

Lema 11.4 *Sea $K = \mathbb{F}_{q^m}$ y α un elemento primitivo de este cuerpo. Entonces $\beta = \alpha^s$ es una solución del polinomio $x^n - 1$ y el orden de β es n.*

Demostración.

Tenemos que $\beta^n = (\alpha^s)^n = \alpha^{q^m - 1} = 1$. No puede existir ningún valor n' menor que n tal que $\beta^{n'} = 1$, pues entonces se tendría $\alpha^{sn'} = 1$ y α no podría ser un elemento primitivo. $\square$

Como β es de orden n, sus potencias

$$\beta^0 = 1, \beta, \beta^2, \ldots, \beta^{n-1}$$

son distintas y también son soluciones del polinomio $x^n - 1$. El polinomio descompone completamente en $\mathbb{F}_{q^m}$.

Definición 11.2 *Dado $k = \mathbb{F}_q$, con n primo con p, el conjunto*

$$U_n = \{\beta^0, \beta, \beta^2, \ldots, \beta^{n-1}\}$$

contenido en el cuerpo $K = \mathbb{F}_{q^m}$ es el conjunto de las raices n-ésimas de la unidad.

De la definición dada se deduce rápidamente que U_n es un grupo cíclico. Los generadores de este grupo se llaman raices n-ésimas primitivas de la unidad. Dependiendo del valor de n, pueden existir una o varias raices primitivas

Ya hemos encontrado las raices del polinomios $x^n - 1$ en una extensión del cuerpo base. Ahora nos interesa, basándonos en dicha información, encontrar los factores irreducibles del polinomio en el subcuerpo k. Los factores

irreducibles resultarán ser los polinomios mínimos de las raices n-ésimas. Introducimos las definiciones necesarias.

Dada una extensión finita $k \hookrightarrow K$, y dado un elemento $\alpha \in K$ construimos la aplicación

$$\begin{array}{rcl} k[x] & \to & K \\ p(x) & \to & p(\alpha) \end{array}$$

Esta aplicación es un morfismo de anillos y su núcleo es un ideal. Como la extensión es finita, la aplicación no puede ser inyectiva y el ideal es no nulo. Además es claro que la imagen de la aplicación es un dominio de integridad. El núcleo es un ideal primo y su generador mónico es un polinomio irreducible, que denotaremos $m_\alpha(x)$. Decimos que $m_\alpha(x)$ es el polinomio mínimo de α sobre el cuerpo k.

Hemos hablado de polinomios mínimos y hemos probado su existencia. Ahora veremos como se pueden construir. Sea $k \hookrightarrow K$ una extensión finita y $\alpha \in K$. Para construir el polinomio mínimo de α consideramos en K la estructura de espacio vectorial sobre el subcuerpo k. Si $\{1, \alpha\}$ son linealmente dependientes, entonces α está en el subcuerpo y el polinomio mínimo es $(x - \alpha)$. Si α no está en el subcuerpo consideramos el conjunto $\{1, \alpha, \alpha^2\}$. Si es linealmente dependiente, la combinación lineal igualada a cero da lugar a un polinomio de grado 2, que es el polinomio mínimo. Si este conjunto es linealmente independiente, sea r el primer entero tal que $\{1, \alpha, \ldots, \alpha^r\}$ es un conjunto de vectores dependiente. La combinación lineal igualada a cero da lugar al polinomio mínimo de α. Es evidente que ningún polinomio de grado menor puede tener a α como raíz, pues ello implicaría la dependencia lineal de un conjunto menor de potencias del elemento.

Aunque el método anterior nos permite construir polinomios mínimos, existe otro método, basado en ideas de la Teoría de Galois, que nos será más útil. Esbozaremos únicamente los conceptos necesarios.

Definición 11.3 *Sea $k \hookrightarrow K$ una extensión finita con $|k| = q$. Llamamos*

morfismo de Frobenius *de la extensión* a

$$\begin{array}{rccc} \phi_q: & K & \to & K \\ & a & \to & a^q \end{array}$$

Como q es múltiplo de la característica, aplicando el binomio de Newton se obtiene

$$(a+b)^q = a^q + b^q$$

Con este resultado, se demuestra que ϕ_q es un morfismo de anillos, inyectivo y epiyectivo. Es lo que se denomina un automorfismo del cuerpo K. La principal propiedad que cumple el automorfismo de Frobenius es

$$\phi_q(a) = a \text{ si } a \in k$$

Esto se expresa diciendo que k es un cuerpo fijo para el automorfismo.

La composición de este automorfismo da lugar a nuevos automorfismos de cuerpos, que presentan la propiedad anterior. Este modo de proceder no da lugar a un conjunto infinito de automorfismos, pues existe un valor n de tal forma que $(\phi_q)^n = \mathrm{Id}$. El menor valor de n que cumple dicha propiedad es el orden del automorfismo.

Proposición 11.5 *El orden del automorfismo de Frobenius coincide con el grado de la extensión* $k \hookrightarrow K$.

Demostración.

Si la dimensión de K sobre k es n (esto se suele denotar $[K:k] = n$), entonces K posee q^n elementos

$$(\phi_q)^n(a) = (a^q)^n = a^{q^n} = a \qquad \Rightarrow \qquad (\phi_q)^n = \mathrm{Id}$$

y necesariamente el grado del morfismo es menor que el grado de la extensión.

Además no puede existir ningún $m < n$ que cumpla dicha propiedad, pues si α es un elemento primitivo entonces $(\phi_q)^m(\alpha) = \alpha^{q^m} \neq \alpha$. $\square$

Cada automorfismo σ de un cuerpo K da lugar a un morfismo de anillos en su anillo de polinomios $K[x]$. Dicho morfismo lo seguiremos denotando σ y actua de la siguiente forma

$$\sigma(p_0 + p_1 x + \cdots + p_r x^r) = \sigma(p_0) + \sigma(p_1)x + \cdots + \sigma(p_r)x^r$$

Las siguientes observaciones conducen a un método de cálculo del polinomio mínimo.

- Si $p(x)$ tiene coeficientes en el subcuerpo k, entonces

$$\phi_q(p(x)) = p(x)$$

- Sea $p(x)$ un polinomio con coeficientes en el subcuerpo y α una raíz

$$p_0 + p_1\alpha + \cdots + p_r\alpha^r = 0$$

  Si a esta expresión le aplicamos el automorfismo de Frobenius

$$\begin{aligned} &\phi_q(p_0) + \phi_q(p_1)\phi_q(\alpha) + \cdots + \phi_q(p_r)\phi_q(\alpha)^r = 0 \\ \Rightarrow \qquad &p_0 + p_1\phi_q(\alpha) + \cdots + p_r\phi_q(\alpha)^r = 0 \end{aligned}$$

  Obtenemos que $\phi_q(\alpha)$ es también una raíz.

Definición 11.4 *Dada una extensión $k \hookrightarrow K$, se dice que dos elementos $\alpha, \beta \in K$ son* conjugados *si existe r tal que $(\phi_q)^r(\alpha) = \beta$.*

El concepto de conjugación establece una relación de equivalencia en el conjunto K. Las clases de equivalencia se llaman clases ciclotómicas. Si la extensión tiene grado n, el morfismo de Frobenius tiene orden n. Entonces las clases ciclotómicas no pueden tener más de n elementos. Si un polinomio con coeficientes en el subcuerpo k tiene una raíz, entonces sus conjugados también son raices.

Proposición 11.6 *Sea $\{\alpha_1, \alpha_2 \ldots, \alpha_r\}$ una clase ciclotómica. El polinomio*

mínimo de cualquiera de los elementos de dicha clase es

$$\prod_{1=i}^{r}(x-\alpha_i) = (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_r)$$

Demostración.

La demostración, que no realizaremos, se basa en que dicho polinomio es invariante por el automorfismo de Frobenius, lo que implica que los coeficientes del polinomio pertenecen a k. □

Ejemplos.

- Sea α el elemento primitivo de $\mathbb{F}_8$. El primer conjugado de α es α^2, el segundo conjugado es $(\alpha^2)^2 = \alpha^4$. El siguiente conjugado es $\alpha^8 = \alpha$. La clase de conjugación respecto a $\mathbb{Z}_2$ es $\{\alpha, \alpha^2, \alpha^4\}$. El polinomio mínimo es

  $$(x-\alpha)(x-\alpha^2)(x-\alpha^4)$$

  Utilizando que $\alpha^3 = \alpha + 1$, el polinomio que se obtiene es

  $$m_\alpha(x) = x^3 + x + 1$$

- En el ejemplo anterior, sea $\beta = \alpha^3$. Su clase de conjugación es

  $$\{\beta = \alpha^3, \beta^2 = \alpha^6, \beta^4 = \alpha^{12} = \alpha^5\}$$

- Sea $K = \mathbb{F}_{64}$ y α un elemento primitivo. La clase es

  $$\{\alpha, \alpha^2, \alpha^4, \alpha^8, \alpha^{16}, \alpha^{32}\}$$

  Como tiene 6 elementos, el polinomio mínimo es de grado 6.

- Es costumbre olvidarnos de escribir la raíz primitiva de la unidad y trabajar únicamente con los exponentes. La clase anterior se escribe entonces como

  $$\{1, 2, 4, 8, 16, 32\}$$

- Si α y β son conjugados, aplicando el morfismo de Frobenius al polinomio mínimo, $m_\alpha(x)$, obtenemos que β es también solución. Como es irreducible, resulta que es el polinomio mínimo de β. El recíproco también es cierto. Esto conduce a una nueva definición de la relación de conjugación: dos elementos $\alpha, \beta \in K$ son conjugados respecto a k, si sus polinomios mínimos son iguales.

Si tenemos el polinomio $x^n - 1$ y α es una solución (lo que equivale a que α sea una raíz n-ésima) entonces m_α divide al polinomio y es irreducible. La descomposición prima es el producto

$$x^n - 1 = m_{\alpha_1} \cdots m_{\alpha_r}$$

Utilizando la proposición anterior, el conocimiento de las clases de equivalencia ciclotómicas sobre k de las raices de la unidad, da lugar a la descomposición en factores primos del polinomio. De esta forma quedan clasificados todos los códigos cíclicos.

Ejemplos.

- Sea $q = 2$ y $n = 7$

  $$\{\beta^0\} \quad \{\beta^1, \beta^2, \beta^4, \beta^8 = \beta\} \quad \{\beta^3, \beta^6, \beta^{12} = \beta^3\}$$

  Las clases ciclotómicas son

  $$\{0\} \quad \{1, 2, 4\} \quad \{3, 6, 5\}$$

  El polinomio $x^7 - 1$ tiene tres factores irreducibles.

- Sea $q = 3$ y $n = 11$. Las clases son

  $$\{0\} \quad \{1, 3, 9, 5, 4\} \quad \{2, 6, 7, 10, 8\}$$

12. Códigos cíclicos II

Aunque la teoría general de los códigos cíclicos ya está esbozada en una sección anterior, no está de más ofrecer enfoques alternativos de algunos conceptos y aplicaciones específicas de lo estudiado anteriormente.

El método de codificación usado en la sección anterior produce codificaciones no sistemáticas. Analizaremos ahora otro método, basado exclusivamente en el álgebra del anillo de polinomios, que da lugar a codificaciones sistemáticas y puede simplificar alguno de los desarrollos.

Dado un polinomio $p(x)$, denotamos por

$$p(x) \bmod g(x)$$

al resto de la división de $p(x)$ entre $g(x)$. Esta operación presenta las suguientes propiedades:

- Si llamamos $s(x)$ a dicho polinomio, entonces $s(x)$ es el único polinomio de grado menor que el de $g(x)$ que cumple $p(x) = c(x) \cdot g(x) + s(x)$.
- Sea $g(x)$ de grado r. La aplicación

$$\begin{array}{rcl} k[x] & \to & R_r \\ p(x) & \to & p(x) \bmod g(x) \end{array}$$

  es lineal.

Con esta nueva notación podemos construir una nueva función lineal cuya imagen es el código cíclico. Al cambiar la función estamos cambiando la forma de codificar, aunque el código siga siendo el mismo. Sea la función

$$\begin{array}{rcl} R_m & \to & R_n \\ p(x) & \to & x^r \cdot p(x) - ([x^r \cdot p(x)] \bmod g(x)) \end{array}$$

donde $r = n - m$ es el grado del polinomio generador.

Proposición 12.1 *La aplicación anterior es una función de codificación para el código con polinomio generador $g(x)$.*

Demostración.

En primer lugar veamos que la imagen de la aplicación está contenida en el código. Efectuamos la división

$$x^r \cdot p(x) = c(x) \cdot g(x) + s(x)$$

donde $s(x) = [x^r \cdot p(x)] \bmod g(x)$. La resta $x^r \cdot p(x) - s(x)$ es un múltiplo de $g(x)$ y pertenece al código.

La aplicación es lineal, debido a las propiedades de la operación mod. También es inyectiva, pues $x^r \cdot p(x)$ es de grado mayor que r (y menor que n) y sin embargo $s(x)$ es de grado menor que r. Su diferencia nunca se puede anular. □

La matriz generadora asociada a esta aplicación se llama matriz sistemática del código. Se puede siempre escribir como

$$G = \begin{pmatrix} a_{00} & a_{01} & \dots & a_{0,r-1} & 1 & 0 & 0 & \dots & 0 \\ a_{10} & a_{11} & \dots & a_{1,r-1} & 0 & 1 & 0 & \dots & 0 \\ a_{20} & a_{21} & \dots & a_{2,r-1} & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m-1,0} & a_{m-2,1} & \dots & a_{m-1,r-1} & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Esta matriz es de la forma $G = (A, \mathrm{Id})$. Hemos comenzado las numeraciones en el cero, convenio que seguiremos siempre. Con este convenio en mente, si denotamos por A_i la fila i-ésima de la matriz A tenemos

$$A_i = -(x^{i+r} \bmod g(x))$$

Si queremos conseguir una matriz en forma estandard, basta con permutar cíclicamente todas las filas y ya se obtiene una matriz en la forma (Id, A). Sin embargo, en los códigos cíclicos, es más útil trabajar en la forma $G = (A, \mathrm{Id})$, en parte debido al siguiente resultado.

Corolario 12.2 *Dada la matriz sistemática* $G = (A, \mathrm{Id})$ *su matriz de paridad asociada,* $H = (\mathrm{Id}, -A^t)$ *cumple (la numeración empezando en* 0*)*

$$h_i = x^i \ mod\ g(x)$$

siendo h_i *la columna situada en la posición* i*.*

Demostración.

Si $i < r$ entonces $x^i \bmod g(x) = x^i$ y se obtiene la matriz identidad. La columna h_r se obtiene trasponiendo la fila 0 de la matriz y cambiando el signo. Entonces

$$h_r = -(A_0)^t = x^r \bmod g(x)$$

y con las columnas posteriores se sigue el mismo razonamiento. □

Ejemplos.

- Consideremos el código de longitud 7 con generador $x^3 + x + 1$. Para obtener la primera fila de la matriz generadora, calculamos la imagen de 1.

  $$x^3 \cdot 1 - (x^3 \bmod [x^3 + x + 1]) = x^3 + x + 1$$

  La primera fila es el vector 1101000. Para obtener la segunda fila repetimos el proceso

  $$x^3 \cdot x - (x^4 \bmod [x^3 + x + 1]) = x^4 + x^2 + x$$

  La matriz final es

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

- En el ejemplo anterior la matriz asociada se puede construir directa-

mente aplicando $H = (\mathrm{Id}, -A^t)$

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

o bien realizando las divisiones $x^i \bmod g(x)$.

Esta matriz de paridad nos permite calcular síndromes. Pero debido a la peculiar estructura de H, el síndrome se puede calcular sin realizar productos de matrices, utilizando únicamente operaciones con polinomios.

Proposición 12.3 *El* síndrome *de un polinomio* $p(x)$ *es*

$$p(x) \ mod \ g(x)$$

Demostración.

La demostración se basa en las propiedades de la operación mod. Denotamos por h_i la columna i-ésima de la matriz H. Si realizamos la multiplicación matricial obtenemos

$$p_0 h_0 + p_1 h_1 + \cdots + p_{n-1} h_{n-1}$$

Pero, con los convenios adoptados, $h_i = x^i \bmod g(x)$. Aplicando la linealidad de la operación mod tenemos

$$(p_0 + p_1 x + p_2 x^2 + \cdots + p_{n-1} x^{n-1}) \bmod g(x)$$

que es el resultado que se deseaba demostrar. □

La decodificación de un código cíclico sigue los mismos pasos que en el caso lineal. La invariancia por permutaciones cíclicas simplifica cada uno de los pasos, pero en general la simplificación no se puede considerar muy significativa. Sin embargo el siguiente caso particular puede ser sumamente útil.

Proposición 12.4 *Sea $\mathcal{C}$ un código cíclico de distancia mínima $d > 2t+1$. Si en una transmisión $p(x) \rightsquigarrow q(x)$ se producen menos de t errores y el síndrome tiene peso menor o igual que t, entonces el error cometido es exactamente igual que el síndrome.*

Demostración.

Denotemos por $e(x)$ el error y por $s(x)$ el síndrome. En virtud de las definiciones adoptadas se cumple

$$e(x) = q(x) - p(x) \qquad\qquad q(x) = c(x) \cdot g(x) + s(x)$$

Por lo tanto $e(x) - s(x) = c(x) \cdot g(x) - p(x) \in \mathcal{C}$. Pero el peso de esta resta cumple

$$\omega(e(x) - s(x)) \leq \omega(e(x)) + \omega(s(x)) \leq 2t < d$$

y como $e(x) - s(x)$ pertenece al código, necesariamente es nulo. □

Vamos ahora a demostrar que todo código de Hamming binario es equivalente a un código cíclico. Todo cuerpo k de característica 2 tiene un subcuerpo isomorfo a $\mathbb{Z}_2$. Si k posee 2^n elementos entonces, k como espacio vectorial sobre $\mathbb{Z}_2$ es de dimensión n, e isomorfo por tanto a $(\mathbb{Z}_2)^n$. Sea α un elemento primitivo del cuerpo k. Entendido como vector sobre $\mathbb{Z}_2$, α es un vector de longitud n. Formemos la siguiente matriz de paridad, cuyas entradas son números binarios

$$H = (1, \alpha, \alpha^2, \ldots, \alpha^{n-1})$$

Por el teorema del elemento primitivo, en esta matriz aparecen todos elementos no nulos del espacio vectorial. Es por tanto un código de Hamming de longitud n. Esta construcción es válida para cualquier valor de n. Si demostramos que este código que hemos construido es cíclico, se concluye que todo código de Hamming binario es cíclico.

Hagamos actuar H sobre un vector, cuya numeración la empezaremos en 0. Si $p = p_0 p_1 p_2 \ldots p_{n-1}$ entonces

$$Hp = p_0 \cdot 1 + p_1 \cdot \alpha + p_2 \cdot \alpha^2 + \cdots + p_{n-1}\alpha^{n-1} = 0$$

Si entendemos el vector como un polinomio, el resultado anterior es equivalente a $p(\alpha) = 0$. El código que hemos construido es canónicamente isomorfo al conjunto de polinomios (sobre $\mathbb{Z}_2$) de grado menor que n y que se anulan sobre α. Pero este conjunto es claramente un ideal, como prueba el siguiente

Lema 12.5 *Sea $k \hookrightarrow K$ una extensión del cuerpo. Dado una colección finita $\alpha_i \in K$, el conjunto*

$$\{p(x) \in k[x] \text{ tales que } p(\alpha_i) = 0 \text{ para todo } i\}$$

es un ideal.

Demostración.

Si $p(x)$ y $p'(x)$ se anulan sobre todos los α_i, su suma también se anula. Si ahora $q(x)$ es un polinomio arbitrario entonces $q(\alpha_i) \cdot p(\alpha_i) = 0$ y es un ideal. □

La construcción anterior es el caso más simple de los llamados códigos BCH, cuya construcción esbozamos a continuación.

Definición 12.1 *Dado un código cíclico de longitud n, decimos que $\alpha \in U_n$ es una* raíz del código *si*

$$c(\alpha) = 0 \text{ para todo } c \in \mathcal{C}$$

El conjunto de todas las raices de $\mathcal{C}$ se denota $Z(\mathcal{C})$.

Esta definición utiliza para su comprobación todos los elementos del código. Sin embargo, como todos los elementos del código son múltiplos del generador, se tiene el siguiente

Corolario 12.6 *Las raices de un código $\mathcal{C}$ coinciden con las raices de su polinomio generador.*

Hemos visto que conociendo el código, podemos conocer cuales son sus raices. Recíprocamente, si conocemos las raices, somos capaces de reconstruir

el polinomio $g(x)$, que es un divisor de $x^n - 1$ y que genera por lo tanto un código cíclico.

Dado un código cíclico, sea $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ una familia de raices del código. Con ayuda de estas raices formamos la siguiente matriz

$$\hat{H} = \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \alpha_1^3 & \ldots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \alpha_2^3 & \ldots & \alpha_2^{n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \alpha_r & \alpha_r^2 & \alpha_r^3 & \ldots & \alpha_r^{n-1} \end{pmatrix}$$

Si multiplicamos esta matriz por un vector $p_0 p_1 p_2 \ldots p_{n-1}$ y lo igualamos a cero, obtenemos un sistema de ecuaciones, $p(\alpha_i) = 0$. Por lo tanto si un vector a pertenece a $\mathcal{C}$ entonces $\hat{H} \cdot a = 0$. La matriz que hemos construido tiene cierta similitud con la matriz de paridad, pero en realidad no es una matriz de paridad, puesto que sus elementos pertenecen al cuerpo K y no a k. Sin embargo esta matriz es útil para calcular pesos mínimos, pues un vector de peso h da lugar a una combinación lineal de h de columnas igualadas a cero. Para conseguir que la distancia mínima sea mayor que un valor r debe ocurrir que todos los determinantes de tamaño r sean no nulos. La mejor manera de conseguir esto es utilizando los determinantes de Vandermonde.

Definición 12.2 *Decimos que un código tiene r raices consecutivas si*

$$\{\alpha^b, \alpha^{b+1}, \alpha^{b+2}, \ldots, \alpha^{b+r-1}\} \subset Z(\mathcal{C})$$

para cierto valor de b.

Teorema 12.7 (Cota BCH) *Si $\mathcal{C}$ tiene r raices consecutivas, entonces su distancia mínima es mayor o igual que $r + 1$.*

Demostración.

Formamos la matriz $\hat{H}$ con estas raices. Si tomamos cualquier determinante de orden r, tras unas manipulaciones algebraicas no encontramos con un determinante de Vandermonde, que es no nulo. No pueden existir vectores cuyo peso sea r o menor. □

Esta idea nos permite definir códigos cuya distancia mínima está acotada inferiormente. Para construir un código cíclico de longitud n, cuya distancia mínima sea superior a $r+1$, simplemente tomamos r raices consecutivas

$$\{\alpha^b, \alpha^{b+1}, \alpha^{b+2}, \ldots, \alpha^{b+r-1}\}$$

y tomamos como polinomio generador

$$g(x) = \text{mcm}(m_{\alpha^b}, m_{\alpha^{b+1}}, \ldots, m_{\alpha^{b+r-1}})$$

Los códigos que admiten esta construcción son los códigos BCH.

Hasta ahora hemos trabajado siempre con un generador particular del código, que presentaba dos propiedades interesantes. Por una parte era el elemento de menor grado del código y por otro era un divisor de $x^n - 1$. Trataremos ahora con otro generador distinto.

Un elemento e de un anillo es idempotente si $e^2 = e$. Los elementos idempotentes son muy importantes en la descomposición de los denominados anillos semisimples, relacionados también con la teoría de representaciones de grupos finitos. No haremos uso de esta teoría general y adaptaremos las demostraciones al anillo R_n.

Proposición 12.8 *Sea $\mathcal{C}$ un código cíclico. Existe un único elemento e idempotente que genera el ideal.*

Demostración.

Sea g el generador del ideal (el mónico de grado mínimo). Tenemos la descomposición $x^n - 1 = g \cdot h$. Como estamos suponiendo que n y la característica son primos, resulta que g y h también son primos. Aplicando el lema de Bezout

$$1 = ag + bh$$

Denotemos por e al producto ag. Si multiplicamos la identidad de Bezout

por e y pasamos al cociente, obtenemos

$$e = ee + bhe$$

pero el último sumando es nulo (pues $gh = 0$). Ya tenemos que e es idempotente.

Sea c un elemento cualquiera del código. Multiplicando la identidad de Bezout por c

$$c = ec + bhc$$

pero el último sumando vuelve a ser nulo. Ya sabemos que es generador. Además e se comporta como un elemento neutro en el conjunto. Si otro elemento idempotente genera el mismo ideal, también se comportará como un elemento neutro para el producto, de donde se deduce la unicidad. □

Si e es idempotente, entonces $1-e$ también es idempotente, y por lo tanto genera otro ideal. Este nuevo ideal está en posición de suma directa. Esta es la característica que define a los anillos semisimples.

Corolario 12.9 *Cada ideal de R_n es un sumando directo.*

Demostración.

Si I tiene como generador idempotente e entonces

$$R_n = R_n e \oplus R_n(1-e) = I \oplus R_n(1-n)$$

y la suma es directa. □

Tomemos un ideal minimal y descomponemos es suma directa. Del sumando directo volvemos a tomar un ideal minimal y continuamos el proceso. Al final se obtiene que R_n es una suma directa de ideales minimales. Este es el teorema de descomposición de anillos semisimples, aplicado al caso particular del anillo R_n. De la misma forma obtenemos que todo ideal es suma directa de los ideales minimales. Además, como cada ideal minimal carece de ideales contenidos en él, cada ideal minimal es un cuerpo.

Unas comprobaciones rutinarias demuestran que si e es el generador idempotente de $\mathcal{C}$, y e' es el generador idempotente de $\mathcal{C}'$ entonces el generador idempotente de $\mathcal{C} \cap \mathcal{C}'$ es $e \cdot e'$ y el generador idempotente de la suma $\mathcal{C} + \mathcal{C}'$ es $e + e' - e \cdot e'$.

Denotemos por $\hat{\mathcal{C}}_i$ los ideales minimales de R_n (recordemos que su polinomio anulador es un factor irreducible de $x^n - 1$), y por $\hat{e}_i$ sus generadores idempotentes. Como la intersección de estos ideales es nula, $\hat{e}_i \cdot \hat{e}_j = 0$ si $i \neq j$. Esto simplifica enormente el cálculo del generador idempotente de la suma. Como todo ideal es suma de ideales minimales, conociendo los generadores de los ideales minimales, y simplemente sumando, se pueden construir el generador idempotente de cualquier ideal. Los generadores de los ideales minimales se denominan por esta razón idempotentes primitivos.

Ejemplos.

- Sea $q = 2$ y $n = 7$. Tomemos una clase ciclotómica, por ejemplo $\{1, 2, 4\}$. Formamos un polinomio, donde la componente de x_i es no nula si i está en la clase. En nuestro caso el polinomio es

  $$x + x^2 + x^4$$

  Elevar al cuadrado en el caso binario es simple, pues $(a+b)^2 = a^2 + b^2$. Ahora es fácil comprobar que el polinomio es idempotente. En realidad este procedimiento es válido en general para cualquier código binario.

- Si tomamos una unión de clases ciclotómicas y seguimos el mismo procedimiento, obtenemos otro idempotente. Este nuevo idempotente es la suma de los idempotentes asociados a cada clase de equivalencia.

- No es difícil demostrar que todo idempotente binario se construye del modo anterior.

- Sea e el generador idempotente de un código $\mathcal{C}$ de dimensión m. Probemos que

  $$\{e, x \cdot e, \cdots, x^{m-1} \cdot e\}$$

es una base del código. Dada una combinación lineal igualada a cero, obtenemos un polinomio de grado menor que m que cumple

$$a(x) \cdot e(x) = 0$$

Multiplicando por el generador mónico

$$a(x) \cdot e(x) \cdot g(x) = a(x) \cdot g(x) = 0$$

pero esto es imposible por razones de grado. Necesariamente $a(x) = 0$ y los vectores son linealmente independientes. Esto nos permite construir matrices generadoras.

13. Códigos de Reed-Solomon

En esta sección, la longitud del código n, será un número entero menor o igual que el cardinal del cuerpo y m, la dimensión del código, un número estrictamente menor que n ($m < n \leq q$)). Denotaremos por R_m el conjunto de polinomios de grado menor que m. Tomemos n elementos distintos $\alpha_1, \alpha_2, \ldots, \alpha_n$ del cuerpo (ello es posible pues $n \leq q$). Con su ayuda construimos la función

$$\begin{array}{rcl} R_m & \rightarrow & k^n \\ p(x) & \rightarrow & (p(\alpha_1), p(\alpha_2), \ldots, p(\alpha_n)) \end{array}$$

La imagen de un polinomio $p(x)$ la denotaremos $\mathbf{p}$. Para obtener $\mathbf{p}$ hemos evaluado $p(x)$ en n puntos (que hemos denotado por α_i), prefijados de antemano y hemos colocado los resultados en forma de vector.

Definición 13.1 *Llamamos* código de Reed-Solomon *a la imagen de la aplicación anterior y lo denotamos* $RS(n, m)$.

La aplicación construida anteriormente es claramente lineal. Veamos que también es inyectiva. Si $\mathbf{p} = 0$, entonces $p(x)$ se anula sobre los n puntos α_i. Ello implica que tiene n raices. Como m es estrictamente menor que n, el polinomio no puede poseer tantas raices, salvo si es nulo. El código de Reed-Solomon $RS(n, m)$ es entonces de tipo $[n, m]$. Naturalmente si cambiamos los valores de los puntos α_i, o simplemente permutamos dichos valores, en general, se obtendrá otro código distinto, pero siempre con los mismos parámetros.

Para calcular su distancia mínima, calculemos el peso mínimo. Dado un vector $\mathbf{p}$, el peso de este vector es n menos el número de 0 que posea. Pero cada 0 del vector es una raíz de $p(x)$, que tiene, como mucho, $m-1$ raices. El peso del vector es entonces mayor o igual que $n - (m - 1)$. Si consideremos el polinomio

$$(x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_m)$$

observamos que su peso es exactamente $n - (m - 1)$. Hemos demostrado la

Proposición 13.1 *Todo código* $RS(n, m)$ *tiene distancia mínima* $d = n - m + 1$.

Como se alcanza la cota de Singleton, todo código de Reed-Solomon es MSD. Podemos construir códigos con la distancia mínima tan grande como queramos, pero ello nos lleva a emplear cuerpos con muchos elementos.

El proceso de codificación consiste en evaluar en ciertos puntos un polinomio. El proceso de decodificación se puede llevar a cabo siguiendo los pasos habituales de los códigos lineales. Pero en este caso podemos proceder de otro modo.

Si tomamos r puntos (distintos) α_i y r "imágenes" (arbitrarias, incluso repetidas) β_i, podemos construir un polinomio f, de grado menor que r, que cumpla $f(\alpha_i) = \beta_i$ para todo i. Un método rápido de construcción es el polinomio interpolador de Lagrange. Introduzcamos la siguiente notación

$$L(x) = \prod_{i=1}^{r}(x - \alpha_i) \qquad L_i = \frac{L(x)}{(x - \alpha_i)} = \prod_{i \neq j}(x - \alpha_j)$$

$L(x)$ es un polinomio mónico de grado r y $L_i(x)$ lo es de grado $r - 1$. Por construcción, $L_i(\alpha_j) = 0$ si $j \neq i$ y además $L_i(\alpha_i)$ no es nulo. El polinomio

$$f = \sum_{i=1}^{n} \frac{L_i(x)}{L_i(\alpha_i)}\,\beta_i$$

es de grado menor que r y cumple que $f(\alpha_j) = \beta_j$. Este es el polinomio interpolador de Lagrange, que es único, pues si otro polinomio g cumpliese lo mismo, el polinomio resta $f - g$ tendría más raices que su grado.

De esta forma, si tenemos el vector $\mathbf{p}$, tomamos m valores de este vector, y el polinomio interpolador de Lagrange nos permite recuperar $p(x)$.

Los polinomios $\{1, x, x^2, \ldots, x^{m-1}\}$ forman la base canónica de R_m. Su imagen es una base del código, que da lugar a la matriz generadora canónica

$$G = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \alpha_2 & \ldots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \alpha_2^2 & \ldots & \alpha_n^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha_1^{m-1} & \alpha_2^{m-1} & \alpha_2^{m-1} & \ldots & \alpha_n^{m-1} \end{pmatrix}$$

Entre los códigos de Reed-Solomon, el caso más interesante se presenta cuando $n = q - 1$. Si α es un elemento primitivo de k, entonces podemos tomar los $q - 1$ elementos no nulos del cuerpo, expresados en el orden

$$\alpha^0 = 1, \alpha, \alpha^2, \ldots, \alpha^{q-2}$$

El código de Reed-Solomon construido es de tipo $[q-1, m, q-m]$ y lo denotaremos por $RS(m)$ (se sobreentiende que la longitud es $q - 1$).

En R^n (con $n = q - 1$) podemos también considerar el código

$$\mathcal{C}_m = \{p(x) \in R_n \text{ tales que } p(\alpha) = p(\alpha^2) = \cdots = p(\alpha^{n-m}) = 0\}$$

formado por todos los polinomios que tienen como raices a los $n-m$ elementos $\alpha, \alpha^2, \ldots, \alpha^{n-m}$. Este es un código de tipo BCH, cuyo generador es

$$g(x) = (x - \alpha)(x - \alpha^2) \ldots (x - \alpha^{n-m})$$

El polinomio generador $g(x)$ tiene grado $n - m$. Ello implica que la dimensión del código es m. Ya tenemos que $RS(m)$ y $\mathcal{C}_m$ son códigos lineales de la misma dimensión. Para demostrar que son iguales necesitamos un

Lema 13.2 *Sea $n = q - 1$ y k el cuerpo con q elementos. Dado un elemento λ no nulo y distinto de la unidad, la suma*

$$1 + \lambda + \lambda^2 + \cdots + \lambda^{n-1}$$

es nula.

Demostración.

Todos los elementos no nulos de k son solución de la ecuación $x^n - 1$. Factorizamos esta ecuación

$$x^n - 1 = (x - 1)(1 + x + x^2 + \cdots + x^{n-1})$$

Todos los elementos no nulos y distintos de la unidad son raices del segundo factor. Pero eso es precisamente lo que se pretendía demostrar. □

Proposición 13.3 *El código $RS(m)$ es igual a $\mathcal{C}_m$.*

Demostración.

Como ambos son de la misma dimensión, basta con probar una de las inclusiones. Pero para probar que $RS(m) \subset \mathcal{C}_m$ es suficiente con demostrar que la base del primero está contenida dentro del segundo. Veamos que efectivamente $\mathbf{x^i} = (1, \alpha^i, \alpha^{2i}, \ldots, \alpha^{(n-1)i})$ está contenido en $\mathcal{C}_m$. Para ello lo primero que debemos hacer es escribir $\mathbf{x^i}$ como un polinomio

$$1 + \alpha^i x + \alpha^{2i} x^2 + \alpha^{3i} x^3 + \cdots + \alpha^{(n-1)i} x^{n-1}$$

Ahora debemos comprobar que este polinomio se anula al sustituir la variable por α^j siendo j un valor comprendido entre 1 y $n - m$.

$$1 + \alpha^i \alpha^j + \alpha^{2i} \alpha^{2j} + \alpha^{3i} \alpha^{3j} + \cdots + \alpha^{(n-1)i} \alpha^{(n-1)j}$$

que reescribimos de otra forma para aplicar el lema

$$1 + (\alpha^i \alpha^j) + (\alpha^i \alpha^j)^2 + \cdots + (\alpha^i \alpha^j)^{n-1}$$

Debido a los valores que pueden tomar i y j el elemento $a^i a^j = a^{i+j}$ no es nunca la unidad y se concluye por el lema anterior. □