Introducción a la Teoría de Códigos

M.A.García, L. Martínez, T.Ramírez

Facultad de Ciencia y Tecnología. UPV/EHU

Resúmenes Teóricos

Mayo de 2017

Tema 1

Preliminares sobre Álgebra Lineal

1 Espacios vectoriales: definiciones básicas

Definición Sea $(K, +, \cdot)$ un cuerpo, V un conjunto con una operación interna $+$ y una aplicación $f : K \times V \to V$. Se dice que $(V, +, f)$ es un K-**espacio vectorial** si se cumple:

1. $(V, +)$ es un grupo abeliano.
2. La aplicación f satisface las siguientes propiedades:
   (a) $f(1_K, v) = v$, $\forall v \in V$
   (b) $f(\lambda_1 + \lambda_2, v) = f(\lambda_1, v) + f(\lambda_2, v)$, $\forall \lambda_1, \lambda_2 \in K$, $\forall v \in V$.
   (c) $f(\lambda, v_1 + v_2) = f(\lambda, v_1) + f(\lambda, v_2)$, $\forall \lambda \in K$, $\forall v_1, v_2 \in V$.
   (d) $f(\lambda_1 \lambda_2, v) = f(\lambda_1, f(\lambda_2, v))$, $\forall \lambda_1, \lambda_2 \in K$, $\forall v \in V$.

Notas:

1. A la aplicación $f : K \times V \to V$ se le denomina **multiplicación por un escalar** y se suele emplear la siguiente notación:

   $$f((\lambda, v)) = \lambda v.$$

   Con esta notación las propiedades (a)–(d) de la definición de espacio vectorial pueden escribirse de la siguiente manera:

   (a) $1_K v = v$, $\forall v \in V$
   (b) $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$, $\forall \lambda_1, \lambda_2 \in K$, $\forall v \in V$.
   (c) $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$, $\forall \lambda \in K$, $\forall v_1, v_2 \in V$.

(d) $(\lambda_1\lambda_2)v = \lambda_1(\lambda_2 v)$, $\forall \lambda_1, \lambda_2 \in K$, $\forall v \in V$.

2. Si tenemos $(V, +, f)$ un K-espacio vectorial y no hay lugar a dudas de la operación interna $+$ y de la multiplicación por un escalar f usada, por abuso del lenguaje diremos simplemente "V es un K-espacio vectorial", sin especificar ni la operación interna ni la multiplicación por un escalar. Más aún, si no hay dudas del cuerpo de escalares K en el que se trabaja, nos referiremos a V como **"espacio vectorial"** en lugar de **"K-espacio vectorial"**.

3. Cuando V es un K-espacio vectorial, a los elementos de V se les llama **vectores** y a los del cuerpo K **escalares**.

Ejemplos

1. Se considera un cuerpo $(K, +, \cdot)$ y el grupo abeliano $(\mathrm{Mat}_{n\times m}(K), +)$, siendo

$$\mathrm{Mat}_{n\times m}(K) = \{(a_{ij}) \mid a_{ij} \in K, 1 \leq i \leq n,\ 1 \leq j \leq m\}$$

y $+$ la suma usual de matrices:

$$\forall (a_{ij}), (b_{ij}) \in \mathrm{Mat}_{n\times m}, \quad (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}).$$

Definimos la multiplicación por un escalar siguiente:

$$\forall \lambda \in K, \forall (a_{ij}) \in \mathrm{Mat}_{n\times m}, \quad \lambda(a_{ij}) = (\lambda a_{ij}).$$

Es fácil ver que esta multiplicación por un escalar cumple (a)–(d), así que $\mathrm{Mat}_{n\times m}(K)$ es un K-espacio vectorial.

2. Sea $(K, +, \cdot)$ un cuerpo y

$$K^n = \{(k_1, \ldots, k_n) | k_i \in K, \forall i \in \{1, 2, \ldots, n\}\}.$$

Se define en K^n la siguiente operación interna:

$$\forall (x_1, \ldots, x_n), (y_1, \ldots, y_n) \in K^n,$$

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

y la multiplicación por un escalar:

$$\forall \lambda \in K, \forall (x_1, \ldots, x_n) \in K^n,\ \lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n).$$

Es fácil ver que K^n es un K-espacio vectorial. En particular, si $K = \mathbb{F}_q$ el cuerpo finito con q elementos, siendo $q = p^t$, para algún p primo, se tiene que $\mathbb{F}_q^n$ es un $\mathbb{F}_q$-espacio vectorial para todo $n \in \mathbb{N}$.

3. Sea $(K, +, \cdot)$ un cuerpo, $n \in \mathbb{N} \cup \{0\}$ y

$$\mathcal{P}_n(K) = \{a_0 + a_1 x + \cdots + a_n x^n \mid a_i \in K,\ i \in \{0, 1, \cdots, n\}\}$$

el conjunto de los polinomios en la variable x con coeficientes en el cuerpo K y grado menor o igual a n. Si tomamos en $\mathcal{P}_n(K)$ la suma usual de polinomios y definimos la multiplicación por un escalar $\forall \lambda \in K$, $\forall a_0 + a_1 x + \cdots + a_n x^n \in \mathcal{P}_n(K)$,

$$\lambda \cdot (a_0 + a_1 x + \cdots + a_n x^n) = \lambda a_0 + \lambda a_1 x + \cdots + \lambda a_n x^n,$$

resulta que $(\mathcal{P}_n(K), +, \cdot)$ es un K-espacio vectorial.

4. Sea $(K, +, \cdot)$ un cuerpo y

$$K[x] = \{\sum_{i=0}^{m} a_i x^i | m \in \mathbb{N} \cup \{0\}, a_i \in K,\ i \in \{0, 1, \cdots, n\}\}$$

el conjunto de los polinomios en la indeterminada x con coeficientes en el cuerpo K. Si tomamos en $K[x]$ la suma usual de polinomios y definimos la multiplicación por un escalar $\forall \lambda \in K$, $\forall \sum_{i=0}^{m} a_i x^i \in K[x]$,

$$\lambda \cdot \sum_{i=0}^{m} a_i x^i = \sum_{i=0}^{m} \lambda a_i x^i,$$

resulta que $(K[x], +, \cdot)$ es un K-espacio vectorial.

En la siguiente proposición vemos algunas propiedades que se cumplen en los espacios vectoriales, cuando elegimos escalares o vectores especiales:

Proposición 1.1 *Sea V un K-espacio vectorial. Entonces,*

1. $\lambda 0_V = 0_V$, $\forall \lambda \in K$.
2. $0_K v = 0_V$, $\forall v \in V$.
3. $(-1_K)v = -v$, $\forall v \in V$.
4. *Si $\lambda \in K$ y $v \in V$ verifican que $\lambda v = 0_V$, entonces $\lambda = 0_K$ ó $v = 0_V$.*

Nota: En la Proposición anterior hemos denotado por 0_V el elemento neutro para la operación interna $+$ de V, 0_K el elemento neutro para la operación interna $+$ de K y -1_K es el elemento opuesto del elemento neutro 1_K de la segunda operación interna $\cdot$ del cuerpo K y $-v$ es el elemento opuesto de v para la suma de vectores.

2 Subespacios vectoriales

Definición Sea V un K-espacio vectorial y S un subconjunto de V no vacío. Se dice que S es un K-**subespacio vectorial** de V, y se denota por $S \leq V$, si S con las operaciones suma y multiplicación por un escalar restringidas a S (esto es, la suma de vectores de S es otro vector de S y la multiplicación de un escalar por un vector de S es un vector de S) es un K-espacio vectorial.

En la siguiente proposición, damos una caracterización equivalente del concepto de subespacio vectorial, que facilitará el comprobar si un subconjunto de un K-espacio vectorial es subespacio vectorial.

Proposición 2.1 (Caracterizaciones equivalentes) *Sea V un K-espacio vectorial y S un subconjunto de V no vacío. Entonces, son equivalentes*

1. *S es un K- subespacio vectorial de V.*
2. *$\forall s_1, s_2 \in S, \quad s_1 + s_2 \in S$ y $\forall \lambda \in K, \forall s \in S, \quad \lambda s \in S$.*
3. *$\forall \lambda_1, \lambda_2 \in K, \forall s_1, s_2 \in S, \quad \lambda_1 s_1 + \lambda_2 s_2 \in S$.*

Ejemplos

1. Si tomamos el $\mathbb{R}$-espacio vectorial $\mathbb{R}^3$, cualquier recta que pase por el $(0,0,0)$ es un subespacio vectorial de $\mathbb{R}^3$. Por ejemplo, $S = \{(x,y,z) \in \mathbb{R}^3 \mid x - 2y = 0, z + y = 0\}$ es un subespacio de $\mathbb{R}^3$ ya que para cualesquiera $(x_1,y_1,z_1), (x_2,y_2,z_2) \in S$ y cualesquiera escalares $\alpha, \beta \in \mathbb{R}$ se cumple

$$\alpha(x_1,y_1,z_1) + \beta(x_2,y_2,z_2) = (\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2, \alpha z_1 + \beta z_2)$$

que satisface

$$\alpha x_1 + \beta x_2 - 2(\alpha y_1 + \beta y_2) = \alpha(x_1 - 2y_1) + \beta(x_2 - 2y_2) = 0$$

y

$$\alpha z_1 + \beta z_2 + \alpha y_1 + \beta y_2 = \alpha(z_1 + y_1) + \beta(z_2 + y_2) = 0$$

por ser $(x_1,y_1,z_1), (x_2,y_2,z_2) \in S$.

2. Si tomamos el $\mathbb{R}$-espacio vectorial de los polinomios en la variable x con coeficientes reales y grado menor o igual a 4, $\mathcal{P}_4(\mathbb{R})$, es fácil ver que el subconjunto $S = \{a + bx + ax^3 | a, b \in \mathbb{R}\}$ es un subespacio vectorial de $\mathcal{P}_4(\mathbb{R})$.

Si tenemos un K-espacio vectorial V y tomamos vectores $v_1, v_2, \dots v_s \in V$, con $s \geq 1$, podemos construir un subespacio vectorial de V que contenga a estos vectores de la manera siguiente:

$$< v_1, v_2, \dots, v_s >= \{\sum_{i=1}^{s} \lambda_i v_i | \lambda_i \in K \, \forall i \in \{1, \dots, s\}\}.$$

En efecto, claramente el subconjunto $< v_1, v_2, \dots, v_s >$ es no vacío, ya que al menos se encuentran los vectores v_i, para $i = 1, \dots, s$. Además, si elegimos $\sum_{i=1}^{s} \lambda_i v_i, \sum_{i=1}^{s} \delta_i v_i \in < v_1, v_2, \dots, v_s >$ y tomamos dos escalares $\alpha, \beta \in K$, tenemos

$$\alpha \sum_{i=1}^{s} \lambda_i v_i + \beta \sum_{i=1}^{s} \delta_i v_i = \sum_{i=1}^{s} (\alpha \lambda_i + \beta \delta_i) v_i$$

y como para $i = 1, \dots, s$ se cumple que $\alpha\lambda_i + \beta\delta_i \in K$, se sigue que $\alpha \sum_{i=1}^{s} \lambda_i v_i + \beta \sum_{i=1}^{s} \delta_i v_i$ es un elemento de $< v_1, v_2, \dots, v_s >$. Por tanto, el subconjunto $< v_1, v_2, \dots, v_s >$ es un K-subespacio vectorial de V, al que llamaremos **subespacio generado por los vectores** $\{v_1, v_2, \dots, v_s\}$ y tiene la propiedad de ser el menor subespacio de V que contiene a los vectores $v_1, v_2, \dots, v_s$.

En la siguiente proposición mostramos que la suma e intersección de subespacios vectoriales es otro subespacio vectorial.

Proposición 2.2 *Sea V un K-espacio vectorial y S_1, S_2 dos K-subespacios de V. Entonces, $S_1 \cap S_2$ y $S_1 + S_2 = \{s_1 + s_2 \mid s_i \in S_i, i = 1, 2\}$ con la suma y la multiplicación por un escalar restringidas a ellos son K-subespacios vectoriales de V.*

A $S_1 \cap S_2$ se le llama **subespacio intersección** de S_1 y S_2 y a $S_1 + S_2$ **subespacio suma** de S_1 y S_2.

Si S_1, S_2 son dos subespacios de V tales que $S_1 \cap S_2 = \{0_V\}$ y $S_1 + S_2 = V$, diremos que V se expresa como **suma directa** de los subespacios S_1 y S_2. Cuando V sea suma directa de S_1 y S_2 escribiremos $V = S_1 \oplus S_2$ y diremos que S_1 (respectivamente, S_2) es un subespacio suplementario de S_2 (respectivamente, S_1).

Esto se puede extender cuando trabajamos con un número finito de subespacios vectoriales de un mismo espacio vectorial, en lugar de con dos, como sigue:

Definición Sea V un K-espacio vectorial y S_1, S_2, …, S_r r K-subespacios vectoriales de V. Se llama **subespacio intersección** de S_1, S_2, …, S_r a

$$\bigcap_{i=1}^{r} S_i = \{v \in V \mid v \in S_i, \, \forall i \in \{1, \dots, r\}\}.$$

Definición Sea V un K-espacio vectorial y S_1, S_2, ..., S_r r K-subespacios vectoriales de V. Se llama **subespacio suma** de S_1, S_2, ..., S_r a

$$\sum_{i=1}^{r} S_i = \{v_1 + \cdots + v_r \in V | v_i \in S_i \forall i \in \{1, \ldots, r\}\}.$$

Definición Sea V un K-espacio vectorial y S_1, S_2, ..., S_r r K-subespacios vectoriales de V. Se dice que V es **suma directa** de S_1, S_2, ..., S_r, y se escribe $V = S_1 \oplus S_2 \oplus \cdots \oplus S_r$, si se cumplen las dos siguientes condiciones:

1. $\sum_{i=1}^{r} S_i = V$.
2. $\forall i \in \{1, \cdots, r\}$, $S_i \cap \sum_{\substack{j=1 \\ j \neq i}}^{r} S_j = \{0_V\}$.

3 Base y dimensión de un espacio vectorial

En este apartado, para aquellos espacios vectoriales que lo admitan, vamos a ver cómo localizar subconjuntos minimales del espacio a partir de los cuales sea posible obtener todos los elementos del espacio vectorial usando la suma y la multiplicación por escalares. Necesitamos unas definiciones previas:

Definición Sea V un K-espacio vectorial, $v_1, \ldots, v_r \in V$ y $\lambda_1, \ldots, \lambda_r \in K$. Se llama K-**combinación lineal** de los vectores $v_1, \ldots, v_r$ con escalares $\lambda_1, \ldots, \lambda_r$ al vector $v = \lambda_1 v_1 + \cdots + \lambda_r v_r$.

Definición Un subconjunto $T \subseteq V$ del K-espacio vectorial V se dice que es un K-**sistema generador** de V si cualquier vector de V se expresa como combinación lineal de vectores de T.

Definición Se dice que un espacio vectorial es **finitamente generado** si admite un sistema generador finito.

Definición Un subconjunto $T \subseteq V$ se dice que es K-**libre** ó que sus vectores son K-**linealmente independientes** si se cumple la siguiente condición: $\forall \lambda_1, \ldots, \lambda_r \in K, \forall v_1, \ldots, v_r \in T$,

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = 0_V \Longrightarrow \lambda_1 = \cdots = \lambda_r = 0_K.$$

Definición Si un subconjunto T no es libre se dice que es K-**ligado** o que sus vectores son K-**linealmente dependientes**.

Definición Un subconjunto $\mathcal{B} \subseteq V$ del K-espacio vectorial V se dice que es una **base**, si $\mathcal{B}$ es K-sistema generador de V y $\mathcal{B}$ es K-libre.

Ejemplos

1. En el $\mathbb{R}$-espacio vectorial $\mathbb{R}^3$ el conjunto

$$\mathcal{B} = \{(1,0,0),(0,1,0),(0,0,1)\}$$

es una base de $\mathbb{R}^3$ ya que

(a) Para todo $(x,y,z) \in \mathbb{R}^3$ se escribe:

$$(x,y,z) = x(1,0,0) + y(0,1,0) + z(0,0,1)$$

con $x,y,z \in \mathbb{R}$, luego $\mathcal{B}$ es un sistema generador para $\mathbb{R}^3$.

(b) El conjunto $\mathcal{B}$ es libre ya que

$$(0,0,0) = \alpha(1,0,0) + \beta(0,1,0) + \gamma(0,0,1) = (\alpha,\beta,\gamma)$$

implica

$$\alpha = \beta = \gamma = 0.$$

2. Usando el mismo argumento que en el ejemplo anterior, es sencillo ver que en el K-espacio vectorial K^n, donde $n \in \mathbb{N}$, el conjunto

$$\mathcal{B} = \{(1_K,0_K,\ldots,0_K),(0_K,1_K,0_K,\ldots,0_K),\ldots,(0_K,\ldots,0_K,1_K)\}$$

es una base de K^n.

3. En el $\mathbb{R}$-espacio vectorial $\mathcal{P}_3(\mathbb{R})$ es fácil comprobar que el conjunto

$$\mathcal{B} = \{1, x, x^2, x^3\}$$

es una base de $\mathcal{P}_3(\mathbb{R})$.

La siguiente proposición nos da una propiedad de los conjuntos generadores que además sean ligados:

Proposición 3.1 *T un K-sistema generador K-ligado del K-espacio vectorial V, entonces existe un vector $v \in T$ que es K-combinación lineal de vectores de $T - \{v\}$ y $T - \{v\}$ es un sistema generador de V.*

Esta propiedad se utiliza para probar:

Teorema 3.2 (Existencia de base) *Sea V un K-espacio vectorial no nulo finitamente generado. Entonces, existe $\mathcal{B}$ base de V.*

La clave de la demostración del resultado anterior está en elegir un sistema generador finito y analizar si es libre o no. Si es libre, se tiene ya la base buscada. Si no es libre, se construye otro sistema generador a partir de él quitando un vector que sea combinación lineal de los restantes del nuevo sistema generador. Obviamente, este nuevo sistema generador tiene un vector menos que el inicial y podemos reiterar el proceso indicado (eliminar uno que sea combinación lineal de los restantes que queden) hasta llegar a un conjunto que sea libre. El hecho de que V sea no nulo y finitamente generado garantiza que el proceso acabará en un número finito de pasos.

Una vez que sabemos que un K-espacio vectorial no nulo finitamente generado admite una base, nos preocupamos por saber si dos bases de un mismo K-espacio vectorial no nulo finitamente generado tienen el mismo cardinal. La respuesta va a ser positiva. Pero para poder probarlo hay que usar el siguiente resultado que se usa mucho en Álgebra Lineal:

Teorema 3.3 (Teorema del Reemplazamiento) *Sea V un K-espacio vectorial no nulo, $T = \{t_1, \ldots, t_m\} \subseteq V$ un K-sistema generador y $U = \{u_1, \ldots, u_r\} \subseteq V$ un subconjunto K-libre. Entonces,*

1. *$r \leq m$.*
2. *Existe $S \subseteq T$ tal que $|S| = m - r$ y $U \cup S$ es un K-sistema generador de V.*

La clave para probar el Teorema del Reemplazamiento es usar que si V un K-espacio vectorial y si consideramos $\{u_1\} \cup T = \{u_1, t_1, \ldots, t_m\} \subseteq V$ (que es también sistema generador), entonces $\{u_1\} \cup T$ es un subconjunto K-ligado tal que $u_1 \neq 0_V$ y, por tanto, existe $t_j \in T$ tal que t_j es K-combinación lineal de $u_1, t_1, t_2, \ldots, t_{j-1}$. Si llamamos $T' = \{u_1, t_1, t_2, \ldots, t_{j-1}, t_{j+1}, \ldots t_m\}$, resulta que T' es un sistema generador y podemos aplicar el mismo razonamiento que acabamos de hacer, pero ahora con u_2 y T' y así hasta introducir los r vectores del conjunto libre.

Tal y como hemos comentado, el Teorema del Reemplazamiento se utiliza en la demostración del siguiente resultado relativo al cardinal de las bases de un espacio vectorial finitamente generado:

Teorema 3.4 *Sea V un K-espacio vectorial no nulo finitamente generado. Entonces, todas las bases de V poseen en mismo cardinal.*

Al tener todas las bases de un K-espacio vectorial no nulo finitamente generado el mismo cardinal tiene sentido dar la siguiente definición:

Definición Sea V un K-espacio vectorial no nulo finitamente generado. Se llama **dimensión** de V, y se denota por $\dim_K(V)$, al cardinal de cualquier base de V. Si $V = \{0_V\}$, se dice que su dimensión es 0 y su base es el conjunto vacío.

Ejemplos

1. En el $\mathbb{R}$-espacio vectorial $\mathbb{R}^3$ hemos visto que el conjunto

$$\mathcal{B} = \{(1,0,0),(0,1,0),(0,0,1)\}$$

   es una base de $\mathbb{R}^3$, así que $\mathbb{R}^3$ es de dimensión 3.

2. En el K-espacio vectorial K^n, donde $n \in \mathbb{N}$, sabemos que el conjunto $\mathcal{B} = \{(1_K,0_K,\ldots,0_K),(0_K,1_K,0_K,\ldots,0_K),\ldots,(0_K,\ldots,0_K,1_K)\}$ es una base de K^n. Por tanto, K^n tiene dimensión n.

3. El $\mathbb{R}$-espacio vectorial $\mathcal{P}_3(\mathbb{R})$ tiene dimensión 4 ya que el conjunto

$$\mathcal{B} = \{1, x, x^2, x^3\}$$

   es una base de $\mathcal{P}_3(\mathbb{R})$.

El conocer la dimensión de un espacio vectorial nos puede ser útil para saber si subconjuntos del espacio vectorial son o no bases. Lo resumimos en las siguientes proposiciones:

Proposición 3.5 *Sea V un K-espacio vectorial de dimensión n. Si $\{v_1,\ldots,v_n\} \subseteq V$ es un K-sistema generador, entonces $\{v_1,\ldots,v_n\}$ es una base de V.*

Proposición 3.6 *Sea V un K-espacio vectorial de dimensión n. Si $\{v_1,\ldots,v_n\} \subseteq V$ es K-libre, entonces $\{v_1,\ldots,v_n\}$ es una base de V.*

Así, si tenemos un subconjunto de V de cardinal la dimensión de V, entonces basta con que se cumpla una de las dos condiciones de la definición de base para garantizar que también se cumple la otra.

Además, se puede demostrar que los subespacios vectoriales de espacios vectoriales de dimensión finita son también de dimensión finita y ésta es siempre menor o igual que la del espacio vectorial que les contiene. En concreto,

Teorema 3.7 *Sea V un K-espacio vectorial de dimensión finita n y W un K-subespacio de V. Entonces,*

1. *W es finitamente generado y* $\dim_K(W) \leq \dim_K(V)$.

2. *Si* $\mathcal{B}_W = \{w_1, \ldots, w_r\}$ *es una base de* W*, entonces existen* $v_{r+1}, \ldots, v_n \in V$ *tales que* $\mathcal{B}_W \cup \{v_{r+1}, \ldots, v_n\}$ *es una base de* V*.*

El paso clave para demostrar el apartado 2 del Teorema anterior es el siguiente:

Se toma a $\mathcal{B}_W = \{w_1, \ldots, w_r\} = U_1$ como conjunto libre de V y una base $\mathcal{B}_V = \{t_1, \ldots, t_n\} = T_1$ de V como sistema generador de V y se aplica el Teorema del Reemplazamiento. Los vectores de $S \subseteq T_1$ tales que $\mathcal{B}_W \cup S$ es sistema generador de V son los que aparecen en el enunciado.

A este proceso se le llama **completar la base de** W hasta obtener una base de V. Además, si consideramos el subespacio vectorial generado por $v_{r+1}, \ldots, v_n$, esto es, $U =< v_{r+1}, \ldots, v_n >$, resulta que $V = W \oplus U$, es decir, que U es un subespacio suplementario de W.

Por otro lado, como consecuencia de las dos últimas proposiciones y de este último resultado tenemos:

Teorema 3.8 *Sea* V *un* K*-espacio vectorial de dimensión finita* n *y* W *un* K*-subespacio de* V*. Entonces,* $W = V$ *si y sólo si* $\dim_K(W) = \dim_K(V)$.

Nos preguntamos qué sucede con las dimensiones de $U + W$ y de $U \cap W$, si tenemos un K-espacio vectorial de dimensión finita y U y W dos K-subespacios de V. Se puede demostrar que existe una relación entre ellas, tal y como se enuncia en la siguiente proposición:

Proposición 3.9 *Sea* V *un* K*-espacio vectorial de dimensión finita y sean* U*,* W *dos* K*-subespacios de* V*. Entonces,*

$$\dim_K(U + W) = \dim_K(U) + \dim_K(W) - \dim_K(U \cap W).$$

Los pasos necesarios para probar el resultado anterior son:

- Se toma una base de $U \cap W$, que denotamos por $\mathcal{B}_{U \cap W}$.
- Se completa $\mathcal{B}_{U \cap W}$ hasta obtener una base de U, $\mathcal{B}_U = \mathcal{B}_{U \cap W} \cup T_1$ y otra de W, $\mathcal{B}_W = \mathcal{B}_{U \cap W} \cup T_2$.
- Se comprueba que $\mathcal{B}_{U \cap W} \cup T_1 \cup T_2$ es base de $U + W$.

4 Coordenadas de un vector

Cuando estamos trabajando en un K-espacio vectorial V de dimensión finita y tenemos una base de él, sabemos que cada vector de V se expresa como combinación

lineal de los elementos de la base. Pero ¿es esta combinación de vectores de la base la única que nos puede dar ese vector? El siguiente resultado responde a esta cuestión:

Teorema 4.1 *Sean V un K-espacio vectorial de dimensión finita n, v un vector de V y $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base de V. Entonces, existen unos únicos escalares $\lambda_1, \ldots, \lambda_n \in K$ tales que $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$.*

Si entendemos que en una base de un K-espacio vectorial de dimensión finita n el orden en que aparecen los vectores de la base es una característica de ella y que si cambiamos de orden dos vectores de esa base hacen que obtengamos una base diferente (aunque como conjuntos ambos sean el mismo), el resultado anterior justifica que a los escalares que usamos en la combinación lineal que nos da el vector reciban un nombre:

Definición Se llaman **coordenadas del vector** v en la base $\mathcal{B} = \{v_1, \ldots, v_n\}$ a los únicos escalares $\lambda_1, \ldots, \lambda_n \in K$ tales que $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$.

Observamos que el orden que indicamos es fundamental, porque el escalar que ocupa la posición i acompaña al i-ésimo vector de la base. Por convenio, si V es un K-espacio vectorial de dimensión n, $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base de V y v es un vector con coordenadas $\lambda_1, \ldots, \lambda_n$ respecto de $\mathcal{B}$, escribiremos estas coordenadas mediante una matriz fila $(\lambda_1 \ldots \lambda_n)$. Esto nos permitirá expresar el vector v de forma matricial como sigue:

$$v = (\lambda_1 \ldots \lambda_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Ejemplos

1. Se considera $\mathcal{P}_3(\mathbb{R})$ el $\mathbb{R}$-espacio vectorial de los polinomios de grado menor o igual que 3 en la variable x, esto es,

   $$\mathcal{P}_3(\mathbb{R}) = \{a_0 + a_1 x + a_2 x^2 + a_3 x^3 | a_i \in \mathbb{R},\ i \in \{0, \ldots, 3\}\}.$$

   Es fácil ver que $\mathcal{B}_1 = \{1, 1+x, 1+x^2, x^3\}$ es una base de $\mathcal{P}_3(\mathbb{R})$. Si tomamos el vector $v = 5 + x + x^3 \in \mathcal{P}_3(\mathbb{R})$, se tiene que

   $$v = 5 + x + x^3 = 4 \cdot 1 + 1(1+x) + 0(1+x^2) + 1x^3,$$

   luego las coordenadas de $5 + x + x^3$ en la base $\mathcal{B}_1$ vienen dadas por (4 1 0 1). En cambio, si tomamos la base $\mathcal{B}_2 = \{1+x, 1, 1+x^2, x^3\}$, que se ha obtenido cambiando el orden de dos vectores de $\mathcal{B}_1$, entonces las coordenadas del mismo vector $v = 5 + x + x^3$ son (1 4 0 1).

2. Si tomamos el $\mathbb{F}_3$-espacio vectorial $\mathbb{F}_3^4$, el conjunto $\mathcal{B} = \{(1,1,1,1),(0,1,1,1),(0,0,1,1),(0,0,0,1)\}$ es una base de $\mathbb{F}_3^4$. El vector $(1,2,0,1) \in \mathbb{F}_3^4$ se expresa como combinación lineal de vectores de $\mathcal{B}$ mediante:

$$(1,2,0,1) = 1(1,1,1,1) + 1(0,1,1,1) + 1(0,0,1,1) + 1(0,0,0,1),$$

por tanto, las coordenadas de $(1,2,0,1)$ en la base $\mathcal{B}$ son (1 1 1 1). Obsérvese que en $\mathbb{F}_3$, se cumple que 1+1+1=0

En resumen,

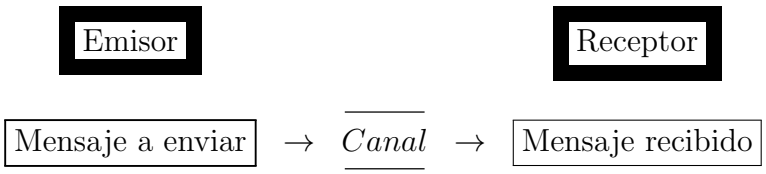
1. Si V es un K-espacio vectorial de dimensión n, las coordenadas de cualquier vector siempre son n escalares ordenados de K, siendo su orden el que marca el orden de los vectores de la base.
2. Las coordenadas de un vector dependen de la base que se elija. El mismo vector puede tener coordenadas diferentes en bases distintas. El único vector que tiene siempre las mismas coordenads en todas las bases es el vector 0_V.
3. Si nos cambian la base, las mismas coordenadas pueden representar vectores diferentes.

Tema 2

Nociones básicas de la Teoría de Códigos

1 Introducción: El problema de la transmisión de la información

En la transmisión de la información nos encontramos con que un emisor manda un mensaje a través de un canal para que le llegue al receptor. Esto es, podemos representar los elementos que tenemos en el proceso de transmisión de una información mediante el siguiente esquema:



Sin embargo, durante la transmisión de la información puede haber problemas debido a interferencias que se produzcan en el canal. Estas interferencias pueden traducirse en dos situaciones diferentes:

1. Al receptor no le llega el mensaje que le ha enviado el emisor, sino que reciba un mensaje diferente del original por haber habido una mala transmisión.
2. El mensaje enviado ha sido interceptado por alguien que no es el destinatario final y este usurpador ha manipulado o ha hecho un uso indebido del mensaje captado.

Así, estas situaciones diferentes que se pueden darse sobre el canal usado nos llevan a realizar una clasificación del canal en dos tipos diferenciados:

1. Canal poco fiable: Hay "ruido" en el canal que dificulta la transmisión. Un ejemplo de este tipo de canal se tiene cuando intentamos leer un CD que presenta huellas de dedos que obstaculizan el acceso a la información grabada en él o cuando recibimos la señal de un satétite muy alejado de la Tierra, por lo que ésta llega muy debilitada, y no sabemos el valor exacto que se ha emitido.

2. Canal poco seguro: El canal que transmite la información soporta el ataque de espías que pueden manipular la información que transcurre por él o hacerse con ella sin ser el destinatario final. Por ejemplo, estaríamos ante un canal poco seguro cuando al hacer una compra por Internet con VISA, el ordenador que utilizamos para enviar la información de nuestra tarjeta tiene instalado un programa espía que captura los datos de nuestra tarjeta y se los pasa a otra persona para que haga otras compras que no hemos autorizado con nuestros datos, ó cuando el que intercepta el mensaje lo manipula antes de que continúe viajando por la red alterando el significado del mismo (p.e., hemos autorizado un pago de 30 euros a un vendedor X y el que lo altera indica al banco que autorizamos un cargo a nuestra cuenta de 300 euros, 30 para el vendedor X y 270 para abonar a Y).

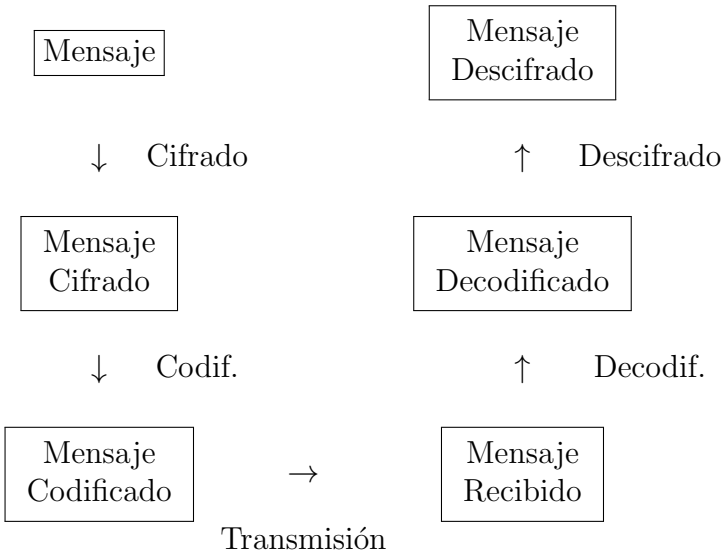
En cualquiera de los dos casos, es necesario disponer de herramientas que nos ayuden a solventar, en la medida de lo posible, este problema de interferencias en el canal. De hecho, las Matemáticas nos permiten solucionar en muchos casos esta dificultad y, tras un proceso más o menos complejo que debe aplicarse a la información que va a ser enviada por el canal, es posible "recuperar" el mensaje inicial (en el caso de canales poco fiables) o "esconder" la información que viaja por el canal, de forma que cualquiera que la intercepte no sepa de qué se trata (cuando el canal es poco seguro). Obviamente, según el tipo de interferencias que podamos sufir en el canal deberemos adoptar acciones diferentes. Así,

1. Si el canal es poco fiable, usaremos los llamados **códigos detectores y correctores de errores** que nos permiten al recibir un mensaje conocer si se ha producido algún fallo en la transmisión (esto es, detecta los errores producidos) y, en ciertos casos, incluso puede saberse cuál fue el mensaje inicial enviado. En esencia, los códigos detectores y correctores añaden información a la inicial, conociendose este proceso como codificación del mensaje, de forma que con la información extra añadida sepamos si se ha producido algún error en el mensaje recibido y, si es así, aplicarle un proceso llamado de decodificación, que permite recuperar (si es posible) el mensaje enviado.

2. Si el canal es poco seguro, la herramienta matemática de la que disponemos es el uso de los conocidos como **sistemas criptográficos**, que antes de enviar la información, la transforman en otra sin sentido para el que pretende hacerse con ella de forma ilegal y al ser recibida por el receptor la información transformada, éste revierte la transformación realizada para obtener el mensaje real que le quería mandar el emisor. El proceso de modificar la información para que carezca de sentido para el que escucha sin ser el receptor se conoce como

proceso de encriptado o cifrado y la acción de revertir el mismo, que lo hace el receptor, es el proceso de desencriptado o descifrado.

Hay veces que por las características de la información a enviar y del canal a utilizar será preciso realizar una combinación de ambas herramientas: en primer lugar se esconde la información, esto es, se cifra. Luego se le añade información extra, es decir, se codifica. A continuación se pasa por el canal y el receptor aplica primero un proceso de decodificación y al mensaje resultante uno de descifrado para llegar al mensaje original que deseaba hacerle llegar el emisor. Si queremos resumir esquemáticamente de este proceso, podemos usar el siguiente diagrama:



En cualquier caso, tanto al usar códigos detectores y correctores de errores como sistemas criptográficos, debemos ser especialmente cuidadosos en la elección del tipo de código o sistema a utilizar: debe ser acorde a las posibles interferencias que soporte el canal. Por ejemplo, si queremos mandar una información a través de un canal no muy fiable del que sabemos que cada 10 dígitos que pasamos por él se produce un error y queremos solventar este problema, de entre los códigos correctores de errores deberemos elegir uno que se ajuste a lo que nosotros necesitamos. Así, sería factible elegir uno que nos permita corregir uno o dos errores en cada bloque de 10 dígitos, pero no uno que permita corregir un error producido cada 100 datos (sería escaso) u otro que nos permita corregir uno de cada dos (podría ser demasiado costoso). Del mismo modo, cuando debamos trabajar con un canal poco seguro, deberemos valorar el tiempo y coste necesarios de los procesos de cifrado /descifrado de mensaje para adecuarlos al canal que usamos.

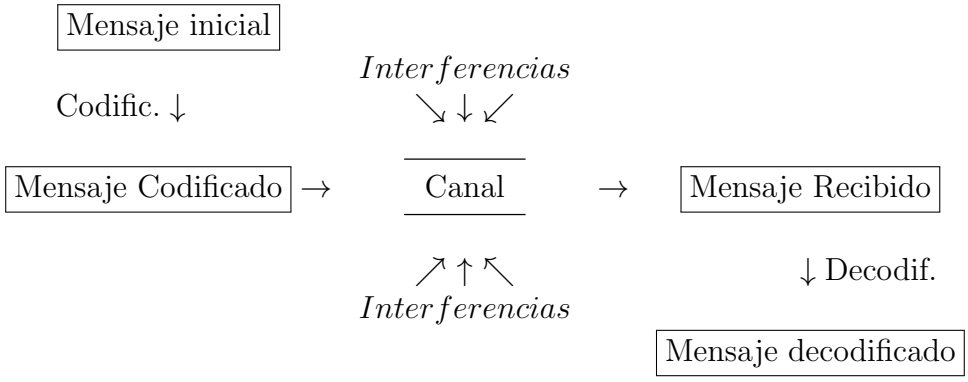
En este curso nos centraremos en estudiar algunos de los códigos detectores y correctores de errores, que son estudiados por la parte de las Matemáticas conocida como Teoría de Códigos, mientras que el estudio de los sistemas criptográficos, de los que se ocupa la Criptografía, se dejará para cursos posteriores. Asimismo, dentro del estudio de los códigos detectores y correctores de errores nos centraremos en analizar las propiedades y manejo de los llamados códigos lineales, que son los más sencillos

de trabajar. El estudio de otros tipos de códigos correctores de errores, como por ejemplo, los algebro-geométricos que usan curvas elípticas en sus definiciones, transciende el objetivo de este curso que lo que pretende es exponer de forma rigurosa, desde el punto de vista matemático, las características fundamentales de los códigos lineales, proporcionando de este modo una primera aproximación a la matemática más sencilla que está detrás de la Teoría de Códigos. No obstante, en este capítulo nos preocupamos por formalizar los conceptos más importantes sin restringirnos, en muchos del ellos, al marco exclusivo de los códigos lineales, sino que estas definiciones son válidas para los códigos detectores y correctores en general.

2 Códigos correctores y detectores de errores

Como se ha indicado en la sección anterior, el problema con el que nos enfrentamos es que deseamos enviar un mensaje a un receptor a través de un canal que puede ser poco fiable, esto es, el mensaje que le llega no tiene que ser el emitido y puede contener errores de transmisión. Lo que nos interesa es que el receptor, tras recibir el mensaje, pueda detectar si el mensaje que le ha llegado es el original y si no lo es, intentar recuperar el mensaje inicial.

De forma esquemática, resumimos las fases que componen todo el proceso como sigue:



Buscamos que el mensaje decodificado sea lo más parecido posible al mensaje codificado. Es más, nos interesa haber elegido bien el código de acuerdo al canal utilizado para poder tener una probabilidad muy alta de que el mensaje codificado y el decodificado sean iguales para que pueda el receptor deducir cuál era el mensaje inicial a partir del decodificado, sabiendo el código que se ha empleado.

En el estudio de los códigos correctores y detectores de errores usaremos algunas definiciones básicas que conviene que precisemos a qué se refieren cuando las empleemos. Por ello, a continuación vamos a formalizar matemáticamente conceptos que aparecen de forma frecuente, como los siguientes:

1. **Alfabeto**: Es un conjunto finito de elementos, llamados letras. Así, matemáticamente un alfabeto es un conjunto

$$A = \{a_1, \ldots, a_m\},$$

siendo cada $a_i \in A$ una letra del alfabeto A y $m \in \mathbb{N}$ el cardinal del alfabeto A.

2. **Palabra**: Fijado un alfabeto A, una palabra sobre el alfabeto A es un elemento formado por la concatenación de un número finito de letras de A. Salvo que se indique lo contrario, denotaremos las palabras por letras minúsculas del alfabeto latino en negrita. Por ejemplo, $\mathbf{x}= x_1 \ldots x_n$, donde $x_i \in A$ para $i = 1, \ldots, n$ es una palabra sobre A que se ha formado al poner de forma consecutiva las letras x_1,...,x_n. En muchos textos a estas palabras se les denomina "*palabras positivas*" para distinguirlas de aquellas en las que las letras x_i pueden venir afectadas del exponente -1.

3. Ω_A: Es el semigrupo libre con base A, esto es, el conjunto de todas las palabras (positivas) que se pueden formar con las letras del alfabeto A.

4. **Código**: Es un conjunto de palabras sobre un mismo alfabeto, es decir, $C \subseteq \Omega_A$. A los elementos de C se les llama palabras del código C.

5. **Longitud de una palabra** $l(\mathbf{x})$: Número de letras que tiene una palabra. Si $\mathbf{x}= x_1 \ldots x_n$, entonces $l(\mathbf{x}) = n$.

6. **Igualdad de palabras**: Dos palabras sobre el mismo alfabeto A $\mathbf{x}$ e $\mathbf{y}$ son iguales si tienen la misma longitud y coinciden los valores de cada letra en cada posición. Esto es,

$$\mathbf{x} = \mathbf{y} \Longleftrightarrow \mathbf{x} = x_1 \cdots x_n, \quad \mathbf{y} = y_1 \cdots y_n \text{ y } x_i = y_i, \forall i = 1, \ldots, n.$$

7. **Código de bloque**: Dado un código $C \subseteq \Omega_A$, se dice que C es un código de bloque si todas las palabras de C tienen la misma longitud n. También se suele llamar código de longitud n.

8. **Distancia de Hamming**: Dadas $\mathbf{x}$ e $\mathbf{y}$ dos palabras de la misma longitud se llama distancia de Hamming entre $\mathbf{x}$ e $\mathbf{y}$ al número de componentes en que difieren ambas palabras. Esto es, si $\mathbf{x}= x_1 \ldots x_n$ e $\mathbf{y}= y_1 \ldots y_n$,

$$d(\mathbf{x}, \mathbf{y}) = |\{i \in \{1, \ldots, n\} \,|\, x_i \neq y_i\}|,$$

donde $d(\mathbf{x}, \mathbf{y})$ denota la distancia de Hamming entre las palabras $\mathbf{x}$ e $\mathbf{y}$).

9. **Distancia mínima de un código de bloque** C **de longitud** n: Dado $C \subseteq \Omega_A$ tal que C es un código de bloque, se llama distancia mínima de C a la menor de las distancias entre palabras diferentes de C, esto es,

$$d(C) = \min\{d(x_1 \ldots x_n, y_1 \ldots y_n) \,| \quad x_1 \ldots x_n, y_1 \ldots y_n \in C, \; x_1 \ldots x_n \neq y_1 \ldots y_n\}$$

10. **Peso de una palabra**: Si $\mathbf{x} \in \mathbb{F}_q^n$, se llama peso de $\mathbf{x}$, y se denota por $w(\mathbf{x})$, al número de componentes no nulas de $\mathbf{x}$.

11. **Peso mínimo de un código**: Si $C \subseteq \mathbb{F}_q^n$, se llama peso mínimo de C, y se denota por $w(C)$, a

$$w(C) = \min\{w(\mathbf{x}) \,|\, \mathbf{x} \in C - \{\mathbf{0}\}\}.$$

12. **Detectar hasta t errores**: Saber que se ha producido t fallos en la transmisión, esto es, que si trabajamos con un código C que detecta t errores, entonces cualquier palabra $\mathbf{y}$ que se obtenga de una palabra de $\mathbf{c} \in C$ cambiando el valor que figura en $\mathbf{c}$ en a lo sumo t posiciones, resulta que $\mathbf{y} \notin C$.

13. **Corregir hasta t errores**: Ser capaz de recuperar la palabra original cuando se han producido t fallos en la recepción de la misma, es decir que si $\mathbf{y} \notin C$ se ha obtenido a partir de $\mathbf{c} \in C$ cambiando el valor que figura en $\mathbf{c}$ en a lo sumo t posiciones, entonces la única palabra de C que dista de $\mathbf{y}$ a lo sumo t es la propia $\mathbf{c}$. En este caso diremos que $\mathbf{c}$ es la decodificación (única) de $\mathbf{y}$.

14. **Principio de máxima verosimilitud**: En las transmisiones es más probable que se produzca siempre el menor número de fallos posible. Seguiremos este principio en todo el curso.

Ejemplos

1. El cuerpo finito $\mathbb{F}_q$, con $q = p^t$, siendo p número primo y $t \in \mathbb{N}$, es un ejemplo de alfabeto que utilizaremos en este curso. Si $q = 2$, esto es, el alfabeto es $\mathbb{F}_2 = \{0, 1\}$ y tomamos un código C de $\Omega_{\mathbb{F}_2}$, se dice que el código C es binario.

2. $\mathbf{x}$=01010 es una palabra de longitud 5 usando como alfabeto $\mathbb{F}_2$ o en general $\mathbb{F}_q$.

3. Si tomamos $A = \mathbb{F}_2 = \{0, 1\}$ y

$$C = \{\mathbf{x} \in \Omega_A \mid \mathbf{x} \text{ contiene un número par de “1”}\},$$

entonces C es un código binario. Una palabra de C es $\mathbf{x}$= 011001010 y la longitud de $\mathbf{x}$ es 9. C no es un código de bloques porque las palabras $\mathbf{y}$= 011 y $\mathbf{z}$= 1010 están en C y tienen longitud distinta. Por otro lado, la palabra $\mathbf{y}$= 011001000 $\in \Omega_A$ no pertenece a C. Otra forma de caracterizar los elementos de C es la siguiente:

$$C = \{\mathbf{x} = x_1 \dots x_m \in \Omega_A \mid m \in \mathbb{N}, \sum_{i=1}^{m} x_i \equiv 0 \bmod 2\}.$$

4. Los siguientes ejemplos son ejemplos conocidos de códigos de bloque:

- Código ASCII: Es un código que se emplea para transmitir la información desde el teclado del ordenador a la CPU. A cada letra y símbolo del teclado se le asigna un número entre 0 y 127. Este número se representa en el sistema binario como una 7-tupla, que se completa con un octavo dígito (0, ó 1) de forma que la 8-tupla resultante tenga un número par de "1". Por ejemplo a la letra A le corresponde el 65 que se representa por la 7-tupla 1000001 y se añade a esta 7-tupla un "0" de forma que la 8-tupla resultante tenga un número par de 1, dando lugar en este caso a la 8-tupla 10000010 que es la que correspondería a la letra A.
- Código ISBN: Es un código que identifica de forma única a los libros publicados. A cada libro se le asocia un número de diez cifras $a_1 \dots a_{10}$ en la que las nueve primeras, que toman valores entre 0 y 9, dan información sobre el libro (país, editorial, título,...) y la última se obtiene exigiendo que $\sum_{i=1}^{10} i a_i \equiv 0 \bmod 11$. Si a_{10} toma el valor 10, se sustituye éste por X, por lo que $a_{10} \in \{0, 1, \dots, 9, X\}$. Por ejemplo, una palabra que forma parte de este código ISBN es 0-19-853803-0.
- Código EAN: Es un código que permite identificar de forma única a los productos que se vendenen Europa. A cada producto se le asocia un número de trece cifras $a_0 \dots a_{12}$, con $a_i \in \{0, 1, \dots, 9\}$ para $i = 0, 1, \dots, 12$, donde los 3 primeros dígitos representan el código del país en donde radica la empresa que lo comercializa, los 4 o 5 siguientes son el código de empresa, que identifica al propietario de la marca y es asignado por la asociación de fabricantes y distribuidores (AECOC). Los 5 o 4 siguientes son el código de producto y el último dígito (a_{12}) es un dígito, llamado de control, que se obtiene de forma que $3\sum_{i=0}^{5} a_{2i+1} + \sum_{i=0}^{6} a_{2i} \equiv 0 \bmod 10$.

3 Distancia de Hamming

Hemos introducido la distancia de Hamming que nos mide en cuántas componentes difieren dos palabras de la misma longitud. El utilizar el nombre de "distancia" está justificado desde el punto de vista topológico, ya que es fácil demostrar que:

Proposición 3.1 *Sea $A = \{a_1, \dots, a_m\}$ un alfabeto, T_n el conjunto de las palabras sobre el alfabeto A de longitud n, esto es, $T_n = \{x_1 \dots x_n \,|\, x_i \in A, i = 1, \dots, n\}$ y $d : T_n \times T_n \to \{0, 1, 2, \dots, n\}$ la aplicación definida por*

$$d(x_1 \dots x_n, y_1 \dots y_n) = |\{i \in \{1, \dots, n\} \,|\, x_i \neq y_i\}|,$$

para todo $x_1 \dots x_n, y_1 \dots y_n \in T_n$. Entonces, (T_n, d) es un espacio métrico.

Al ser (T_n, d) un espacio métrico tiene sentido construir sus bolas cerradas. Recordemos que, dado $r \in \mathbb{N} \cup \{0\}$ y $\mathbf{x} \in T_n$, se llama **bola cerrada de centro x y radio**

r al conjunto

$$\overline{B}(\mathbf{x}, r) = \{\mathbf{y} \in A^n \mid \mathrm{d}(\mathbf{x}, \mathbf{y}) \leq r\}.$$

Obviamente, si $r \geq n$, $\overline{B}(\mathbf{x}, r) = T_n$. Más aún, es sencillo determinar el número de elementos de una bola cerrada cuando $r = 0, 1, \ldots, n$, tal y como nos lo indica el siguiente resultado:

Lema 3.2 *Sea $A = \{a_1, \ldots, a_m\}$ un alfabeto, $\boldsymbol{x} = x_1 \ldots x_n \in T_n$ y $0 \leq r \leq n$. Entonces,*

$$|\overline{B}(\boldsymbol{x}, r)| = \sum_{i=0}^{r} \binom{n}{i} (m-1)^i.$$

El Lema anterior va a ser útil para demostrar la conocida Cota de Hamming:

Proposición 3.3 (Cota de Hamming) *Sea $C \subseteq T_n$ un código de longitud n sobre el alfabeto $A = \{a_1, \ldots, a_m\}$ con distancia mínima d. Entonces,*

$$|C| \sum_{i=0}^{\left[\frac{d-1}{2}\right]} \binom{n}{i} (m-1)^i \leq m^n.$$

Para probarlo simplemente hay que darse cuenta que

$$\overline{B}(\mathbf{c}, \left[\frac{d-1}{2}\right]) \cap \overline{B}(\mathbf{c}', \left[\frac{d-1}{2}\right]) = \emptyset, \quad \forall \mathbf{c}, \mathbf{c}' \in C \text{ tales que } \mathbf{c} \neq \mathbf{c}'$$

y usar el Lema 3.2.

Relacionamos ahora el número de errores que puede detectar y corregir un código de bloque de longitud n con su distancia mínima. Se puede demostrar que:

Proposición 3.4 *Sea $C \subseteq T_n$ un código de longitud n sobre el alfabeto A con distancia mínima d. Entonces*

1. *C puede detectar hasta $d-1$ errores.*
2. *C puede corregir hasta $\left[\frac{d-1}{2}\right]$ errores.*

Ejemplo Consideramos el código $C \subseteq \mathbb{F}_2^7$ cuyas palabras son

$$\begin{aligned} C \quad = \{&0000000, 0110100, 0011010, 0001101, \\ &1000110, 1001011, 1011100, 0010111, \\ &1010001, 1110010, 0111001, 1111111, \\ &0101110, 1101000, 0100011, 1100101\}. \end{aligned}$$

Este código C tiene distancia mínima 3, ya que haciendo las distancias de pares de palabras distintas de C, se observa que el valor mínimo es este valor. Si le aplicamos ahora la Proposición 3.4, deducimos que puede detectar hasta dos errores y corregir uno. Por ejemplo, si recibimos la palabra $\mathbf{x} = 1000111$ vemos que no está en el código, por tanto se ha(n) producido error(es) en la transmisión. Pero nos fijamos que 1000110 sí pertenece al código y difiere solo en un dígito con $\mathbf{x}$, luego podemos deducir que es la que se ha emitido cuando hemos recibido $\mathbf{x}$, ya que este código corrije un error. De hecho, si calculamos la distancia que hay entre $\mathbf{x}$ y las diferentes palabras de C, vemos que la única palabra de C que dista 1 de $\mathbf{x}$ es 1000110. Más aún, C alcanza la cota de Hamming ya que

$$|C| \sum_{i=0}^{\left[\frac{d-1}{2}\right]} \binom{n}{i} (m-1)^i = 2^4 \sum_{i=0}^{1} \binom{7}{i} 1^1 = 2^4 2^3 = 2^7.$$

Estudiamos los códigos que alcanzan la Cota de Hamming en el siguiente apartado y veremos qué significa alcanzar esta cota en términos de poder determinar la palabra emitida cuando recibimos una palabra de $\mathbb{F}_2^7$.

4 Códigos perfectos

Un código $C \subseteq T_n$ con distancia mínima d que alcance la Cota de Hamming se dice que es un **código perfecto**. Esto es, si $C \subseteq T_n$ es un código de bloque sobre el alfabeto $A = \{a_1, \ldots, a_m\}$ de distancia mínima d, entonces se dice que C es perfecto, si cumple $|C| \sum_{i=0}^{\left[\frac{d-1}{2}\right]} \binom{n}{i} (m-1)^i = m^n$.

En el último ejemplo del apartado anterior, hemos comentado que el código binario

$$\begin{aligned} C \ = \{&0000000, 0110100, 0011010, 0001101, \\ &1000110, 1001011, 1011100, 0010111, \\ &1010001, 1110010, 0111001, 1111111, \\ &0101110, 1101000, 0100011, 1100101\}. \end{aligned}$$

era un código perfecto y habíamos visto que si recibiamos la palabra $\mathbf{x} = 1000111$, podíamos conocer la palabra emitida que le correspondía. En la siguiente proposición vemos que esta característica la tienen todos los códigos perfectos:

Proposición 4.1 *Sea $C \subseteq T_n$ un código de bloque de longitud n sobre un alfabeto A de distancia mínima d. Entonces, C es perfecto si y, solo si, $\dot{\bigcup}_{\boldsymbol{c} \in C} \overline{B}(\boldsymbol{c}, \left[\frac{d-1}{2}\right]) = T_n$.*

La interpretación del resultado anterior es la misma que la del último ejemplo del apartado anterior: cada elemento de T_n se encuentra en una única bola cerrada de centro $\mathbf{c}$ y radio $\left[\frac{d-1}{2}\right]$ y, como las bolas son disjuntas, esto implica que cada elemento de T_n admite decodificación única, que es precisamente el centro $\mathbf{c}$ de la bola a la que pertenece el elemento de T_n que hemos considerado.

En el siguiente resultado vemos que la distancia mínima de un código perfecto no puede tomar cualquier valor:

Proposición 4.2 *Si $C \subseteq T_n$ es un código perfecto, su distancia mínima es un número impar.*

Como consecuencia de la Proposición anterior, podemos deducir la no existencia de códigos perfectos con distancia mínima un número par.

Ejemplo Consideramos de nuevo el código binario de longitud 7 cuyas palabras son:

$$\begin{aligned} C \ = \{&0000000, 0001101, 1111111, 1110010,\\ &1101000, 1000110, 0010111, 0111001,\\ &0110100, 0100011, 1001011, 1011100,\\ &0011010, 1010001, 1100101, 0101110\} \end{aligned}$$

Como ya hemos comentado, este código tiene distancia mínima 3. Además, sabemos que si $\mathbf{c}$ y $\mathbf{c}'$ son dos palabras distintas de C, se tiene que $\overline{B}(\mathbf{c},1) \cap \overline{B}(\mathbf{c}',1) = \emptyset$ y podemos calcular para cada $\mathbf{c} \in C$ el cardinal de $\overline{B}(\mathbf{c},1)$ que es 8. Así que C es un código perfecto ya que

$$\dot{\bigcup}_{\mathbf{c}\in C} \overline{B}(\mathbf{c},1) = \mathbb{F}_2^7.$$

Esto implica que cada palabra de T_n admite, por tanto, decodificación única.

En el Tema 3 estudiaremos una familia de códigos perfectos: los códigos de Hamming, que son códigos de bloque sobre el cuerpo finito $\mathbb{F}_q$ de distancia mínima 3 y longitud $\frac{q^r-1}{q-1}$, siendo r un número natural y q potencia de un primo. De hecho, el código binario del ejemplo anterior va a ser un código de tipo Hamming.

5 Códigos equivalentes

Cuando consideramos dos códigos de bloque C_1 y C_2 con la misma longitud n sobre el mismo alfabeto A nos interesará no solo que C_1 y C_2 sean diferentes como subconjuntos de T_n, sino que también nos preocuparemos de que las capacidades correctoras de ambos u otras características intrínsecas a ellos sean diferentes. Por ello, vamos a introducir la definición de códigos equivalentes que va a recoger esta idea. Necesitamos primero introducir dos tipos de aplicaciones de T_n en sí mismo:

1. Sean $i, j \in \{1, \ldots, n\}$ con $i < j$. Se define

$$\begin{array}{rccc} \phi_{ij}: & T_n & \to & T_n \\ & x_1 \ldots x_n & \mapsto & x_1 \ldots x_{i-1} x_j x_{i+1} \ldots x_{j-1} x_i x_{j+1} \ldots x_n. \end{array}$$

   Es decir, ϕ_{ij} intercambia las letras que aparecen en las posiciones i y j de cada palabra $\mathbf{x} \in T_n$.

2. Dada una permutación $g : A \to A$, esto es, g es una aplicación biyectiva de A en sí mismo, y un índice $k \in \{1, \ldots, n\}$, se define

$$\begin{array}{rccc} \psi_{g,k}: & T_n & \to & T_n \\ & x_1 \ldots x_n & \mapsto & x_1 \ldots x_{k-1} g(x_k) x_{k+1} \ldots x_n. \end{array}$$

   Esto es, $\psi_{g,k}$ lo que hace es aplicar la permutación g a las letras que aparecen en la posición i-ésima.

Definición Sea A un alfabeto y $C_1, C_2 \subseteq T_n$. Se dice que C_2 es **equivalente** a C_1, si se puede obtener C_2 como la imagen de C_1 mediante una aplicación h, que es la composición de un número finito de aplicaciones del tipo ϕ_{ij} y $\psi_{g,k}$ para $g \in \Sigma_A$ y $i, j, k \in \{1, \ldots, n\}$.

Observamos que al ser tanto las aplicaciones del tipo ϕ_{ij} como $\psi_{g,k}$ biyectivas con inversa del mismo tipo es evidente que si C_2 es equivalente a C_1, entonces C_1 es equivalente a C_2. Por tanto, diremos simplemente que C_1 y C_2 son equivalentes. Además, si consideramos $\mathcal{T}_n$ el conjunto de los códigos de longitud n sobre el alfabeto A, se tiene que el ser equivalentes es una relación de equivalencia en $\mathcal{T}_n$, que denotamos por $\mathcal{R}$. Podemos calcular el conjunto cociente $\mathcal{T}_n/\mathcal{R}$ y la clase de un código $C \in \mathcal{T}_n$ vendrá dada por:

$$[C] = \{C_2 \mid C \mathcal{R} C_2\}.$$

Una propiedad interesante de los códigos equivalentes es la que relaciona las distancias mínimas de los códigos equivalentes y que la resumimos en la siguiente Proposición:

Proposición 5.1 *Sea $C_1 \subseteq T_n$ un código de longitud n sobre el alfabeto A y distancia mínima d. Si $C_2 \subseteq T_n$ es un código equivalente a C_1, entonces la distancia mínima de C_2 es también d.*

Para demostrar la Proposición anterior basta fijarse que si elegimos $\mathbf{x}, \mathbf{y} \in T_n$, se cumple

1. $d(\mathbf{x}, \mathbf{y}) = d(\phi_{ij}(\mathbf{x}), \phi_{ij}(\mathbf{y}))$
2. $d(\mathbf{x}, \mathbf{y}) = d(\psi_{g,k}(\mathbf{x}), \psi_{g,k}(\mathbf{y}))$

Por otro lado, si nos fijamos en la Proposición anterior, podemos deducir que una condición necesaria para que dos códigos de bloque de longitud n con alfabeto A sean equivalentes es que ambos tengan la misma distancia mínima.

Por otro lado, se puede demostrar la siguiente proposición:

Proposición 5.2 *Sea $C_1 \subseteq T_n$ un código de longitud n y $\boldsymbol{u} \in T_n$. Entonces, existe $C_2 \subseteq T_n$ equivalente a C_1 tal que $\boldsymbol{u} \in C_2$.*

Como consecuencia del resultado anterior, deducimos que

Corolario 5.3 *Sea $C_1 \subseteq \mathbb{F}_q^n$ un código de longitud n. Entonces, existe $C_2 \subseteq \mathbb{F}_q^n$ equivalente a C_1 tal que $00\ldots00 \in C_2$.*

Esto será interesante cuando deseemos construir un código de bloque de longitud n sobre $\mathbb{F}_q$ que contenga al $0\ldots0$ y sea equivalente a uno dado $C_1 \subseteq \mathbb{F}_q^n$, que no contiene al $0\ldots0$.

Tema 3

Códigos Lineales

1 Definición y primeras propiedades

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

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

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

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

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

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

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

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

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

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

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

Ejemplo El código binario

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

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

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

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

2 Matriz generadora de un código lineal

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

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

Ejemplo El (7,4)-código lineal

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

tiene por matriz generadora a

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

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

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

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

Ejemplo El (7,4)-código lineal

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

tiene por matriz generadora a

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

que no está dada en forma estándar.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Empleando las propiedades de los subespacios vectoriales, se demuestra:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

4 Codificación y decodificación de un código lineal

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

Ilustramos el proceso de codificación con el siguiente ejemplo:

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

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

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

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

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

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

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

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

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

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

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

Describimos ahora ambos métodos de decodificación:

4.1 Método de decodificación basado en los líderes

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

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

Observamos que

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

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

4.2 Método de decodificación mediante síndromes

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

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

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

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

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

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

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

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

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

5 Ejemplo de códigos lineales: Códigos de Hamming

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

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

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

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

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

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

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

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

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

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

Tema 4

Códigos cíclicos

1 Definición y construcción de códigos cíclicos

Sea $C \subseteq \mathbb{F}_q^n$ un (n,k)-código lineal. Se dice que C es **cíclico** si se satisface la siguiente propiedad:

$$\forall c_0 \dots c_{n-1} \in C,\ c_{n-1}c_0 \dots c_{n-2} \in C.$$

Observamos que si C es un código cíclico, entonces dada $c_0 \dots c_{n-1} \in C$, se tiene que las palabras $c_{n-1}c_0 \dots c_{n-2}$, $c_{n-2}c_{n-1}c_0 \dots c_{n-3}$, $\dots$, y $c_1 \dots c_{n-1}c_0$ están también en C.

Por otro lado, también podemos caracterizar los códigos cíclicos fijandonos solamente en lo que sucede en una base del código, tal y como se indica en el siguiente resultado:

Proposición 1.1 (Caracterización de los códigos cíclicos) *Sea $C \subseteq \mathbb{F}_q^n$ un (n,k)-código lineal con base $\mathcal{B} = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_k\}$. Entonces, C es cíclico si y, solo si, para todo $\boldsymbol{x}_i \in \mathcal{B}$, con $i = 1, \dots, k$, se tiene $x_{in-1}x_{i0} \dots x_{in-2} \in C$, siendo $\boldsymbol{x}_i = x_{i0} \dots x_{in-2}x_{in-1}$.*

Si dada una palabra $\mathbf{x} = x_0 \dots x_{n-1} \in \mathbb{F}_q^n$ llamamos a $x_{n-1}x_0 \dots x_{n-2}$ la **traslación cíclica** de $\mathbf{x}$, la proposición anterior nos indica que un código lineal es cíclico si y, solo si, la traslación cíclica de las palabras de una base están también en C. Esta caracterización nos simplificará el estudio de si un código es cíclico, cuando conozcamos una base del mismo.

Ejemplo Sea $C \subseteq \mathbb{F}_2^7$ el código lineal cuya matriz generadora viene dada por:

$G = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$. Entonces, una base de C es

$$\mathcal{B} = \{1101000, 0110100, 0011010, 0001101\}$$

y para las tres primeras palabras de esta base se cumple que sus traslaciones cíclicas son otra palabra de la misma base, por lo que también son palabras de C. Por tanto, solo nos queda estudiar lo que sucede con la traslación cíclica de la última palabra de la base, que es 0001101. Ahora, como estamos trabajando en $\mathbb{F}_2$, se cumple que

$$1000110 = 1101000 + 0110100 + 0011010,$$

así que la traslación cíclica de la última palabra de la base es también otra palabra de C. Por consiguiente, aplicando la Proposición 1.1, podemos afirmar que C es un código cíclico.

Además de la caracterización que hemos visto de los códigos cíclicos estudiando únicamente lo que sucede en un base, vamos a encontrar otra fijandonos en una estructura subyacente de los códigos cíclicos. Para ello, necesitamos recordar cómo se contruye el anillo cociente $\mathbb{F}_q[x]/(x^n - 1)$, donde $\mathbb{F}_q[x]$ el anillo de los polinomios en la variable x, y se puede relacionar con $\mathbb{F}_q^n$. En $\mathbb{F}_q[x]$ definimos la relación de equivalencia

$$\forall f(x), g(x) \in \mathbb{F}_q[x], \quad f(x)\mathfrak{R}g(x) \Leftrightarrow (x^n - 1)|f(x) - g(x),$$

esto es, $f(x) - g(x)$ es múltiplo de $x^n - 1$. Entonces el conjunto cociente, denotado por $\mathbb{F}_q[x]/(x^n - 1)$, está definido por

$$\mathbb{F}_q[x]/(x^n - 1) = \{\overline{f(x)} \mid f(x) \in \mathbb{F}_q[x]\}$$

y en cada clase de equivalencia $\overline{f(x)}$ podemos elegir como representante el polinomio de grado a lo sumo $n-1$ que esté en ella, que será el resto de dividir $f(x)$ por $x^n - 1$.

Por tanto,

$$\mathbb{F}_q[x]/(x^n - 1) = \{\overline{a_0 + a_1x + \cdots + a_{n-1}x^{n-1}} \,|\, a_i \in \mathbb{F}_q,\ i = 0, \ldots, n-1\}.$$

Observamos que hay tantas clases de equivalencia como polinomios de grado menor que n en $\mathbb{F}_q[x]$.

Podemos establecer un isomorfismo de espacios vectoriales entre $\mathbb{F}_q^n$ y $\mathbb{F}_q[x]/(x^n - 1)$ mediante $\psi(a_0 \ldots a_{n-1}) = \overline{a_0 + a_1x + \cdots + a_{n-1}x^{n-1}}$.

Si $C \subseteq \mathbb{F}_q^n$ es un código, denotamos por $C(x)$ a $\psi(C)$, esto es,

$$C(x) = \{\overline{a_0 + a_1x + \cdots + a_{n-1}x^{n-1}} \in \mathbb{F}_q[x]/(x^n - 1)|a_0a_1 \ldots a_{n-1} \in C\}.$$

Observamos que si C es un código cíclico, entonces dada $c_0 \dots c_{n-1} \in C$, se tiene que $c_{n-1}c_0 \dots c_{n-2}$, $c_{n-2}c_{n-1}c_0 \dots c_{n-3}, \dots, c_1 \dots c_{n-1}c_0 \in C$ y esto implica que si $\overline{\mathbf{c}(x)} = \overline{\sum_{i=0}^{n-1} c_i x^i}$, se sigue que $x\overline{\mathbf{c}(x)} = \overline{c_{n-1} + \sum_{i=0}^{n-2} c_i x^{i+1}} \in C(x)$ y en general $x^k\overline{c(x)} = \overline{\sum_{i=1}^{k} c_{n-i}x^{k-i} + \sum_{i=0}^{n-k-1} c_i x^{i+k}} \in C(x)$ para $k < n$. Utilizamos esta propiedad de los códigos cíclicos para caracterizarlos:

Proposición 1.2 (Caracterización de los códigos cíclicos) *Sea $C \subseteq \mathbb{F}_q^n$ un código lineal. Entonces, C es cíclico si y, solo si, $C(x)$ es un ideal de $\mathbb{F}_q[x]/(x^n - 1)$.*

La importancia de la proposición anterior quedará de manifiesto en el siguiente apartado, donde explotaremos esta característica de los códigos cíclicos.

2 Polinomio generador y matriz generadora de un código cíclico

Como hemos indicado, usando Teoría de Anillos, vamos a poder conocer más sobre la estructura de $C(x)$, cuando C es un código cíclico:

Proposición 2.1 *Sea $C \subseteq \mathbb{F}_q^n$ un código cíclico. Entonces, existe un único polinomio mónico $g(x)$ de grado mínimo tal que*

$$C(x) = (\overline{g(x)}) = \{\overline{t(x)g(x)} \in \mathbb{F}_q[x]/(x^n - 1) \mid t(x) \in \mathbb{F}_q[x]\}.$$

Además, $g(x)$ es un factor de $x^n - 1$ en $\mathbb{F}_q[x]$.

Cuando $C \subseteq \mathbb{F}_q^n$ es un código cíclico, al polinomio mónico $g(x)$ de grado mínimo tal que $C(x) = (\overline{g(x)})$ se le llama **polinomio generador de** C. Por tanto, es fácil determinar todos los códigos cíclicos de una longitud determinada n sobre $\mathbb{F}_q$: basta con hallar los polinomios mónicos que dividen a $x^n - 1$ y tomar cada uno de ellos como polinomio generador del código cíclico buscado.

Ejemplo Para calcular los códigos cíclicos de longitud 7 en $\mathbb{F}_2$, debemos determinar los factores irreducibles sobre $\mathbb{F}_2$ de $x^7 - 1$. Ahora,

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

Entonces, hay 2^3 códigos distintos $C_i(x) = (\overline{g_i(x)})$, con $i = 1, \dots, 8$, donde

$$\begin{array}{ll} g_1(x) = 1, & g_2(x) = x + 1 \\ g_3(x) = x^3 + x + 1 & g_4(x) = x^3 + x^2 + 1 \\ g_5(x) = x^4 + x^2 + x + 1 & g_6(x) = x^4 + x^3 + x^2 + 1 \\ g_7(x) = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1 & g_8(x) = x^7 - 1 \end{array}$$

Observamos que $C_1(x) = \mathbb{F}_2[x]/(x^7 - 1)$ y, por tanto, $C_1 = \mathbb{F}_2^7$ y $C_8(x) = 0$, luego $C_8 = \{0000000\}$.

Este polinomio generador de un código cíclico nos sirve para determinar una matriz generadora del mismo, tal y como nos indica el siguiente resultado:

Proposición 2.2 (Matriz generadora) *Sea $C \subset \mathbb{F}_q^n$ un código cíclico con polinomio generador $g(x) = \sum_{i=0}^{n-k} g_i x^i$ de grado $n-k$. Entonces, C es un código de dimensión k y una matriz generadora es*

$$\begin{pmatrix} g_0 & g_1 & \dots & g_{n-k} & 0 & 0 & \dots & 0 \\ 0 & g_0 & g_1 & \dots & g_{n-k} & 0 & \dots & 0 \\ & & & & & & & \\ 0 & \dots & 0 & g_0 & g_1 & g_2 & \dots & g_{n-k} \end{pmatrix}.$$

Ejemplo Si consideramos C_3 el código binario cíclico de longitud 7 con polinomio generador $g_3(x) = x^3 + x + 1$, entonces aplicando la Proposición 2.2 una matriz generadora de C_3 viene dada por

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

Obviamente, la matriz generadora G_3 es de tamaño 4×7 y con rango 4, pues C es dimensión 4.

En cambio, si tomasemos el código cíclico binario C_1 de longitud 7 con polinomio generador $g_1(x) = 1$, entonces de la Proposición 2.2 deducimos que una matriz generadora de C_1 es la matriz identidad I_7.

3 Polinomio de control y matriz de control de un código cíclico

Si C es un código cíclico de longitud n y dimensión k con polinomio generador $g(x)$, que será de grado $n-k$, sabemos que $g(x)$ es un divisor de $x^n - 1$ y que existe $h(x) \in \mathbb{F}_q[x]$, de grado precisamente k, tal que $g(x)h(x) = x^n - 1$. A este polinomio $h(x)$, que también es mónico y que verifica $g(x)h(x) = x^n - 1$, se le denomina **polinomio de control del código cíclico** C.

Justificamos en la siguiente proposición el denominar a $h(x)$ polinomio de control:

Proposición 3.1 ***(Matriz de control)*** *Sea $C \subset \mathbb{F}_q^n$ un código cíclico con polinomio de control $h(x) = \sum_{i=0}^{k} h_i x^i$ de grado k. Entonces, una matriz de control de C es*

$$\begin{pmatrix} h_k & h_{k-1} & \dots & h_0 & 0 & 0 & \dots & 0 \\ 0 & h_k & h_{k-1} & \dots & h_0 & 0 & \dots & 0 \\ & & & & & & & \\ & & & & & & & \\ 0 & \dots & 0 & h_k & h_{k-1} & h_{k-2} & \dots & h_0 \end{pmatrix}.$$

Observamos que si $C \subset \mathbb{F}_q^n$ es un código cíclico con polinomio de control $h(x) = \sum_{i=0}^{k} h_i x^i$, la matriz H que figura en la proposición anterior nos permite deducir que $C^\perp$ es otro código cíclico puesto que una base de él es la formada por las filas de H y esta base verifica la Proposición 1.1. También lo podríamos verificar estudiando si $h_0^{-1} \sum_{i=0}^{k} h_i x^{k-i}$, que lo obtenemos de la expresión de H, es un divisor de $x^n - 1$ y comprobando que este polinomio genera el ideal de $C^\perp(x)$. En cualquier caso, deducimos que un polinomio generador para $C^\perp$ es $h_0^{-1} \sum_{i=0}^{k} h_i x^{k-i}$. Obviamente, $C^\perp$, que tiene dimensión $n - k$ si C es de dimensión k por ser su dual, tiene un polinomio generador de grado k.

Ejemplo Si consideramos C_3 el código binario cíclico de longitud 7 con polinomio generador $g_3(x) = x^3 + x + 1$, sabemos que su polinomio de control viene dado por $h_3(x) = x^4 + x^2 + x + 1$. Entonces, aplicando la Proposición 3.1 una matriz de control de C_3 viene dada por

$$H_3 = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

Vemos que la matriz de control H_3 es de tamaño 3×7 y con rango 3, pues C es de dimensión 4. Además, esta matriz genera a $C_3^\perp$, que es también cíclico, y el polinomio generador de $C_3^\perp$ viene dado por $1 + x^2 + x^3 + x^4$, que es precisamente el polinomio $g_6(x)$ del ejemplo de la sección anterior, en el que se calculaban todos los polinomios generadores de códigos cíclicos binarios de longitud 7. Observamos con este ejemplo que el polinomio de control $h(x)$ de un código cíclico C no tiene que ser necesariamente el polinomio generador de $C^\perp$, aunque $h(x)$ tenga el grado adecuado para generar $C^\perp$, sea mónico y divisor de $x^n - 1$.

4 Codificación y decodificación de un código cíclico

Sea $C \subseteq \mathbb{F}_q^n$ un código cíclico de dimensión k con polinomo generador $g(x)$. Supongamos que queremos transmitir un mensaje que contiene la palabra $\mathbf{a} = a_0 \dots a_{k-1} \in \mathbb{F}_q^k$ y que queremos codificar $\mathbf{a}$ usando el código C. Podemos hacerlo de dos modos:

1. Si $\mathbf{a} = a_0 \dots a_{k-1}$, la codificamos como $(a_0 \dots a_{k-1})G$, donde G es una matriz generadora de C. Observamos que G se obtiene de forma fácil a partir de

$g(x)$, usando la Proposición 2.2. Pero para la palabra de longitud n que tiene el receptor, una vez que la haya decodificado para corregir los errores que se hayan producido en la transmisión, no es inmediato calcular sus coordenadas en la base formada por las palabras que constituyen la matriz generadora G y que nos darían, precisamente, $a_0 \dots a_{k-1}$.

2. Si $\mathbf{a} = a_0 \dots a_{k-1}$, construimos el polimonio $b(x) = \sum_{i=0}^{k-1} a_i x^{n-1-i}$. Dividimos $b(x)$ por $g(x)$ y obtenemos $r(x)$ de grado menor que $n-k$ tal que $b(x) = t(x)g(x)+r(x)$. Entonces, codificamos $\mathbf{a}$ usando $b(x)-r(x)$, que es una palabra del $C(x)$, con la ventaja frente al método anterior de que en las k últimas posiciones de esta palabra van las letras de a y esto facilita los cálculos que debe hacer el receptor, una vez que se han corregido los errores que se hayan podido producir en la transmisión, para llegar a $\mathbf{a}$.

Nos preocupamos ahora por cómo puede corregir el receptor los errores producidos en la transmisión, esto es, del proceso de decodificar la palabra recibida. Al ser también códigos lineales, cuando usamos un código cíclico podemos aplicar cualquiera de los dos métodos generales que se utilizan en la decodificación de palabras codificadas con códigos lineales: el método de los líderes y el método de los síndromes. Pero para los códigos cíclicos existe otro método que explota el que las traslaciones cíclicas de palabras de C estén en C. Es el llamado método de decodificación cíclica que explicamos a continuación.

4.1 Método de decodificación cíclica

Sea $C \subseteq \mathbb{F}_q^n$ un código cíclico cuya matriz de control es H y su distancia mínima es d. Supongamos que hemos recibido una palabra $\mathbf{y} = y_0 \dots y_{n-1} \in \mathbb{F}_q^n$ tal que su síndrome es no nulo, esto es, $\mathbf{y} \notin C$ y queremos decodificarla. Como ya se ha indicado, podríamos emplear, por ejemplo, la decodificación por síndromes, que es un método válido para cualquier código lineal, puesto que disponemos de una matriz de control de C. Así, siguiendo el procedimiento explicado en el tema anterior deberíamos calcular la tabla de síndromes (dando un líder para cada síndrome) y localizar aquel líder que tuviese el mismo síndrome que la palabra recibida, decodificando ésta como la diferencia entre la recibida y el líder. Pero en el caso de emplear códigos cíclicos, el proceso anterior puede simplificarse utilizando el llamado método de decodificación cíclica, que está basado en el hecho de que si una palabra $\mathbf{c} = c_0 \dots c_{n-1} \in C$, entonces también pertenece a C la palabra $\mathbf{c}^{(1)} = c_{n-1}c_0 \dots c_{n-2} \in C$. En esencia, el método consiste en fijar una posición de la palabra (por ejemplo, la última) y calcular los síndromes de palabras líder que tengan en esa posición un valor no nulo (lo que llamaremos tabla reducida de síndromes) y compararlo con el síndrome de la palabra recibida. Si coincide, significa que en las posiciones no nulas de las palabras líder se ha producido error y lo podemos corregir. Si no coincide el síndrome de nuestra palabra con ninguna de los de la tabla reducida, significa que en esas posiciones fijadas no se ha producido error y se repite el proceso con $y_{n-1}y_0 \dots y_{n-2}$, $y_{n-2}y_{n-1}y_0 \dots y_{n-3}$, etc. Al igual que en

el cálculo de la tabla de síndromes, siempre empezaremos calculando los síndromes de palabras líder con menor peso y que éste sea a lo sumo $\left[\frac{d-1}{2}\right]$, que son las clases para las que se puede garantizar decodificación única.

Algoritmo del método de decodificación cíclica

Dada $\mathbf{y} = y_0 \dots y_{n-1} \in \mathbb{F}_q^n$ e $i = 1, \dots, n-1$, denotaremos por

$$\mathbf{y}^{(i)} = y_{n-i} y_{n-i+1} \cdots y_{n-1} y_0 \cdots y_{n-i-1}.$$

Así, $\mathbf{y}^{(1)} = y_{n-1} y_0 \dots y_{n-2}$, $\mathbf{y}^{(2)} = y_{n-2} y_{n-1} y_0 \dots y_{n-3}, \dots$. A $\mathbf{y}$ se la denotará por $\mathbf{y}^{(0)}$.

Para describir este método, la componente de la palabra en la que nos fijamos es la última, aunque se puede realizar el proceso fijandose en cualquiera de las componentes.

- Paso 1: Se construye una tabla reducida de síndromes para los líderes con peso menor o igual que $[\frac{d-1}{2}]$ y que tengan última componente no nula.
- Paso 2: Se toma $i = 0$.
- Paso 3: Se calcula $S(\mathbf{y}^{(i)})$. Si $S(\mathbf{y}^{(i)})$ coincide con alguno de los síndromes $S(\mathbf{e})$ de la tabla reducida, significa que se ha producido en $y^{(i)}$ el error $\mathbf{e}$, por lo que $x^{(i)} = (\mathbf{y}^{(i)} - \mathbf{e}) = x_{n-i} x_{n-i+1} \dots x_{n-1} x_0 \dots x_{n-i-1}$ es la i-ésima traslación de la palabra emitida y, por tanto, $\mathbf{x} = x_0 \dots x_{n-1}$, finalizando el proceso. En caso contrario, se va al paso 4.
- Paso 4: Si $S(\mathbf{y}^{(i)})$ no coincide con ninguno de los $S(\mathbf{e})$ de la tabla reducida significa que la última posición de $y^{(i)}$ es correcta y se va al paso 5.
- Paso 5: Se aumenta en una unidad el índice i, verificando que el nuevo índice sea, a lo sumo, $n-1$ y se reitera el paso 3. Si el nuevo índice es n, se va al Paso 6.
- Paso 6: Si no hemos sido capaces de calcular $\mathbf{x}$ mediante el proceso anterior, significa que la palabra recibida tiene más errores que los que es capaz de corregir C, por lo que ampliaremos la tabla reducida con líderes de peso mayor que $[\frac{d-1}{2}]$ y que tengan la última componente no nula y síndrome diferente a los que figuran en la tabla reducida. Entonces, reiteraremos el procedimiento de los pasos anteriores con esta nueva tabla de síndromes hasta lograr la decodificación, que podrá ser no única.

Ejemplo Se considera el código cíclico binario de longitud 7, distancia mínima 3 y cuya matriz de control viene dada por

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

Supongamos que hemos recibido la palabra $\mathbf{y} = 0011110$. Entonces, $S(\mathbf{y}) = 111$, por lo que sabemos que $\mathbf{y} \notin C$.

1. Calculamos la tabla reducida de síndromes:

Líder	**Síndrome**
0000001	001

2. $S(\mathbf{y}^{(0)}) = 111 \neq 001 = S(0000001)$, luego la última componente de $\mathbf{y}$ es correcta.
3. $S(\mathbf{y}^{(1)}) = 011 \neq 001 = S(0000001)$, luego la última componente de $\mathbf{y}^{(1)}$ es correcta.
4. $S(\mathbf{y}^{(2)}) = 001 = S(0000001)$, luego la última componente de $\mathbf{y}^{(2)}$ debe corregirse para obtener $\mathbf{x}^{(2)} = \mathbf{y}^{(2)} - 0000001 = 1000110$ y entonces $\mathbf{x} = 0011010$.

5 Ejemplo de códigos cíclicos: Códigos BCH

Por Teoría de Anillos, si n y q son coprimos entre sí, sabemos que $x^n - 1$ se puede expresar como producto de polinomios mónicos $f_i(x)$ irreducibles sobre $\mathbb{F}_q$, esto es, $x^n - 1 = \prod f_i(x)$ y que si α_i es una raíz de f_i, entonces $t(\alpha_i) = 0$ si y, solo si, $t(x) = f_i(x)a(x)$ para algún polinomio $a(x)$. Esto nos sirve para caracterizar a los códigos cíclicos, usando las raíces de su polinomio generador de la forma siguiente:

Proposición 5.1 *Sea $C \subset \mathbb{F}_q^n$ un código cíclico con polinomio generador $g(x) = \prod_{i=1}^{s} f_i(x)$, siendo f_i polinomio irreducible sobre $\mathbb{F}_q$, y sea α_i una raíz de $f_i(x)$ en una extensión apropiada de $\mathbb{F}_q$. Entonces,*

$$C(x) = \{\overline{t(x)} \in \mathbb{F}_q[x]/(x^n - 1) \,|\, t(\alpha_1) = \cdots = t(\alpha_s) = 0\}.$$

A las raíces del polimonio generador de un código cíclico se les llaman **ceros** del código. Observamos que los ceros siempre son raíces n-ésimas de la unidad. Estos ceros nos sirven para dar otra construcción de los códigos cíclicos: si partimos de un subconjunto de raíces n-ésimas de la unidad, para cada una de ellas le localizamos el polinomio irreducible del que es raíz y construimos el código C cuyo polinomio generador sea el mínimo común múltiplo de de estos polinomios irreducibles. Además, si $\alpha_1, \ldots, \alpha_r$ son los ceros del código cíclico que buscamos, la matriz

$$H_1 = \begin{pmatrix} 1 & \alpha_1 & \ldots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \ldots & \alpha_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \alpha_r & \ldots & \alpha_r^{n-1} \end{pmatrix}$$

funciona como una "matriz de control" en el sentido que $c \in C$ si y, solo si, $cH_1^t = 0$.

Un tipo de códigos que se contruyen de esta manera son los códigos BCH: Sean n, $q = p^r$ dos números naturales coprimos entre sí. Sea m el orden multiplicativo de q módulo n, es decir, el número natural más pequeño tal que $q^m \equiv 1 \bmod n$. Sea δ un número natural tal que $2 \leq \delta \leq n$ y sea $\alpha \in \mathbb{F}_{q^m}$ una raíz primitiva n-ésima de la unidad. Se llama **código BCH en sentido estricto sobre $\mathbb{F}_q$ de longitud n y distancia mínima prevista** δ al código cíclico cuyo polinomio generador tiene por raíces a $\{\alpha, \alpha^2, \ldots, \alpha^{\delta-1}\}$. Se puede demostrar que la distancia mínima del código así construido es, al menos, δ.

Ejemplo Consideramos $q = 2$, $n = 2^3 - 1 = 7$ y $\delta = 3$. Entonces, debemos considerar un código cíclico cuyo polinomio generador tenga por raíces a α, y α^2, siendo α una raíz primitiva séptima de la unidad. Ahora, si $f(x) \in \mathbb{F}_2[x]$, entonces $f(\alpha^2) = f(\alpha)^2$, luego α y α^2 tienen el mismo polinomio irreducible. Además, $m = 3$, luego el polinomio irreducible que tiene a α como raíz es de grado 3. Pero $x^7 - 1 = (x-1)(x^3 + x^2 + 1)(x^3 + x + 1)$ en $\mathbb{F}_2[x]$, así que α es raíz de $(x^3 + x^2 + 1)$ ó de $(x^3 + x^2 + 1)$. Supongamos que α es raíz de $(x^3 + x^2 + 1)$. Entonces, este polinomio será el polinomio generador del código BCH buscado. Si aumentaramos en una unidad la distancia prevista, esto es, pidieramos $\delta = 4$, entonces para hallar el código BCH C_1 que tiene también por cero a α, raíz primitiva séptima de la unidad con polinomio irreducible (x^3+x^2+1), debemos incluir en el polinomio generador de C_1 al polinomio irreducible que tenga a α^3, como raíz. Pero el polinomio irreducible de α^3 es, en este caso, $(x^3 + x + 1)$, con lo que C_1 tendrá por polinomio generador a $(x^3 + x^2 + 1)(x^3 + x + 1)$.

Anexo: CUERPOS FINITOS

A.1. Algunas estructuras algebraicas interesantes

En este anexo introduciremos algunas estructuras algebraicas básicas, constituidas por ciertos conjuntos con una o más leyes de composición internas que cumplen unos determinados axiomas.

Definición Un **grupo** es un par $(G, +)$, donde $+$ es una ley de composición interna en G (es decir, una aplicación de $G \times G$ en G), que cumple:

1. $(a + b) + c = a + (b + c)$ $\forall a, b, c \in G$ (propiedad asociativa).
2. $\exists 0_G \in G$ tal que, $a + 0_G = 0_G + a = a$ $\forall a \in G$ (existencia de elemento neutro).
3. $\forall a \in G$ $\exists a' \in G$ tal que $a + a' = a' + a = 0_G$ (existencia de elemento opuesto).

Si además se cumple que $a + b = b + a$ $\forall a, b \in G$ (propiedad conmutativa), el grupo se dice **conmutativo**, o también **abeliano**.

Hay un único elemento en el grupo que cumple la propiedad indicada en el segundo axioma, es decir, el elemento neutro es único.

También, para cada $a \in G$, el elemento a' indicado en el tercer axioma es único. Al opuesto a' del elemento a se le suele denotar por $-a$.

Es indiferente el símbolo que se use para denotar la ley de composición interna; de hecho, se podría haber utilizado cualquier otro, como $\star, \circ$, etc. Cuando el símbolo usado es $+$, se dice que la notación es aditiva. También es habitual denotarlo por $\cdot$ (o, más brevemente, simplemente por yuxtaposición), en cuyo caso se dice que la notación es multiplicativa. En este caso, al elemento neutro se le suele denotar por 1_G (o simplemente por 1 si no hay lugar a confusión), y al elemento a' que verifica $aa' = a'a = 1$ se le suele denotar por a^{-1}, y se le llama el **inverso de** a.

Ejemplos A.1.1

1. El grupo más simple de todos es el grupo $\{0\}$ con un sólo elemento y con la

única operación posible definida por $0 + 0 = 0$, llamado grupo trivial; este grupo es abeliano.

2. El conjunto $\mathbb{Z}$ de los números enteros, con la operación habitual de suma de enteros, es un grupo abeliano. En cambio, $\mathbb{Z}$ con la operación producto habitual no es un grupo, ya que los únicos elementos que tienen inverso son 1 y -1.

3. El conjunto $\mathbb{Q}$ de los números racionales, con la operación habitual de suma, es un grupo abeliano. Con la operación producto no es grupo, ya que 0 no tiene inverso, pero si omitimos este elemento, es obvio que los números racionales no nulos con la multiplicación sí forman un grupo, el cual es abeliano.

4. El conjunto $\mathbb{R}$ de los números reales, con la operación habitual de suma, es un grupo abeliano.

5. El conjunto $\mathbb{C}$ de los números complejos, con la operación habitual de suma, es un grupo abeliano.

6. El conjunto $G = \{a, b, c, d, e, f\}$, con la operación definida por

*	a	b	c	d	e	f
a	a	b	c	d	e	f
b	b	c	a	f	d	e
c	c	a	b	e	f	d
d	d	e	f	a	b	c
e	e	f	d	c	a	b
f	f	d	e	b	c	a

es un grupo. No es abeliano ya que, por ejemplo, $b * d = f$, pero $d * b = e$.

Definición Un **subgrupo** de un grupo G es un subconjunto H de G que cumple:

1. $a + b \in H \ \forall a, b \in H$.
2. $0_G \in H$.
3. $-a \in H \ \forall a \in H$.

Un subgrupo H de un grupo G se dice **normal**, y se denota $H \trianglelefteq G$, si cumple

$$-a + b + a \in H, \ \forall a \in G, \forall b \in H.$$

Obviamente, si G es abeliano, todo subgrupo suyo es normal.

Si H es un subgrupo de G, la relación definida por

$$a \equiv b, \text{ si } a - b \in H$$

es de equivalencia. Dado un $a \in G$, la clase de equivalencia representada por a, es igual a $\{h+a \mid h \in H\}$. A este conjunto se le suele denotar por $H+a$, y a este tipo de conjuntos se les llama **coclases a derecha**.

Análogamente, la relación definida por

$$a \equiv b \text{ si } -a+b \in H$$

es de equivalencia. Dado un $a \in G$, la clase de equivalencia representada por a, es igual a $\{a+h \mid h \in H\}$. Este conjunto se denota por $a+H$, y a estos conjuntos se les llama **coclases a izquierda**.

Si H es subgrupo normal de G, ambas relaciones son iguales, y las coclases a derecha e izquierda son iguales. En este caso, el conjunto cociente formado por las clases de equivalencia, es decir, por las coclases, se suele representar por G/H. En este conjunto cociente la operación binaria siguiente está bien definida:

$$(a+H)+(b+H)=(a+b)+H.$$

Por bien definida queremos decir que el resultado no depende de los representantes elegidos en las coclases que forman los sumandos, es decir, que si $a+H=a'+H$ y $b+H=b'+H$, entonces $(a+b)+H=(a'+b')+H$.

Con esta operación se tiene lo siguiente:

Proposición A.1.1 *$(G/H,+)$ es un grupo. Además, si G es abeliano, G/H también lo es.*

Al grupo anteriormente introducido se le llama **grupo cociente del grupo G entre el subgrupo normal H.**

Cuando no hay ambigüedad respecto al subgrupo H que se está considerando, se suele denotar la coclase $a+H$ simplemente por $\overline{a}$. Con esta notación, la operación del grupo cociente es $\overline{a}+\overline{b}=\overline{a+b}$.

Ejemplo A.1.2 Si n es un número entero, el conjunto

$$n\mathbb{Z}=\{nm \mid m \in \mathbb{Z}\}$$

es un subgrupo del grupo aditivo $(\mathbb{Z},+)$ de los enteros. Si $n>0$, las coclases de $\mathbb{Z}/n\mathbb{Z}$ son, sin repeticiones ni que falte ninguna, $\{0+n\mathbb{Z}, 1+n\mathbb{Z}, \ldots, (n-1)+n\mathbb{Z}\}$. El resultado de sumar dos coclases $\overline{a}$ y $\overline{b}$ en este cociente es $\overline{a+b}$. Para expresar el resultado en la forma indicada anteriormente, se pone $\overline{a+b}=\overline{r}$, donde r es el resto de la división de $a+b$ entre n.

Por ejemplo, en $\mathbb{Z}/7\mathbb{Z}$, se tiene que $\overline{3}+\overline{6}=\overline{9}=\overline{2}$. La tabla completa de la suma en

$\mathbb{Z}/7\mathbb{Z}$ es

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{4}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{5}$	$\overline{5}$	$\overline{6}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{6}$	$\overline{6}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$

Otra estructura algebraica importante, en la que intervienen dos leyes de composición, es la de anillo:

Definición Un **anillo** es una terna $(A, +, \cdot)$, donde $+$ y $\cdot$ son leyes de composición internas en A, que cumple:

1. $(A, +)$ es grupo abeliano.
2. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ $\forall a, b, c \in A$ (propiedad asociativa).
3. $a \cdot (b + c) = a \cdot b + a \cdot c$ y $(b + c) \cdot a = b \cdot a + c \cdot a$ $\forall a, b, c \in A$ (propiedad distributiva).

0_A denotará al elemento neutro de la suma, y $-a$ al elemento opuesto de a. El anillo se dice **conmutativo** si $\forall a, b \in A,\ a \cdot b = b \cdot a$, y unitario si existe un elemento $1_A \in A$ que cumple $1_A \cdot a = a \cdot 1_A = a$ $\forall a \in A$.

Ejemplos A.1.3

1. El anillo $\{0\}$, con las únicas operaciones posibles $0 + 0 = 0$ y $0 \cdot 0 = 0$, es un anillo conmutativo y unitario, llamado **anillo nulo o anillo trivial**. En él, los elementos neutros de la suma y del producto coinciden (se puede ver que es el único caso en el que pasa esto).
2. Los conjuntos $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, con sus operaciones habituales de suma y producto, son anillos conmutativos y unitarios.
3. El conjunto
$$\{a + bI \mid a, b \in \mathbb{Z}\},$$
donde I es la unidad imaginaria, es un anillo conmutativo y unitario con sus operaciones habituales como números complejos. Se le llama **anillo de los enteros gaussianos.**
4. Si A es un anillo, el conjunto
$$A[x] = \{\sum a_i x^i \mid a_i \in A\ \forall i, a_i = 0 \text{ a partir de un cierto } n_0\},$$

es un anillo con las operaciones $\sum a_i x^i + \sum b_i x^i = \sum (a_i + b_i) x^i$ y $(\sum a_i x^i)(\sum b_i x^i) = (\sum c_i x^i)$, con $c_i = \sum_{j+k=i} a_j b_k$, es un anillo. Si A es conmutativo (respectivamente, unitario), entonces $A[x]$ también lo es.

5. Dado un número natural n, el conjunto $M_n(\mathbb{R})$ de matrices cuadradas de orden n cuyas entradas son números reales, es un anillo unitario con sus operaciones usuales de suma y producto de matrices. No es conmutativo, salvo en el caso $n = 1$.

A veces se suele omitir el punto en el producto, y éste se denota simplemente por yuxtaposición.

Aunque en el último ejemplo hemos mostrado, con propósitos ilustrativos, un anillo no conmutativo, en lo sucesivo supondremos, aunque no se diga explícitamente, que los anillos que consideramos son conmutativos y unitarios.

En cualquier anillo, la propiedad asociativa garantiza que el resultado de hacer $a + \cdots + a$, donde a aparece n veces, no depende de cómo se distribuyan los paréntesis al hacer la suma. A este elemento se le denota por na. Esta definición se puede extender del modo siguiente: $0a = 0_A$ y, si n es un entero negativo, $na = -((-n)a)$.

De forma similar, la asociatividad del producto garantiza que el producto $a \cdots a$, con n factores, no depende de cómo se distribuyan los paréntesis al hacer los productos parciales. A este elemento se le denota por a^n. Se conviene también que $a^0 = 1_A$.

Proposición A.1.2 *1.* $0_A a = 0_A \ \forall a \in A$.

2. $(-a)b = a(-b) = -(ab) \ \forall a, b \in A$.

3. $(-a)(-b) = ab \ \forall a, b \in A$.

4. $a(b - c) = ab - ac \ \forall a, b, c \in A$.

Definición Un elemento a de un anillo A es una **unidad** si es inversible para la multiplicación, es decir, si existe un $b \in A$ tal que $ab = 1$.

Al inverso de un elemento a se le suele denotar por a^{-1}.

Ejemplos A.1.4

1. En el anillo $\mathbb{Z}$ de los enteros los únicos elementos inversibles, es decir, las únicas unidades, son 1 y -1.

2. En el anillo de los enteros gaussianos hay exactamente 4 unidades: $1, -1, I, -I$.

3. En el anillo $\mathbb{Q}$ de los números enteros, las unidades son los elementos no nulos.

Se pueden definir potencias de exponente negativo de elementos inversibles de la siguiente forma: si a es una unidad y n es un entero negativo, $a^n = (a^{-1})^{-n}$.

Veremos ahora los objetos que son el objetivo principal de estudio en este apéndice, los cuerpos:

Definición Un **cuerpo** es un anillo no nulo K en el que todo elemento no nulo es una unidad.

Ejemplos A.1.1 *1. El conjunto $\mathbb{Q}$ de los números racionales, con las operaciones habituales, es un cuerpo.*

2. El conjunto $\mathbb{R}$ de los números reales, con las operaciones habituales, es un cuerpo.

3. El conjunto $\mathbb{C}$ de los números complejos, con las operaciones habituales, es un cuerpo.

Así como en el caso de los grupos la subestructura que nos permitía definir los grupos cocientes era la de subgrupo normal, en el de los anillos, la subestructura que nos permitirá definir los anillos cociente es la de ideal:

Definición Si A es un anillo, un **ideal** de A es un subconjunto $\mathfrak{a}$ de A que cumple:

1. $(\mathfrak{a}, +)$ es subgrupo de $(A, +)$
2. $ax \in \mathfrak{a}\ \forall a \in A, \forall x \in \mathfrak{a}$

Si un ideal $\mathfrak{a}$ de un anillo A contiene al elemento 1_A, entonces $\mathfrak{a} = A$, ya que si $a \in A$, entonces $a = a \cdot 1_A \in \mathfrak{a}$, por definición de ideal.

Ejemplos A.1.5

1. Tanto $\{0_A\}$ como A siempre son ideales de un anillo A. A éstos se les llama **ideales triviales de** A. A los que son distintos de $\{0_A\}$ y de A se les llama **ideales no triviales**, y a los que son distintos de A, **ideales propios**.
2. Si n es un número entero, el conjunto

$$n\mathbb{Z} = \{nm \mid m \in \mathbb{Z}\}$$

es un ideal del anillo $\mathbb{Z}$. De hecho, éstos son todos los ideales del anillo $\mathbb{Z}$ de los enteros.

Si $\mathfrak{a}$ es un ideal de A, entonces $(\mathfrak{a}, +)$ es un subgrupo normal del grupo aditivo, luego podemos considerar la relación de congruencia $x \equiv y$ si $x - y \in \mathfrak{a}$ y, como dijimos antes, el conjunto cociente $A/\mathfrak{a}$, con la operación $(x + \mathfrak{a}) + (y + \mathfrak{a}) = (x + y) + \mathfrak{a}$ es un grupo, llamado el **grupo cociente de A entre el subgrupo $\mathfrak{a}$**. La definición siguiente de producto de clases de equivalencia es consistente y no depende de los representantes elegidos:

$$(x + \mathfrak{a})(y + \mathfrak{a}) = (xy) + \mathfrak{a}.$$

Proposición A.1.3 *Si A es un anillo y $\mathfrak{a}$ es un ideal de A, entonces $(A/\mathfrak{a}, +, \cdot)$ es anillo.*

Al anillo anterior se le llama **anillo cociente de A entre el ideal $\mathfrak{a}$**. Sus elementos neutros para la suma y el producto son $0_A + \mathfrak{a}$ y $1_A + \mathfrak{a}$, respectivamente.

A los $x + \mathfrak{a}$ se les llama también **coclases módulo $\mathfrak{a}$**, o simplemente **coclases**, y se suelen denotar, cuando no haya ambigüedad respecto al ideal que se esté considerando, poniendo simplemente una barra encima de un representante, como $\overline{x}$. Cuando $x + \mathfrak{a} = y + \mathfrak{a}$ (lo cual ocurre si y sólo si $x - y$ está en $\mathfrak{a}$), se dice que x es **congruente con y módulo $\mathfrak{a}$**, lo cual se denota por $x \equiv y \pmod{\mathfrak{a}}$.

Ejemplo A.1.6 Si, dado un número natural n, consideramos el ideal $n\mathbb{Z}$, sabemos por lo visto en los grupos cociente que

$$\mathbb{Z}/n\mathbb{Z} = \{\overline{1}, \overline{2}, \ldots, \overline{n-1}\},$$

y que se tiene la operación suma

$$\overline{r} + \overline{s} = \overline{r + s}.$$

Al ser $n\mathbb{Z}$ un ideal, también se tiene la operación producto

$$\overline{r} \cdot \overline{s} = \overline{r \cdot s}.$$

Veremos ahora un tipo de ideales que serán muy útiles posteriormente para la construcción de cuerpos:

Definición Un ideal $\mathfrak{M}$ de un anillo A es **maximal** si es ideal propio y los únicos ideales que lo contienen son $\mathfrak{M}$ y A.

Ejemplo A.1.7

1. $6\mathbb{Z}$ no es ideal maximal de $\mathbb{Z}$, ya que $6\mathbb{Z} \subset 3\mathbb{Z}$, y $3\mathbb{Z} \neq 6\mathbb{Z}, 3\mathbb{Z} \neq \mathbb{Z}$.
2. $2\mathbb{Z}$ sí es ideal maximal de $\mathbb{Z}$, ya que $1 \notin 2\mathbb{Z}$, luego $2\mathbb{Z} \neq \mathbb{Z}$, y cualquier ideal que contenga estrictamente a $2\mathbb{Z}$ contiene un entero impar, luego también al 1, de donde se deduce que el ideal tiene que ser todo $\mathbb{Z}$.

Se pueden caracterizar los ideales maximales en términos de la estructura del anillo cociente:

Proposición A.1.4 *Un ideal $\mathfrak{M}$ de un anillo A es maximal si y sólo si $A/\mathfrak{M}$ es cuerpo.*

Generalizando lo visto en el segundo ejemplo, se tiene que los ideales maximales de $\mathbb{Z}$ son los de la forma $p\mathbb{Z}$, donde p es un número primo. De ahí se deduce lo siguiente:

Proposición A.1.5 *Un anillo cociente $\mathbb{Z}/n\mathbb{Z}$ es cuerpo si y sólo si n es un número primo.*

La proposición anterior nos proporciona un método de construcción de cuerpos finitos de cardinal un número primo. En la sección siguiente veremos, más en general, cómo construir cuerpos con cardinal una potencia de un primo.

Ejemplo A.1.8 Como 7 es un número primo, el anillo $\mathbb{Z}/7\mathbb{Z}$ es un cuerpo con 7 elementos. Ya vimos en el ejemplo A.1.2 la tabla de la suma en este cuerpo. Veremos ahora la tabla del producto. Para confeccionarla, hay que tener en cuenta que para multiplicar $\overline{i}$ por $\overline{j}$, se hace primero el producto ij, se toma luego el resto r de la división de ij por 7, y el resultado de la multiplicación es $\overline{r}$. Por ejemplo, $\overline{5}\cdot\overline{4} = \overline{20} = \overline{6}$, ya que $20 = 2\cdot 7 + 6$. La tabla del producto es:

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{4}$	$\overline{6}$	$\overline{1}$	$\overline{3}$	$\overline{5}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{6}$	$\overline{2}$	$\overline{5}$	$\overline{1}$	$\overline{4}$
$\overline{4}$	$\overline{0}$	$\overline{4}$	$\overline{1}$	$\overline{5}$	$\overline{2}$	$\overline{6}$	$\overline{3}$
$\overline{5}$	$\overline{0}$	$\overline{5}$	$\overline{3}$	$\overline{1}$	$\overline{6}$	$\overline{4}$	$\overline{2}$
$\overline{6}$	$\overline{0}$	$\overline{6}$	$\overline{5}$	$\overline{4}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

Cuando trabajemos en $\mathbb{Z}/n\mathbb{Z}$, si no hay ambigüedad, denotaremos a la coclase $\overline{r}$, simplemente por r.

A.2. Construcción de cuerpos finitos

En la sección anterior vimos cómo, dado un número primo p, el cociente $\mathbb{Z}/p\mathbb{Z}$ es un cuerpo finito de cardinal p. Ahora veremos cómo se pueden construir cuerpos finitos de cardinal una potencia p^n de un primo p tomando un cociente apropiado del anillo de polinomios $\mathbb{Z}/p\mathbb{Z}[x]$ en la indeterminada x con coeficientes en $\mathbb{Z}/p\mathbb{Z}$

entre el ideal formado por los polinomios múltiplos de un polinomio irreducible de grado n. Primero necesitaremos dar algunas definiciones:

Definición Sea K un cuerpo, y sea $P(x) = \sum_i a_i x^i$ un polinomio no nulo de $K[x]$. El **grado de** $P(x)$ es el mayor entero no negativo n que cumple $a_n \neq 0$.

El grado de $P(x)$ se suele denotar por $\delta(P(x))$.

Es obvio que los polinomios de grado 0 son las unidades de $K[x]$. A estos se les llaman **polinomios constantes no nulos** (el polinomio nulo también es constante, pero no se le asigna grado).

Definición Sea K un cuerpo. Un polinomio $P(x) \in K[x]$ de grado estrictamente positivo es **irreducible** si $P(x) = Q_1(x)Q_2(x)$ implica que $\delta(Q_1(x)) = 0$ ó $\delta(Q_2(x)) = 0$.

Un polinomio no nulo $P(x)$ de $K[x]$, donde K es un cuerpo, se dice mónico si el coeficiente de $x^{\delta(P(x))}$ es 1_K (dicho coeficiente se llama **coeficiente director** de $P(x)$).Todo polinomio $P(x)$ de grado positivo de $K[x]$, donde K es un cuerpo, se puede descomponer de forma única, salvo por el orden de los factores, en la forma

$$P(x) = aP_1(x) \cdots P_m(x),$$

donde $m \in \mathbb{N}, a \in K$ y los polinomios $P_1(x), \ldots, P_m(x)$ son mónicos e irreducibles.

Ejemplos A.2.1

1. El polinomio $x^2 + \overline{1}$ no es irreducible en $\mathbb{Z}/2\mathbb{Z}[x]$, ya que $x^2 + \overline{1} = (x + \overline{1})^2$, pero $x + \overline{1}$ no es un polinomio constante.
2. En cambio, el polinomio $x^2 + x + \overline{1}$ sí es irreducible en $\mathbb{Z}/2\mathbb{Z}[x]$.

Cuando se trabaja en $\mathbb{Z}/p\mathbb{Z}[x]$ es habitual, cuando no hay ambigüedad, omitir las barras; así, los dos polinomios de los ejemplos se denotarían simplemente por $x^2 + 1$ y $x^2 + x + 1$.

Si K es un cuerpo y $P(x) \in K[x]$, el conjunto

$$(P(x)) = \{P(x)Q(x) \mid Q[x] \in K[x]\}$$

es un ideal, llamado el **ideal generado por el polinomio** $P(x)$. Dado un $Q(x) \in K[x]$, a la coclase $Q(x) + (P(x))$ del anillo cociente $K[x]/(P(x))$ se la suele denotar simplemente por $\overline{Q(x)}$ cuando no hay ambigüedad respecto al ideal $(P(x))$ considerado.

Proposición A.2.1 *El ideal $(P(x))$ es maximal si y sólo si el polinomio $P(x)$ es irreducible.*

De aquí, utilizando la proposición A.1.4, se deduce lo siguiente:

Corolario A.2.2 *El anillo cociente $K[x]/(P(x))$ es cuerpo si y sólo si el polinomio $P(x)$ es irreducible.*

Ejemplo A.2.2 El cociente $\mathbb{Z}/2\mathbb{Z}[x]/(x^2+x+1)$ es un cuerpo. Si $P(x) \in \mathbb{Z}/2\mathbb{Z}[x]$ y $R(x)$ es el resto de la división de $P(x)$ entre x^2+x+1 (¡Cuidado!, la división es en $\mathbb{Z}/2\mathbb{Z}[x]$), entonces $\overline{P(x)} = \overline{R(x)}$ (aquí las barras representan coclases módulo $(P(x))$), luego

$$\mathbb{Z}/2\mathbb{Z}[x]/(x^2+x+1) = \{\overline{a+bx} \mid a, b \in \mathbb{Z}/2\mathbb{Z}\},$$

y está claro que en esta representación no hay repeticiones. Por lo tanto,

$$\mathbb{Z}/2\mathbb{Z}[x]/(x^2+x+1) = \{\overline{0}, \overline{1}, \overline{x}, \overline{x+1}\}$$

es un cuerpo finito con 4 elementos. Para sumar, por ejemplo, $\overline{x}$ y $\overline{x+1}$, se hace

$$\overline{x} + \overline{x+1} = \overline{2x+1} = \overline{1}.$$

Para multiplicar los mismos elementos, se hace

$$\overline{x} \cdot \overline{x+1} = \overline{x(x+1)} = \overline{x^2+x} = \overline{1},$$

ya que $x^2 + x = (x^2+x+1) + 1$.

Para garantizar la existencia de cuerpos de orden potencia de un primo usaremos lo siguiente:

Proposición A.2.3 *Si p es un número primo y $n \in \mathbb{N}$, entonces existe por lo menos un polinomio $P(x) \in \mathbb{Z}/p\mathbb{Z}[x]$ irreducible de grado n.*

De ahí se deduce lo siguiente:

Proposición A.2.4 *Si p es un número primo y $n \in \mathbb{N}$, entonces existe un cuerpo finito de orden p^n.*

No podemos realmente aspirar a una mayor generalidad en los cardinales de los cuerpos finitos, como vemos en la siguiente proposición.

Proposición A.2.5 *Si K es un cuerpo finito, entonces $|K| = p^n$ para algún número primo p y algún número natural n.*

Para cada primo p y cada exponente n hay esencialmente un único cuerpo de orden p^n. Definiremos primero qué queremos decir con 'esencialmente uno'.

Definición Si F, K son dos cuerpos, una aplicación $f : F \longrightarrow K$ es un **isomorfismo** si es biyectiva y además satisface

1. $f(x+y) = f(x) + f(y)\ \forall x, y \in F$,
2. $f(x \cdot y) = f(x) \cdot f(y)\ \forall x, y \in F$.

Cuando existe un isomorfismo entre dos cuerpos F y K, se dice que éstos son **isomorfos**, y se denota por $F \simeq K$. Intuitivamente, esto quiere decir que tienen la misma estructura, es decir, que son indistinguibles desde el punto de vista de las propiedades algebraicas que afectan a la suma y el producto, y que se distinguen tan sólo en la forma de denotar sus elementos.

Proposición A.2.6 *Si F, K son dos cuerpos finitos del mismo cardinal, entonces F y K son isomorfos.*

De lo visto en las proposiciones A.2.4 y A.2.6 se deduce lo siguiente:

Proposición A.2.7 *Para cada número primo p y cada número natural n hay un único cuerpo de orden p^n, salvo isomorfismos.*

Dicho cuerpo se suele denotar por $\mathbb{F}_{p^n}$.

A.3. Ejemplos de cuerpos finitos

En esta sección veremos algunos ejemplos de cuerpos finitos de orden pequeño. Para construir un cuerpo de orden p^n es necesario conocer un polinomio irreducible de grado n con coeficientes en $\mathbb{Z}/p\mathbb{Z}$. Por lo general, no hay un único polinomio irreducible de grado n en $\mathbb{Z}/p\mathbb{Z}[x]$ aunque, por lo visto en la proposición A.2.6, todos ellos dan lugar a cuerpos isomorfos.

Dentro de los polinomios irreducibles $P(x)$, interesan especialmente aquellos que cumplen que las potencias $\overline{x}^i$ permiten obtener todos los elementos no nulos del cuerpo $\mathbb{Z}/p\mathbb{Z}[x]/(P(x))$. Dichos polinomios se llaman **primitivos**. El conocer estos polinomios es importante en algunas áreas de la Teoría de Códigos, como por ejemplo cuando se estudian los códigos BCH.

Por ejemplo, el polinomio $x^2 + 1 \in \mathbb{Z}/3\mathbb{Z}[x]$ es irreducible, pero no es primitivo, ya que al hacer las potencias de $\overline{x}$ obtenemos:

$\overline{x}^0 = \overline{1}$,

$\overline{x}^1 = \overline{x}$,

$\overline{x}^2 = \overline{2}$,

$\overline{x}^3 = \overline{2x}$,

$\overline{x}^5 = \overline{1}$,

y de ahí en adelante se van repitiendo cíclicamente los mismos cuatro elementos. Por lo tanto, al hacer potencias de $\overline{x}$ sólo se obtienen $\overline{1}, \overline{x}, \overline{2}, \overline{2x}$, y no se consiguen todos los 8 elementos no nulos del cuerpo.

En cambio, el polinomio $x^2 + 2x + 2 \in \mathbb{Z}/3\mathbb{Z}$, además de irreducible es primitivo, ya que las potencias de $\overline{x}$ dan:

$\overline{x}^0 = \overline{1}$,

$\overline{x}^1 = \overline{x}$,

$\overline{x}^2 = \overline{x+1}$,

$\overline{x}^3 = \overline{2x+1}$,

$\overline{x}^4 = \overline{2}$,

$\overline{x}^5 = \overline{2x}$,

$\overline{x}^6 = \overline{2x+2}$,

$\overline{x}^7 = \overline{x+2}$,

y en este caso sí obtenemos los 8 elementos distintos de cero del cuerpo. Al seguir haciendo potencias, estos 8 elementos se van repitiendo cíclicamente.

Proposición A.3.1 *Si p es un número primo y n es un número natural, hay por lo menos un polinomio primitivo de grado n en $\mathbb{Z}/p\mathbb{Z}[x]$.*

Por lo general, no hay un sólo polinomio primitivo de grado n en $\mathbb{Z}/p\mathbb{Z}[x]$. Por ejemplo, en $\mathbb{Z}/2\mathbb{Z}[x]$ hay dos polinomios primitivos de grado 3: $1+x+x^3$ y $1+x^2+x^3$.

En la tabla siguiente damos un ejemplo de un polinomio primitivo de grados 2,3,4,5 para los primos 2,3,5,7,11,13.

	$n=2$	$n=3$	$n=4$	$n=5$
$p=2$	x^2+x+1	x^3+x^2+1	x^4+x+1	$x^5+x^4+x^3+x^2+1$
$p=3$	x^2+2x+2	x^3+2x^2+x+1	$x^4+x^3+2x^2+2x+2$	$x^5+x^4+2x^3+x^2+x+1$
$p=5$	x^2+3x+3	x^3+4x+3	x^4+4x^2+x+2	$x^5+2x^4+3x^3+x^2+4x+2$
$p=7$	x^2+x+3	x^3+3x^2+4	x^4+3x^2+4x+3	$x^5+x^4+5x^3+2x^2+4$
$p=11$	x^2+4x+7	x^3+2x^2+6x+9	x^4+4x^2+x+2	$x^5+x^4+9x^3+3x^2+5x+5$
$p=13$	$x^2+4x+11$	x^3+7x^2+2x+6	x^4+4x^2+x+2	$x^5+x^4+x^3+11x^2+8x+6$

Para terminar, veamos un ejemplo de cómo podemos utilizar los datos de la tabla para calcular el inverso en un cuerpo finito. Consideramos un cuerpo de orden 8, utilizando el polinomio primitivo x^3+x^2+1 de dicha tabla, es decir, tomemos

$$\mathbb{Z}/2\mathbb{Z}[x]/(x^3+x^2+1).$$

Vamos a calcular el inverso de $\overline{x^2+1}$. Como x^2+1 y x^3+x^2+1 son primos entre sí en $\mathbb{Z}/2\mathbb{Z}[x]$, utilizando el algoritmo de Euclides extendido obtenemos

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

luego el inverso de $\overline{x^2+1}$ es $\overline{x^2+x+1}$.