Tema 3. Códigos Hamming

Justo Peralta Lópe

Juan Antonio López Ramos
Dpto. Álgebra y Análisis Matemático

Resumen: *En el primer punto, se definen códigos Hamming sobre un alfabeto binario. En él se muestra como se construye, sus parámetros y el algoritmo de decodificación. Además, se muestra cómo se puede modificar este tipo de códigos para aumentar su capacidad de detección y su aplicación en casos más prácticos. En el segundo punto, emplearemos los mismos conceptos que nos permiten definir un código Hamming binario por medio de la matriz de paridad para definir códigos Hamming sobre cualquier alfabeto y daremos los mecanismos necesarios para definir un algoritmo de decodificación basado en la relación del síndrome y las columnas de la matriz de paridad.En ambos casos mostraremos ejemplos prácticos de códigos Hamming y sus circuitos digitales o circuitos lógicos asociados.*

4.1. Caso binario

Definición 4.1.1. *Sea r un número entero positivo y sea H un $r \times (2^r - 1)$ matriz cuyas columnas son todos los vectores distintas de cero de $V(r, 2)$. Entonces el código que tiene a H como matriz de paridad es un código Hamming binario, al cual notaremos por $Ham(r, 2)$. Este tipo de códigos es un código perfecto con parámetros:*

Número de bits de paridad:	r
Longitud de palabra:	$2^r - 1$
Número de bits de paridad:	$2^r - 1 - r$
Distancia mínima:	3

Ejemplo 4.1.1. 1. *El código $Ham(2, 2)$ posee como matriz de paridad $H = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. Este código posee como matriz de generadora $G = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$. Es decir, es un código de repetición binario de longitud de palabra 3.*

2. *$Ham(3,2)$ posee como matriz de paridad a*

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

Se puede comprobar que es un $[7,4,3]$-código.

Teorema 4.1.1. *Un q-código (no necesariamente lineal) con parámetros $(n, M, 2t+1)$ satisface la siguiente ecuación:*

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

Si se verifica la igualdad entonces decimos que el código es perfecto

Otra definición, equivalente a la anterior, de código perfecto sería la siguiente

Definición 4.1.2. *Un código C capaz de corregir t errores se dice que es perfecto, si en su array de decodificación aparece como líderes todos los vectores de peso menor o igual que t, y ninguno de pesos mayor que t.*

Teorema 4.1.2. *Un código Hamming $Ham(r,2)$,con $r \geq 2$ verifica:*

1. *Es un $[2^r - 1, 2^r - 1 - r, 3]$-código lineal.*
2. *Es un código perfecto*

4.2. Decodificación para un $Ham(r,2)$

Ya que $Ham(r,2)$ es un código perfecto, los líderes de las clases son 2^r vectores del espacio vectorial $V(n,2)$ de pesos menor o igual que 1.

Ahora bien, supongamos que recibimos la palabra $y = x + e$ donde x es una palabra del código y e es el error cometido, con $w(e) = 1$. Supongamos también que la posición de e distinta de cero es la posición j. Entonces

$$S(y) = yH^t = (x+e)H^t = xH^t + eH^t = eH^t = (0 \dots 0\overset{j}{1}0 \dots 0)H^t = h_j$$

donde h_j es la j-ésima columna de H. Si ordenamos las columnas de H de menor a mayor, h_j será la representación de j en binario. Esto último lleva a la siguiente decodificación:

Algoritmo:

1. Recibimos la palabra y y calculamos su síndrome $S(y)$.
2. Si $S(y) = 0$ entonces y pertenece al código y no se han cometido errores.
3. Si $S(y) \neq 0$, entonces el síndrome nos indica la posición donde se ha producido el error en binario.

Ejemplo 4.2.1. *Para el código Hamming, $Ham(3,2)$ con matriz de paridad*

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

Si recibimos la palabra $y = 1101011$, $S(y) = 110$. Luego se ha producido un error en la posición 6. Obsérvese que si se produce un error doble, por ejemplo en las posiciones 3 y 7, $S(y) = h_3 + h_7 = (0011 + 0111) = 0100$ y decodificamos incorrectamente.

Para evitar este tipo de errores, podemos transformar H de tal forma que seamos capaces de corregir 1 error y detectar 2, es decir, pasamos de un código del SEC a uno del tipo SEC-DED. Para ello definimos una nueva matriz de paridad H' como sigue

$$H' = \begin{bmatrix} 0 & & \\ \vdots & \mathbf{H} & \\ 0 & & \\ 1 & 1 \ \ldots \ 1 & \end{bmatrix}$$

Ejemplo 4.2.2. *La matriz de paridad de un $Ham(4,2)$ viene dado por*

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

Si $v = v_1v_2v_3v_4v_5v_6v_7v_8v_9v_{10}v_{11}v_{12}v_{13}v_{14}v_{15}$ es la palabra recibida, entonces $S(y) = (s_4, s_3, s_2, s_1)$ con

$$\begin{aligned} s_1 &= v_1 + v_3 + v_5 + v_7 + v_9 + v_{11} + v_{13} + v_{15} \\ s_2 &= v_2 + v_3 + v_6 + v_7 + v_{10} + v_{11} + v_{14} + v_{15} \\ s_3 &= v_4 + v_5 + v_6 + v_7 + v_{12} + v_{13} + v_{14} + v_{15} \\ s_4 &= v_8 + v_9 + v_{10} + v_{11} + v_{12} + v_{13} + v_{14} + v_{15} \end{aligned}$$

Y el diseño del circuito que realiza la decodificación viene dado en la figura 4.1

La nueva matriz de paridad para poder detectar errores dobles viene dada por

$$H = \begin{bmatrix} \mathbf{0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \mathbf{0} & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ \mathbf{0} & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ \mathbf{0} & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \end{bmatrix}$$

Ahora el síndrome posee una ecuación más

$$s_{ov} = v_0 + v_1 + \cdots + v_{15}$$

Obsérvese que ahora, si recibimos una palabra (de longitud 16) con 2 errores, $s_{ov} = 0$, lo cual nos permitirá la detección de 2 errores o un número par de errores. El algoritmo de decodificación queda como sigue (hay que tener en cuenta que la palabra recibida ahora es de longitud 16 $v = v_0v_1 \ldots v_{15}$):

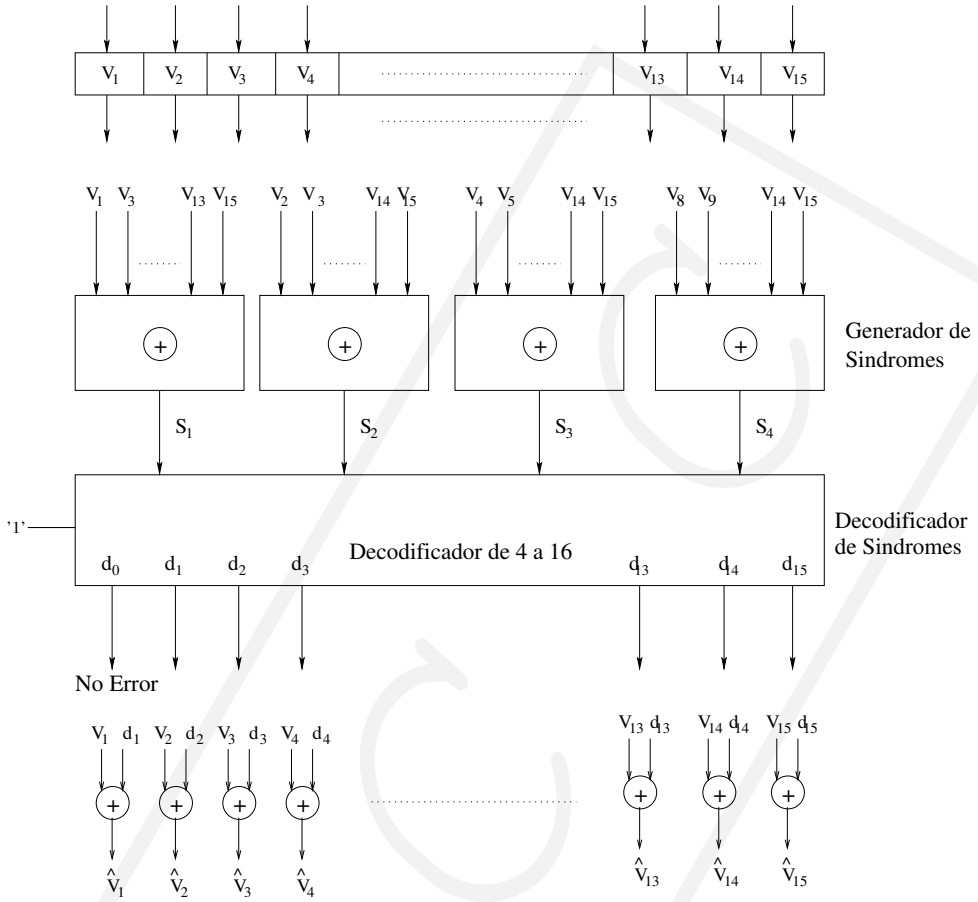

Figura 4.1: Decodificador para Ham(4,2)

1. Si $s = 0$, no se ha cometido ningún error.
2. Si $s_{ov} = 1$, entonces un error simple ha ocurrido y la posición viene dada en binario por (s_4, s_3, s_2, s_1) (ej:si $(s_4, s_3, s_2, s_1) = 0101$ entonces ha ocurrido un error en la posición 5).
3. Si $s_{ov} = 0$ y $(s_4, s_3, s_2, s_1) = 0$, un error doble o par ha ocurrido y no se puede corregir.

El circuito lógico correspondiente a la decoficación anteriormente descrita se puede observar en la figura 4.2

4.3. Código Hamming sobre cualquier alfabeto

Para obtener un código lineal con distancia mínima 3, basta con encontrar una matriz generadora H con cualquier par de columnas linealmente independiente, es decir, ninguna columna puede ser múltiplo por un escalar de otra y ninguna columna puede ser cero.

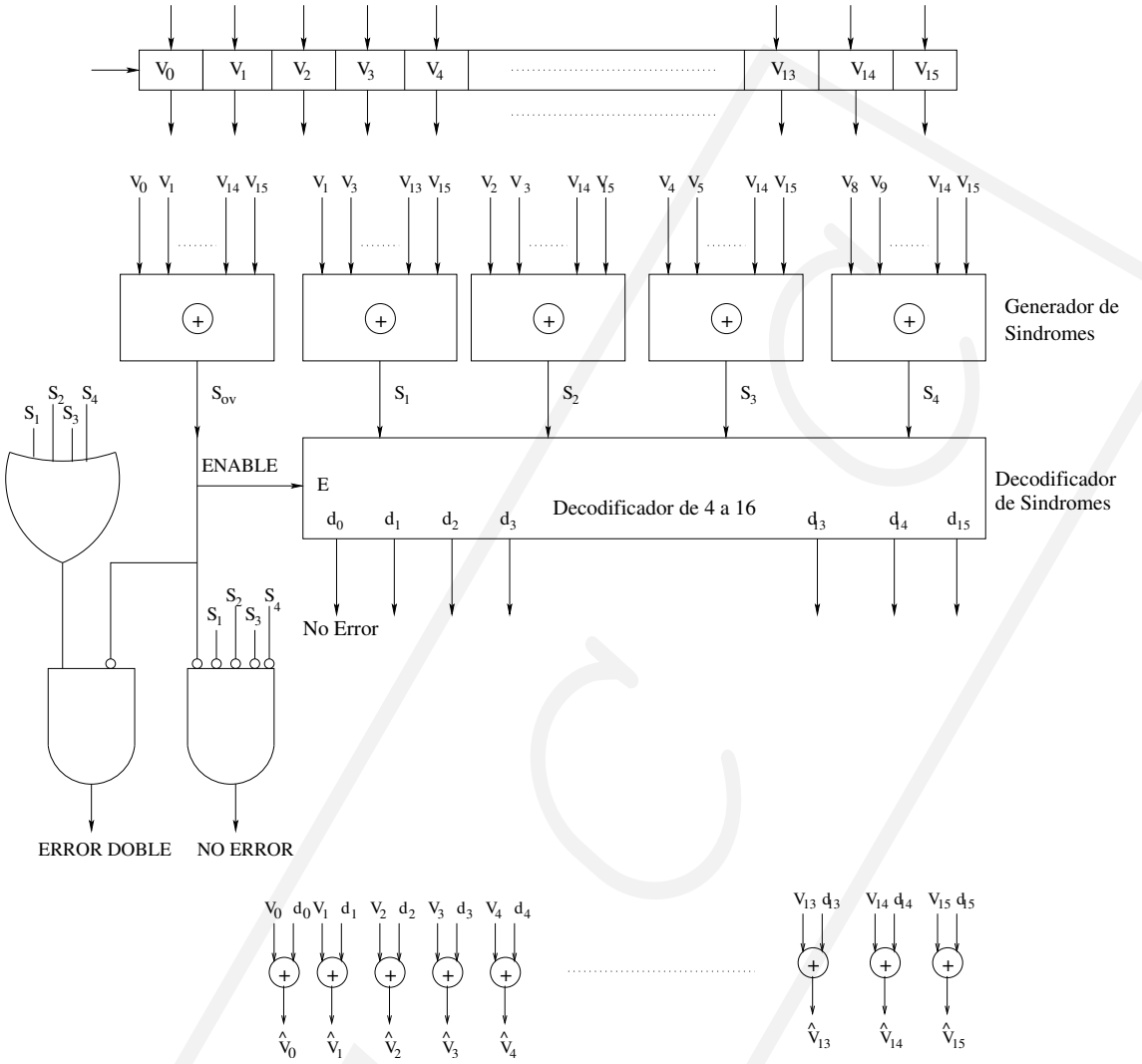

Figura 4.2: Decodificador para Ham(4,2) modificado (SEC-DED)

Para cualquier vector v del espacio vectorial $V(r,q)$, éste posee $q-1$ múltiplos por escalares distintos de cero dados por los conjuntos de la forma $[v] = \{\lambda v | \lambda \in GF(q) \text{y} \lambda \neq 0\}$, con $v \in V(r,q)$. Luego sobre el espacio vectorial $V(r,q)$ se puede definir una partición dado por la clase de equivalencia *ser múltiplo de*, es decir, $u \sim v$ si y sólo si $u = av$ con $a \neq 0 \in GF(q)$ y $u, v \in V(r,q)$. Como hemos visto, cada clase $[v]$ posee $q-1$ elementos, lo que nos da un total de $q^r - 1/q - 1$ clases.

A la hora de construir H, sólo tenemos que tomar un representante o un elemento de cada clase como columna. Esto nos garantiza que una columna no sea múltiplo de otra, ya que de no ser así, ambas columnas estarían en la misma clase. De esta forma, H es la matriz de paridad

de un código hamming sobre un alfabeto $GF(q)$, al cual notamos por $Ham(r,q)$, con parámetros $[q^r-1/q-1,(q^r-1/q-1)-r,3]$.

Ejemplo 4.3.1. *Construcción de la matriz de paridad de un $Ham(2,3)$.*
Las columnas de H son vectores no nulos del espacio vectorial $V(2,3)$. Dicho conjunto viene dado por:

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

Luego las clases de los múltiplos de un vector son:

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

Y la matriz de paridad estará formado por un vector como columna de cada clase

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

Ejemplo 4.3.2.

1. *La matriz de paridad de un $Ham(2,11)$ viene dado por*

$$H=\begin{bmatrix}0&1&1&1&1&1&1&1&1&1&1&1\\1&0&1&2&3&4&5&6&7&8&9&10\end{bmatrix}$$

2. *La matriz de paridad de un $Ham(3,3)$ viene dado por*

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

4.4. Decodificación en un $Ham(r,q)$

Teorema 4.4.1. *Un $Ham(r,q)$ es un código perfecto capaz de corregir errores simples*

Como $Ham(r,q)$ es un código perfecto, los líderes de la tabla de decodificación está formado por todos los vectores de peso menor o igual a 1. Esto quiere decir, sea cual sea el peso del error cometido, siempre decodificaremos como un error de peso 1. Si recibimos la palabra $y=x+e$ con el error $e=(00\ldots0\overset{i}{b}0\ldots0)$, diremos que el error cometido es de peso 1 y magnitud b. En este caso, el síndrome viene dado por:

$$S(y)=S(e)=(00\ldots0\overset{i}{b}0\ldots0)H^t=bh_i$$

donde h_i es la i-ésima columna de H. Luego el algoritmo de decodicación sería el que sigue:

1. Calculamos el síndrome de la palabra recibida $S(y)=yH^t$
2. Si $S(y)=0$ entonces no hay errores.

3. Si $S(y) \neq 0$, entonces al ser un código perfecto, $S(y) = bh_i$, y el error cometido es $e = (00\ldots0\overset{i}{b}0\ldots0)$ y decodificamos por $y - e$

Ejemplo 4.4.1. *Sea H la matriz de paridad de un $Ham(2,5)$, es decir,*

$$H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 & 3 & 4 \end{bmatrix}$$

y $y = 203031$ la palabra recibida. Entonces $S(y) = (2,3) = 2(1,4)$ y como $(1,4)$ es la quinta columna de H el error cometido es $e = 00002$ y decodificamos por $x = y - e = 203034$

4.5. Ejercicios

1. Escribir la matriz de paridad de un $[15,11]$-código Hamming binario. Explicar como se puede usar para corregir un error simple en una palabra. ¿Qué ocurre si se producen dos o más errores ?.
2. Modificar la matriz anterior para poder detectar más de un error y poner un ejemplo donde eso ocurra (construir el array de síndromes).
3. Modificar la matriz de paridad de un $Ham(3,2)$ para corregir un error y detectar dos. Decodificar la palabras $11100000, 01110000, 11000000$ y 00110011
4. Escribir la matriz de paridad para un $[8,6]$-código hamming sobre F_7. Usarlo para decodificar 35234106 y 10521360.
5. Sea C en $GF(5)$ generado por
$$\begin{bmatrix} 1 & 2 & 4 & 0 & 3 \\ 0 & 2 & 1 & 4 & 1 \\ 2 & 0 & 3 & 1 & 4 \end{bmatrix}$$
Calcular la mínima distancia del código.
6. Sea C en $GF(3)$ generado por
$$\begin{bmatrix} 1 & 2 & 0 & 2 & 1 & 0 \\ 2 & 0 & 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 2 & 1 & 2 \end{bmatrix}$$
Calcular la mínima distancia del código.
7. Sea C en $GF(2)$ generado por
$$\left[\begin{array}{c|cccc} & 1 & 1 & 0 & 0 \\ & 1 & 0 & 1 & 0 \\ & 0 & 1 & 1 & 0 \\ \mathbf{I_7} & 1 & 1 & 1 & 1 \\ & 1 & 1 & 0 & 1 \\ & 0 & 1 & 0 & 1 \\ & 1 & 0 & 0 & 1 \end{array}\right]$$
Calcular la mínima distancia del código.

8. Utilizar el Teorema **??** para construir un código con parámetros $[6, 3, 4]$