Universidad del Norte

División de Ciencias Básicas
Departamento de Matemáticas

Códigos de red lineales y la métrica del rango.

Yesneri M Zuleta Saldarriaga.

Trabajo presentado como requisito parcial para optar al título de Magíster en Matemáticas

Director: Prof. Dr. Ismael S Gutiérrez Garcia.
MSc. Jorge Robinson Evilla.

Barranquilla, Enero 10 del 2014

Dedicatoria

Al ser que hace que lleguen a ser todas las cosas y al cual le debemos todas las cosas a Jehová Dios.

Agradecimientos

Agradezco inicialmente a mi director de tesis Ismael Gutiérrez, a quien admiro en el campo académico y quien estuvo dispuesto a brindarme su apoyo, paciencia, acompañamiento, guía y tiempo en la elaboración de este trabajo de tesis.

Agradezco especialmente a COLCIENCIAS por la financiación del proyecto titulado: Códigos de red lineales y algunas implementaciones.

A mis otros maestros, les agradezco cada uno de sus conocimientos compartidos y dedicación al momento de transmitirlos y a quienes me siento en la obligación de mencionar: Agustín Barrios, Jorge Robinson, Jairo Hernández, Ricardo Prato y Carlos Vega.

Deseo agradecerle sin duda alguna a mis padres que son el mejor regalo de mi vida, y a quien no les puedo negar el hecho de que gracias a esa infinita sabiduría que para mi ustedes poseen influyeron en mi madurez y en mis decisiones, lo que me dio el animo y fortaleza suficiente para alcanzar este nuevo logro, es para mi un honor ser su hija y contar con todo ese apoyo y amor que me dieron en el proceso de esta tesis.

También deseo agradecer a todas esas hermosas personas que Dios puso en mi camino para fortalecerme y regalarme una sonrisa, una voz de apoyo y palabras de animo en los momentos en que tal vez me pude sentir agobiada o abatida, mis gracias sinceras a esas personas importantes para mi, Doris, Yaneth, Diana, Monica Zuleta y Eder Martinez a ustedes gracias por estar conmigo en estos tiempos tan importantes, a mis compañeros de casi toda mi vida profesional Edwin Bolaño, Darling Vasquez e Ivan Gonzalez, ha sido muy grato para mi estar junto a ustedes en este proceso.

Agradezco a todos aquellos que crean merecerlo y a quien no por ser menos importantes dejo de mencionar en estas líneas, muchas gracias.

Introducción

El presente trabajo de tesis esta enmarcado en la teoría de códigos de red lineales. Esta teoría relativamente nueva, tuvo probablemente su origen en el trabajo de R. Ahlswede, N. Cai, Shuo-Yen Li y R. W. Yeung [1]. No obstante la inspiración del presente trabajo está en los trabajos de R. Koetter, D. Silva y F. Kschischang [3], [4] y E. Gabidulin [10], [11].

El documento está dividido en tres capítulos y un apéndice. En el primer capítulo se presentan algunos aspectos elementales de la teoría cláasica de códigos de bloques sobre cuerpos finitos. Haciendo énfasis en la descripción de los parámetros de los códigos lineales, algunas técnicas de decodificación, ciertas clases importantes de códigos como por ejemplo los códigos de peso constante.

En el capitulo segundo introducimos los códigos de red y las métrica de subespaciós. Esta métrica resulta de especial importancia en la codificación en red aleatoria lineal. Similar como en la teoría clásica de códigos, establecer cotas para el tamaño de los códigos es un problema central. Presentaremos algunas de ellas como la del empaquetamiento esférico y la cota de Singleton para códigos de dimensión constante.

En el tercer capítulo se consideran códigos matriciales y se introduce la métrica del rango. Seguidamente se establecen relaciones entre las dos métricas y se comparan algunos resultados obtenidos en el capítulo anterior con los que resultan en el nuevo contexto. Posteriormente se consideran los códigos de Gabidulin y su relación con la métrica del rango. Para finalizar el capítulo se presenta una construcción de un código de Gabidulin, utilizando como herramienta computacional el sistema de software matemático libre SAGE. Cabe resaltar que las implementaciones forman parte de algunos objetivos a cumplir dentro de un proyecto de Joven investigador financiado por COLCIENCIAS.

La última parte son dos apéndices en los que se recogen algunos resultados del algebra lineal básica asociados el rango de una matriz y además los códigos de los algoritmos implementados.

Índice general

Capítulo 1

Códigos lineales clásicos

1.1. Los parámetros de un código lineal

Durante la transmisión de información digital a través de un sistema o **canal** se producen errores prácticamente inevitables debido a la presencia de ruido y a otros factores tales como la interferencia, la intermodulación y los ecos. Es por ello necesario establecer maneras, si no para evitar errores, por lo menos para poder reconocer su presencia y si es posible corregirlos. El proceso de control de error es de gran importancia debido a la gran redundancia de los datos digitales, donde un grupo de números o símbolos alfanuméricos erróneos pueden tener un significado, pero dicho significado puede ser muy diferente a la información original. Precisamente la tarea de la teoría de códigos es codificar y decodificar un mensaje.

Suponga que se desea enviar un mensaje x, cuyos caracteres pertenecen a un alfabeto K, a través de un canal de comunicación. Este mensaje original es transformado en forma digital, obteniendo así un **codeword** c o **palabra código**. Este proceso se conoce como **codificación**. Después que el mensaje pasa por el canal y es decodificado el receptor obtiene el mensaje y, donde es usual que $x \neq y$. Un modelo simple de un sistema de comunicación es mostrado en la siguiente figura.

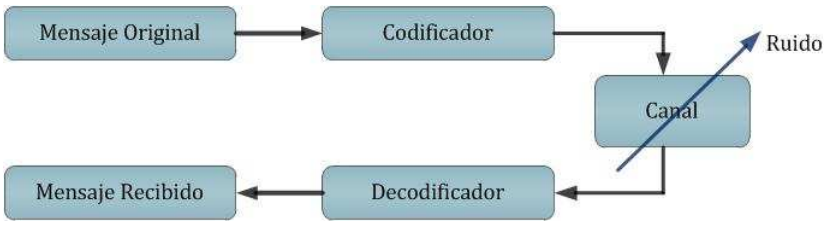

Figura 1.1: Transmisión de la información

Como ejemplo de alfabeto consideraremos siempre un cuerpo finito K, donde $|K| = q$ y q es potencia de algún número primo. Es decir, $K \cong \mathbb{F}_q$ esto es lo que precisamente permite aplicar a los problemas de codificación todos los recursos del algebra y la geometría.

Podemos hacer la distinción de una clase especial de códigos de bloque, conocidos como códigos lineales. Antes recordemos que K^n es el producto cartesiano de n copias de K. Es decir,

$$K^n = \{(a_1, \ldots, a_n) \mid a_j \in K\}.$$

En ocasiones en vez de usar la notación para vector $a = (a_1, \ldots, a_n)$ se usara $a = a_1 \ldots a_n$, con el fin de minimizar espacios.

1.1.1 Definición. Sea K un cuerpo finito y $n \in \mathbb{N}$. Un ***código lineal*** C sobre K es un subespacio vectorial del espacio K^n. Escribiremos para ello $C \leq K^n$. Si $\dim_K(C) = k$, entonces diremos que C es un $[n, k]$-código sobre K.

Si el canal no trabaja libre de errores, entonces un codeword enviado, digamos $c = (c_1, \ldots, c_n)$ es recibido como el vector $v = (v_1, \ldots, v_n)$. El número de errores ocurridos durante la transmisión está dado por el número de elementos del conjunto

$$\{j \mid u_j \neq v_j, j = 1, \ldots, n\}.$$

En la siguiente definición precisamos la noción de distancia entre vectores, para ello definimos una función sobre el espacio vectorial K^n, donde K es un cuerpo finito con q elementos. Esta función se denominara ***distancia de Hamming***[1].

[1]Richard Wesley Hamming (1915 – 1998). Matemático estadounidense, trabajó en temas relacionados con la informática y las telecomunicaciones, también trabajo en Los Álamos en el proyecto de la bomba atómica, pero casi toda su carrera se desarrollo en los laboratorios Bell.

1.1.2 Definición. Sean K un cuerpo finito y $n \in \mathbb{N}$. Para $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n) \in K^n$ definimos la ***distancia de Hamming*** d entre u y v de la siguiente manera

$$d(u, v) := |\{j \mid u_j \neq v_j, j = 1, \ldots, n\}|.$$

Es decir, la distancia entre dos vectores es el número de coordenadas en las que estos difieren.

1.1.3 Ejemplo. para $u = 11101$, $v = 01110 \in \mathbb{F}_2^5$ se verifica a través de la definición anterior

$$d(u, v) = d(11101, 01110) = 3.$$

1.1.4 Teorema. Sean K un cuerpo finito. Entonces la distancia de Hamming d es una métrica. Es decir, para todo $u, v, w \in K^n$ se verifican

(a) $d(u, v) \geq 0$ y $d(u, v) = 0$ si y solo si $u = v$.

(b) $d(u, v) = d(v, u)$.

(c) $d(u, v) \leq d(u, w) + d(w, v)$.

(d) Además d es invariante bajo traslaciones. Es decir,
$d(u + w, v + w) = d(u, v)$.

Demostración. Sean $u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n)$ y $w = (w_1, \ldots, w_n) \in K^n$

(a)

$$\begin{aligned} d(u, v) = 0 &\Leftrightarrow |\{j \mid u_j \neq v_j, j = 1, \ldots, n\}| = 0 \\ &\Leftrightarrow u_j = v_j \\ &\Leftrightarrow u = v. \end{aligned}$$

(b)

$$\begin{aligned} d(u, v) &= |\{j \mid u_j \neq v_j, j = 1, \ldots, n\}| \\ &= |\{j \mid v_j \neq u_j, j = 1, \ldots, n\}| \\ &= d(v, u). \end{aligned}$$

(c) Sea $Q = \{j \mid u_j \neq v_j, j = 1, \ldots, n\}$. Entonces $|Q| = d(u, v)$.
Sean

$$V = \{j \mid u_j \neq v_j \wedge u_j \neq w_j, j = 1, \ldots, n\},$$

$$U = \{j \mid u_j \neq v_j \wedge u_j = w_j, j = 1, \ldots, n\}.$$

Puesto que Q es la unión disjunta de los conjuntos U y V se obtiene lo siguiente:

$$|Q| = d(u, v) = |U| + |V|,$$

por la definición de $d(u, v)$ tenemos que $|V| \leq d(u, w)$ y de U tenemos que $w_j = u_j \neq v_j$. Entonces $|U| \leq d(w, v)$. Por lo tanto

$$d(u, w) = |U| + |V| \leq d(u, w) + d(w, v).$$

(d) Por otro lado,

$$\begin{aligned} d(u, v) &= |\{j \mid u_j \neq v_j, j = 1, \ldots, n\}| \\ &= |\{j \mid v_j + w_j \neq u_j + w_j, j = 1, \ldots, n\}| \\ &= d(v + w, u + w). \square \end{aligned}$$

La distancia de Hamming juega un papel importante en la teoría de códigos, cuando se decodifica el vector recibido escogiendo el codeword más cercano, esta técnica se conoce como decodificación mediante el ***vecino más proximo***. Para ello el canal debe cumplir con ciertas condiciones, las cuales en la practica se satisfacen a menudo. Sea K un cuerpo finito con q elementos.

(a) Todo $a \in K$ tiene la misma probabilidad $p < \frac{q-1}{q}$ de ser recibido con error.

(b) Si un símbolo es recibido con error, entonces cada uno de los $q - 1$ errores posibles es igualmente probable.

Decodificación de máxima verosimilitud

Este método suele llamarse también **decodificación ML**, por su nombre en inglés *Maximun Likelihood Decoding*.

Sea K un cuerpo finito con q elementos y $C \subseteq K^n$. Para $c \in C$ y $v \in K^n$ notemos con $P(v|c)$ la probabilidad condicional que es recibido el vector v, dado que fue enviado el codeword c. Una decodificación de máxima verosimilitud o simplemente una ML-decodificación, decodificada el vector $v \in K^n$ mediante un codeword $c \in C$, para el cual se verifica que

$$P(v|c) = \text{máx}\{P(v|c') \mid c' \in C\}.$$

Es decir, mediante el codeword con mayor probabilidad de haber sido enviado. Si existe mas de un codeword que alcanza dicho máximo, entonces se elige

uno aleatoriamente.

Bajo las condiciones impuestas anteriormente al canal, la probabilidad condicionada $P(v|c)$ con $d(v,c') = j$ esta dada por

$$P(v|c') = \left(\frac{p}{q-1}\right)^j (1-p)^{n-j},$$

ya que existen j posiciones adulteradas. Note que

$$P(v|c') = \left(\frac{p}{(q-1)(1-p)}\right)^j (1-p)^n.$$

De la condición $p < \frac{q-1}{q}$ se sigue que $p < \frac{p}{(q-1)(1-p)} < 1$. Por tanto la función

$$f(j) := \left(\frac{p}{(q-1)(1-p)}\right)^j (1-p)^n$$

es estrictamente decreciente. Por lo tanto se verifica que

$$P(v|c) = \text{máx}\{P(v|c') \mid c' \in C\}.$$

exactamente para los codewords $c \in C$ para los cuales se cumple que

$$d(v,c) := \text{mín}\{d(v,c') \mid c' \in C\}.$$

En el caso de canales simétricos q-arios la ML-decodificación asigna a un vector recibido el codeword mas cercano. Es decir, el codeword para el cual la distancia de Hamming con el vector recibido sea mínima.

1.1.5 Definición. Sea C un $[n,k]$-código sobre un cuerpo finito K. Si $|C| > 1$, entonces llamaremos a

$$d(C) := \text{mín}\{d(c,c')|c,c' \in C, c \neq c'\}$$

la ***distancia mínima*** de C y si $|C| = 1$, entonces definimos $d(C) := 0$.

1.1.6 Definición. Si C es un $[n,k]$-código sobre un cuerpo finito K y $d(C) = d$, entonces decimos que C es un $[n,k,d]$-código sobre K. Llamaremos a $[n,k,d]$ los parámetros de C.

Si $|K| = q$, entonces es usual decir que C es un $[n,k,d]_q$-código.

1.1.7 Definición. Sea K un cuerpo finito y $r \in \mathbb{N}_0$. Para $u \in K^n$ definimos

$$B_r(u) := \{v \mid v \in K^n, d(u,v) \leq r\}$$

y se denomina esfera o bola con centro en u y radio r.

1.1.8 Teorema. (Cota de Singleton) Sea C un código de longitud n sobre un cuerpo finito K, con $|K| = q^k$ y distancia mínima d. Entonces

$$d \leq n - \log_q |C| + 1.$$

O equivalentemente

$$|C| \leq q^{n-d+1}.$$

Demostración. Consideremos la función f.

$$\begin{array}{rccc} f: & K^n & \longrightarrow & K^{n-d+1} \\ & (x_1, \ldots, x_n) & \longmapsto & (x_1, \ldots, x_{n-d+1}) \end{array}$$

Dado que dos codewords distintos tienen distancia mínima por lo menos d, se tiene que la restricción de f a C es inyectiva. En efecto sean $x, y \in K^n$, digamos $x = (x_1, \ldots, x_n)$ y $y = (y_1, \ldots, y_n)$. Si $x \neq y$, entonces forzosamente $(x_1, \ldots, x_{n-d+1}) \neq (y_1, \ldots, y_{n-d+1})$, ya que de lo contrario se tendría que $d(x,y) \leq d-1$, lo que es contradictorio. Así que

$$|C| = |f(C)| \leq |K^{n-d+1}| = q^{n-d+1}.$$

1.1.9 Definición. Sea C un código de longitud n sobre un cuerpo finito K, con $|K| = q$ y distancia mínima d. Diremos que C es un **MDS-código** *(Maximum Distance Separable),* si C alcanza la cota de Singleton. Es decir, si

$$|C| = q^{n-d+1}$$

1.1.10 Definición. Sea $x \in K^n$. El **peso** del vector x notado con $\omega t(x)$ se define de la siguiente manera:

$$\omega t(x) := |\{j \mid x_i \neq 0, j = 1, \ldots, n\}|.$$

Es decir, el peso del vector x es el número de coordenadas no nulas de este. Si $x \neq 0$, entonces $1 \leq \omega t(x) \leq n$.

1.1.11 Ejemplo. Sea $K = \mathbb{F}_2$. Si $x = 1001$, entonces $\omega t(x) = 2$

1.1.12 Observación. Note que la relación entre el peso y la distancia esta dada por:

(a)
$$\begin{aligned} \omega t(x) &= |\{j \mid x_i \neq 0, j = 1, \ldots, n\}| \\ &= d(x, 0). \end{aligned}$$

(b)
$$\begin{aligned} \omega t(x - y) &= |\{j \mid x_i - y_i \neq 0, j = 1, \ldots, n\}| \\ &= |\{j \mid x_i \neq y_i, j = 1, \ldots, n\}| \\ &= d(x, y). \end{aligned}$$

1.1.13 Definición. Sean K un cuerpo finito, $n \in \mathbb{N}$ y $C \subseteq K^n$.

(a) Si $C \neq 0$, entonces se define ***peso mínimo*** de C, notado con $\omega t(C)$, de la siguiente manera

$$\omega t(C) := \text{mín}\{\omega t(x) \mid 0 \neq x \in C\}.$$

Si $C = 0$, entonces $\omega t(C) = 0$.

(b) El ***soporte*** de $x = (x_1, x_2, ..., x_n) \in K^n$ es notado y definido de la siguiente manera:

$$\text{sop}(x) := \{j \mid x_j \neq 0\}.$$

Es decir que el soporte indica cuales coordenadas son distintas de cero, mientras que el peso determina cuantas son distintas de cero. Por lo tanto podemos afirmar que $\omega t(x) = |\text{sop}(x)|$.

En el caso de que la característica del cuerpo sea 2 la resta y la suma coinciden, pudiéndose así calcular las distancias como se muestra a continuación.

1.1.14 Corolario Si $\text{char}(K) = 2$, entonces para todo $x, y \in K^n$

(a) $\omega t(x - y) = \omega t(x + y)$

(b) $d(x, y) = \omega t(x) + \omega t(y) - 2|\text{sop}(x) \cap \text{sop}(y)|$.

1.1.15 Ejemplo. Sea $x = 1001$ e $y = 1101$, entonces $x + y = 0100$ y $|\text{sop}(x) \cap \text{sop}(y)| = 2$. Teniendo así que

$$\omega t(x + y) = 1 = \omega t(x) + \omega t(y) - 2|\text{sop}(x) \cap \text{sop}(y)| = 2 + 3 - 2 * 2.$$

1.1.16 Teorema. Sea K un cuerpo finito, $n \in \mathbb{N}$ y $C \leq K^n$. Entonces

$$d(C) = \omega t(C).$$

Demostración. Si $C = 0$, entonces la afirmación es inmediata.

Supongamos que $C \neq 0$. De la invariancia bajo traslaciones de d se sigue

$$\begin{aligned} d(C) &= \text{mín}\{d(c, c')|c, c' \in C, c \neq c'\} \\ &= \text{mín}\{d(c - c', 0)|c, c' \in C, c \neq c'\} \\ &= \text{mín}\{d(x, 0)|x \in C, x \neq 0\} \\ &= \omega t(C). \end{aligned}$$

El número de errores cometidos durante la transmisión esta dado por la distancia de Hamming, entonces podríamos preguntarnos. ¿Cuantos errores puede corregir un código? daremos la respuesta a esa pregunta mediante los siguientes teoremas.

1.1.17 Teorema. Sea C un código con distancia mínima d y $t \in \mathbb{N}_0$

(a) Si $d \geq t + 1$, entonces C es un t-detector o puede detectar t-errores.

(b) Si $d \geq 2t + 1$, entonces C es un t-corrector o puede corregir t-errores.

Demostración.

(a) Si $d \geq t + 1$ y en la transmisión de $c \in C$ ocurrieron a lo mas t errores, entonces para el vector recibido $v \in K^n$ se verifica que

$$d(c, v) \leq t \leq d - 1 \leq d.$$

Por la tanto $v \in \mathrm{B}_t(c)$ y no puede ser un codeword. En consecuencia el decodificador detecta que han ocurrido errores durante la transmisión.

(b) Sean $\mathrm{B}_t(c)$ y $\mathrm{B}_t(c')$ esferas que contienen a las palabras recibidas c y c' respectivamente; supongamos que

$$\mathrm{B}_t(c) \cap \mathrm{B}_t(c') \neq \emptyset.$$

Para todo $c, c' \in \mathrm{C}$ con $c \neq c'$ existirá x tal que $x \in \mathrm{B}_t(c)$ y $x \in \mathrm{B}_t(c')$, entonces por desigualdad triangular tenemos

$$d(c, c') \leq d(c, x) + d(x, c') \leq 2t < d - 1 < d,$$

lo que contradice la minimalidad de d. Por lo tanto

$$\mathrm{B}_t(c) \cap \mathrm{B}_t(c') = \emptyset. \quad \square$$

1.1.18 Corolario Sea C un código con distancia mínima d y $t \in \mathbb{N}_0$.

(a) C puede ser usado para detectar hasta $d-1$ errores.

(b) C puede ser usado para detectar hasta $\lfloor \frac{d-1}{2} \rfloor$ errores.

Demostración.

(a) Note que $d \geq t+1$ si y solo si $t \leq d-1$.

(b) La desigualdad $d \geq 2t+1$ se verifica si y solo si $t \leq \frac{d-1}{2}$.□

1.2. Matriz generadora y matriz de control de un código lineal

1.2.1 Definición. Sea C un $[n,k]$-código sobre un cuerpo finito K y sean $g_1 = (g_{11}, \ldots, g_{1n}), \ldots, g_k = (g_{k1}, \ldots, g_{kn}) \in K^n$. Si $B = (g_1, \ldots, g_n)$ es una base para C, entonces diremos que

$$G = \begin{pmatrix} g_1 \\ \vdots \\ g_k \end{pmatrix} = \begin{pmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & & \vdots \\ g_{k1} & \cdots & g_{kn} \end{pmatrix} \in \mathrm{Mat}(k \times n, K)$$

es una matriz generadora de C.

1.2.2 Teorema. Sea C un $[n,k]$-código sobre un cuerpo finito K. Entonces $G \in \mathrm{Mat}(k \times n, K)$ es una matriz generadora de C, si y solo si

$$C = \{uG \mid u \in K^k\}.$$

Demostración. Supongamos que

$$G = \begin{pmatrix} g_1 \\ \vdots \\ g_k \end{pmatrix} = \begin{pmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & & \vdots \\ g_{k1} & \cdots & g_{kn} \end{pmatrix} \in \mathrm{Mat}(k \times n, K)$$

es una matriz generadora de C. Si $(u_1, \ldots, u_k) \in K^k$, entonces

$$\begin{aligned} uG &= (u_1 g_{11} + \cdots + u_k g_{k1}, \ldots, u_1 g_{1n} + \cdots + u_k g_{kn}) \\ &= (u_1 g_{11}, \ldots, u_1 g_{1n}) + \cdots + (u_k g_{k1}, \ldots, u_k g_{kn}) \\ &= u_1(g_{11}, \ldots, g_{1n}) + \cdots + u_k(g_{k1}, \ldots, g_{kn}) \\ &= u_1 g_1 + \cdots + u_k g_k \in C. \end{aligned}$$

Esto demuestra que $\{uG \mid u \in K^k\} \subseteq C$. Ahora por ser uG combinación lineal de los elementos de la base,

$$C \subseteq \{uG \mid u \in K^k\}.$$

Por lo tanto

$$\{uG \mid u \in K^k\} = C.$$

Recíprocamente, supongamos que $C = \{uG \mid u \in K^k\}$ y notemos con e_j, con $j = 1, \ldots, n$ los vectores de la base canónica de K^k. Entonces

$$\begin{aligned} e_1 G &= g_1 \\ &\vdots \\ e_k G &= g_k \end{aligned}$$

y se tiene que las filas de G pertenece a C. Por lo tanto

$$C = \{u_1 g_1 + \cdots + u_k g_k | u_j \in K\} = \langle g_1, \ldots, g_k \rangle.$$

Dado que $\dim_K C = k$, se sigue que $B = (g_1, \ldots, g_k)$ es una base para C y se tiene que G es una matriz generadora para C.□

Si G, es una matriz generadora de un $[n, k]$-código C sobre K, entonces al efectuar operaciones elementales sobre las filas de G, se tiene nuevamente una matriz generadora de C.

1.2.3 Definición. Si un $[n, k]$-código C admite una matriz generadora G de la forma

$$G = (I_k \mid B);$$

es decir,

$$G = \begin{pmatrix} 1 & 0 & \cdots & 0 & * & \cdots & * \\ 0 & 1 & \cdots & 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & * & & * \end{pmatrix},$$

entonces esta matriz se le llama matriz generadora de C en ***forma estándar.***

No todo código lineal tiene una matriz generadora en forma estándar. Sin embargo siempre existe un código equivalente, que admite una matriz generadora en forma estándar.

1.2.4 Definición. Sea C un $[n,k]$-código sobre un cuerpo finito K, con $k<n$. Diremos que $H \in \mathrm{Mat}(n-k\times n, K)$ es una ***matriz de control*** para C, si

$$C = \{u \in K^n \mid Hu^t = 0\}.$$

1.2.5 Teorema. Sea C un $[n,k]$-código sobre un cuerpo finito K, con $k<n$. Entonces existe una matriz $H \in \mathrm{Mat}(n-k\times n, K)$ tal que, para $x \in K^n$ se verifica

$$x \in C \Leftrightarrow Hx^t = 0.$$

Además Rang $H = n-k$. Por lo tanto, si G es una matriz generadora de C, entonces $HG^t = 0$.

Demostración. Sea $B = (g_1, \ldots, g_k)$ una base para C, donde para cada $j = 1, \ldots, k$ se tiene que $g_j = (g_{j1}, \ldots, g_{jk})$ y definamos

$$G = \begin{pmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & & \vdots \\ g_{k1} & \cdots & g_{kn} \end{pmatrix}.$$

Consideremos el sistemas de ecuaciones lineales $Gz^t = 0$, con $z = (z_1, \ldots, z_n)$. Esto es,

$$\begin{aligned} g_{11}z_1 + \cdots + g_{1n}z_n &= 0 \\ &\vdots \\ g_{k1}z_1 + \cdots + g_{kn}z_n &= 0. \end{aligned} \tag{1.1}$$

Dado que Rang $G = k$, del algebra lineal se sigue que la dimensión del espacio solución del sistema es $n-k$.

Sea $B' = (h_1, \ldots, h_{n-k})$, una base para el espacio solución del sistema (1.1), donde para cada $j = 1, \ldots, n-k$ se tiene que $h_j = (h_{j1}, \ldots, h_{jn})$. Definamos la matriz $H \in \mathrm{Mat}(n-k\times n, K)$ mediante

$$H := \begin{pmatrix} h_{11} & \cdots & h_{1n} \\ \vdots & & \vdots \\ h_{n-k,1} & \cdots & h_{n-k,n} \end{pmatrix}$$

Entonces $Hg_i^t = 0$, para todo $i = 1, \ldots, k$. Por lo tanto, $Hx^t = 0$, para todo $x \in C$.

Dado que

$$\text{Rang } H = n - \dim(ker(H)) = n-k,$$

se sigue que $\dim(ker(H)) = k$. Es decir, $C = ker(H)$.□

1.2.6 Ejemplo. Sea C un $[7,4]$-código binario con matriz generadora

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

- Inicialmente resolvemos $Gx^t = 0$, entonces
  para $x = (x_1, x_2, x_3, x_4, x_5, x_6, x_7)$

$$Gx^t = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

  Da como resultado el siguiente sistema de ecuaciones

$$\begin{cases} x_1 + x_2 + x_7 = 0 \\ x_3 + x_4 + x_5 + x_7 = 0 \\ x_1 + x_4 + x_5 = 0 \\ x_5 + x_6 + x_7 = 0. \end{cases}$$

  Despejando obtenemos que

$$\begin{cases} x_2 = x_1 + x_7 \\ x_3 = x_1 + x_7 \\ x_4 = x_1 + x_5 \\ x_6 = x_1 + x_2 + x_5 \\ x_7 = x_1 + x_2. \end{cases}$$

- Luego el conjunto solución esta dado por

$$S = \{(x_1, x_2, x_2, x_1 + x_5, x_5, x_1 + x_2 + x_5, x_1 + x_2) \mid x_1, x_2, x_5 \in \mathbb{F}_2\}$$

  Note que

$$S = \{x_1 1001011 + x_2 0110011 + x_5 0001110 \mid x_1, x_2, x_5 \in \mathbb{F}_2\}$$

- Por lo tanto una matriz de control para C es

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

Una relación entre la distancia mínima y la matriz de control la efectúa el siguiente lema

1.2.7 Lema. Un código lineal C con matriz de control H tiene distancia mínima $d \geq s+1$ si, y solo, si cualesquiera s columnas de H son linealmente independientes.

Demostración. Sea C un código lineal con matriz de control H tal que $d \geq s+1$. Supongamos que H tiene s columnas linealmente dependientes si esto ocurre, existirán s palabras códigos no nula de C, entonces $\omega t(C) \leq s$, lo que da lugar a una contradicción. Por tanto s columnas de H son linealmente independientes.

Reciprocamente supongamos que H tiene s columnas linealmente independientes, sin perder generalidad supongamos que son las s primeras columnas. Luego existirán $\lambda_1, ..., \lambda_s \in \mathbb{F}_q$ no todas nulas tales que,

$$\lambda_1 H_1 + ... + \lambda_s H_s = 0.$$

Tomemos $\lambda_i = c_i$ para $i = 1, ..., s$ y $c_i = 0$ para $i = s+1, ..., n$ y construyamos $c = (c_1, ..., c_n)$. Por construcción tendremos que $Hc^t = 0$, luego $c \in C$ y c es no nulo ya que en las s primeras componentes hay por lo menos una que no es nula, es decir $\omega t(c) = d \leq s$. Como H no tiene columnas linealmente dependientes, entonces el código tiene palabras de peso mayor que s o lo que es lo mismo $d > s$, luego $d \geq s+1$.□

Decodificación con códigos lineales

Se detecta un error cuando una de las palabras recibidas no pertenece al código. En el caso de los códigos lineales debido a su estructura algebraica dicho error lo podemos expresar como la diferencia entre dos palabras como se muestra a continuación.

1.2.8 Definición. Si c es un codeword y v es la palabra enviada a través de un canal de comunicación ¨ruidoso¨, entonces

$$e = v - c = e_1, \ldots, e_n.$$

e es llamado error o el vector error.

Todo posible vector error e de un vector recibido v es un vector cercano a v el vector más probable es el vector e con menor peso en el cercano a v. De este modo decodificando v se tiene que

$$c = v - e,$$

donde c es un codeword.

Otra alternativa de decodificación, es la decodificación mediante el síndrome expuesta a continuación.

Para $v \in K^n$, el vector $s_v \in K^{n-k}$ definido por

$$s_v := Hv^t,$$

se denomina síndrome de v, donde H es una matriz de control para C. Note que si $c \in C$, entonces

$$\mathrm{H}c^t = 0 = s_c,$$

luego todo codeword tiene síndrome cero.

Por otro lado si $v \in K^n$ es nuevamente el vector recibido, entonces

$$s_v = \mathrm{H}v^t = \mathrm{H}(e^t + c^t) = \mathrm{H}e^t + \mathrm{H}c^t = \mathrm{H}e^t = s_e$$

Mostremos ahora que dos codeword están en la misma clase lateral de C si, y solo, si tienen igual síndrome.

1.2.9 Lema. Sea K un cuerpo finito y $n \in \mathbb{N}$. Entonces $x, y \in K^n$ tienen el mismo síndrome, si y solo si $x + C = y + C$.

Demostración. Sea $x, y \in K^n$. Entonces

$$\begin{aligned} x + C = y + C &\Leftrightarrow x - y \in C \\ &\Leftrightarrow \mathrm{H}(x-y)^t = 0 \\ &\Leftrightarrow \mathrm{H}x^t - \mathrm{H}y^t = 0 \\ &\Leftrightarrow \mathrm{H}x^t = \mathrm{H}y^t \\ &\Leftrightarrow s_x = s_y. \quad \square \end{aligned}$$

Decodificación del síndrome: Sea C un $[n, k]$-código, y matriz de control H. Para cada palabra e de longitud n se define.

$$e + C = \{e + c \mid c \text{ es un codeword}\}.$$

Lo cual interpretamos como la clase de todas las posibles palabras recibidas $w = e_i + c$, donde c es un codeword enviado a través de un canal ruidoso que produce un error e_i, toda palabra en la clase dada tiene el mismo síndrome como ya se demostró.

Ahora con el fin de decodificar, para cada síndrome es decir para cada palabra s de longitud $n - k$ escogemos un representante para la clase, esto es una palabra e_i de menor peso que tenga el síndrome s. Para resolver el sistema de ecuaciones lineales

$$He^t = s^t.$$

Lo que conduce al siguiente procedimiento de decodificación mediante la tabla de síndromes:

- Cuando recibimos un vector w, calculamos el síndrome $Hw^t = s^t$.
- Si el síndrome encontrado no es nulo, entonces se localiza el vector e_i correspondiente al síndrome calculado.
- Asumimos que el codeword enviado es $c = w - e_i$.

1.2.10 Ejemplo. Sea C un $[5, 3]$-código binario con matriz de control

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

Realizando el proceso descrito para calcular los representantes, obtenemos la tabla de síndromes,

Si recibimos, por ejemplo, el vector $w = 11111$ lo decodificaremos como sigue:

- Encontramos el síndrome

$$Hw^t = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = s_w.$$

Síndrome	Representantes de la clase
00	00000
11	10000
01	00001
10	00010

Figura 1.2: Tabla de representantes de las clases

- El representante con síndrome 11 es $e_2 = 10000$.
- Luego el error es $e = 10000$, entonces la palabra código es

$$c = w - e_2 = 11111 - 10000 = 01111.$$

1.3. Los códigos de Reed-Solomon

En la actualidad, los códigos Reed-Solomon se utilizan para corregir errores en varios sistemas, incluyendo los dispositivos de almacenamiento como: cintas, discos compactos, DVD, etc.
También tiene aplicaciones en las telecomunicaciones como: las inalámbricas o móviles, telefonía celular, enlaces de microondas, comunicaciones satelitales, televisión digital entre otros. Estos códigos, fueron presentados por, Irving S. Reed[2] y Gustave Solomon[3] en el año de 1960.

1.3.1 Definición. Sea K un cuerpo finito, digamos $|K| = q$. Sean $k, n \in \mathbb{N}$ con $1 \leq k \leq n \leq q$ y además

$$K[x]_k := \{f \in K[x] \mid grad(f) < k\}.$$

Es decir $K[x]_k$ denota el conjunto de todos los polinomios con coeficientes en K con grado estrictamente menor que k.

[2]Irving S. Reed(1923-)Matemático e ingeniero que ha realizado numerosas contribuciones en el área de la ingeniería incluyendo radar, procesamiento digital de señales y procesamiento digital de imágenes.

[3]Gustave Solomon(1930-1996) Matemático e ingeniero coinventor de la teoría algebraica de la corrección de errores. Obtiene su Ph.D. en Matemática del MIT en 1956 con la dirección de Kenkichi Iwasawa.

Sea $A = \{a_1, \ldots, a_n\} \subseteq K$, con $a_i \neq a_j$ para $i \neq j$. Definamos

$$C(A) := \{(f(a_1), \ldots, f(a_n)) \mid f \in K[x]_k\} \subseteq K^n.$$

Donde $C(A) \leq K^n$. Este subespacio se denomina ***código Reed- Solomon*** o simplemente un RS-código.

1.3.2 Teorema. Sean K un cuerpo finito, con $|K| = q$ y $1 \leq k \leq n \leq q$. Entonces $C(A)$ es un $[n, k, n-k+1]$-código sobre K.

Demostración. Claramente $C(A)$ tiene longitud n. Consideremos ahora la función

$$\alpha : K[x]_k \longrightarrow C(A)$$

definida por

$$\alpha(f) := (f(a_1), \ldots, f(a_n)).$$

Se verifica sin dificultades que α es un epimorfismo entre espacios vectoriales.

Demostremos ahora que α es inyectiva, por lo tanto es un isomorfismo. Si $f \in ker(\alpha)$, entonces

$$\alpha(f) = (f(a_1, \ldots, a_n)) = (0, \ldots, 0).$$

Dado que $grad(f) < k < n$, se sigue que f tiene a lo mas $k-1$ raíces distintas. Por lo tanto la igualdad anterior se verifica si y solo si $f = 0$.

Es claro que

$$B = (1, x, \ldots, x^{k-1})$$

es una base para $K[x]_k$. Por consiguiente

$$\dim_k K[x]_k = k = K[x]_k C(A).$$

Para calcular $d(C(A))$, consideraremos dos elementos cualesquiera de $C(A)$, digamos $c_f = (f(a_1), \ldots, f(a_n))$ y $c_g = (g(a_1), \ldots, g(a_n))$ y supongamos que la distancia $d(c_f, c_g) = t$. Entonces c_f y c_g coinciden en $n - t$ posiciones, por lo tanto el polinomio $f - g$ tiene $n - t$ raíces. Sabemos que $n - t \leq k - 1$. En consecuencia

$$n - k + 1 \leq t.$$

Esto significa que $n - k + 1$ es una cota inferior, se tiene que

$$n - k + 1 \leq d(C(A)).$$

Para el polinomio

$$f = \prod_{j=1}^{k-1}(x - a_j)$$

se verifica que $\omega t(c_f) = n - k + 1$. Por lo tanto

$$d(C(A)) = n - k + 1.$$

Con esto se concluye que $C(A)$ es un $[n, n-k+1]$-código sobre K.□

Este tipo de códigos resultan ser muy importantes en vista que alcanzan la cota Singleton, lo que hace que puedan detectar un mayor número de errores.

1.3.3 Teorema. Sean K un cuerpo finito, con $|K| = q$, $A = \{a_1, \ldots, a_n\} \subseteq K$, con $a_i \neq a_j$ para $i \neq j$ y $1 \leq k \leq n \leq q$. Entonces la matriz generadora de un $[n, k]$-código de Reed- Solomon está dada por la matriz de Vandermonde

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & & & \\ a_1^{k-1} & a_2^{k-1} & \cdots & a_n^{k-1} \end{pmatrix}.$$

Demostración.

$$\begin{aligned} c_f \in C(A) &\Leftrightarrow \exists f \in K[x]_k \text{ tal que } c_f = (f(a_1), \ldots, (a_n)) \\ &\Leftrightarrow \exists c_0, \ldots, c_{k-1} \in K, \text{ tales que } c_f = (\textstyle\sum_{j=0}^{k-1} c_j a_1^j, \ldots, \sum_{j=0}^{k-1} c_j a_n^j) \\ &\Leftrightarrow c_f = (c_0, \ldots, c_{k-1})G \\ &\Leftrightarrow c_f \in K^k G. \quad \square \end{aligned}$$

1.4. Algunas construcciones de códigos lineales

1.4.1 Definición. Sea C un $[n, k]-$código sobre un cuerpo finito K, con $|K| = q$. Llamaremos a

$$\widehat{C} := \{(c_1, \ldots, c_n, c_{n+1}) \mid (c_1, \ldots, c_n) \in C, \sum_{j=1}^{n+1} c_j = 0\} \subseteq K^{n+1}$$

el código **extensión** de C mediante un bit de control de paridad.

1.4.2 Teorema. Sea C un código $[n,k,d]-$código sobre un cuerpo finito K, con $|K| = q$. Entonces

(a) $\widehat{C}$ es un $[n+1,k,d]-$código sobre K, con $d \leq d' \leq d+1$.

(b) Si C es binario, entonces $d' = d+1$ si y solo si d es impar.

(c) Si H es una matriz de control para C, entonces

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

es una matriz de control para $\widehat{C}$.

Demostración.

(a) Claramente $\widehat{C}$ es un código lineal sobre K con longitud $n+1$. Dado que $|\widehat{C}| = |C| = q^k$, se tiene que $\dim_k \widehat{C} = k$. Para el cálculo de la distancia mínima, note que si $c = (c_1, \ldots, c_{n+1}) \in \widehat{C}$, entonces

$$\omega t(c) = \omega t(c_1, \ldots, c_n) + \omega t(c_{n+1}).$$

En consecuencia $d \leq d' \leq d+1$.

(b) Supongamos que C es binario y supongamos además que $d' = d+1$. Sea $c = (c_1, \ldots, c_n) \in C$ con $\omega t(c) = d$. Si d fuese un número par, entonces $\hat{c} = (c_1, \ldots, c_n, 0) \in \widehat{C}$. Por lo tanto

$$d' \leq \omega t(\hat{c}) = \omega t(c) = d.$$

Usando (a) se tendría que $d' = d$, lo cual es una contradicción.
Recíprocamente, supongamos que d es impar. Sea $\hat{c} \in \widehat{C}$ no nulo , digamos $\hat{c} = (c_1, \ldots, c_n, c_{n+1})$, con $\omega t(\hat{c}) = d\,'$. Definamos $c := (c_1, \ldots, c_n)$.

Caso1. Si $c_{n+1} = 0$, entonces $\omega t(\hat{c}) = \omega t(c)$ es par. Dado que d es impar, se tiene que existe $0 \neq y = (y_1, \ldots, y_n) \in C$ tal que

$$d = \omega t(y) < \omega t(c).$$

Entonces $\hat{y} = (y_1, \ldots, y_n, 1) \in \widehat{C}$ y se tiene que

$$\omega t(\hat{y}) = \omega t(y) + 1 \leq \omega t(c) = \omega t(\hat{c}) = d'.$$

En consecuencia

$$d' = \omega t(\hat{y}) = \omega t(y) + 1 = d + 1.$$

Caso2. Si $c_{n+1} \neq 0$, entonces $\omega t(c) < \omega t(\hat{c}) = d'$. Usando (a) se tiene que

$$d \leq \omega t(c) < d' \leq d + 1.$$

Por lo tanto

$$d' = d + 1.$$

(c) Sean $\hat{c} = (c_1, \ldots, c_n, c_{n+1}) \in K^{n+1}$ y $c = (c_1, \ldots, c_n) \in K^{n+1}$. Entonces

$$\hat{c} \in \widehat{C} \quad \Leftrightarrow \quad (Hc^t = 0 \text{ y } \sum_{j=1}^{n+1} c_j = 0)$$

$$\Leftrightarrow \quad \begin{pmatrix} 1 & \cdots & 1 & 1 \\ & & & 0 \\ & H & & \vdots \\ & & & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_{n+1} \end{pmatrix} = 0.$$

Con lo cual se tiene la afirmación deseada. □

1.4.3 Definición. Sean C un $[n, k]-$código sobre un cuerpo finito K, con $n \geq 2$ y $|K| = q$.

(a) Para $1 \leq i \leq n$. Definimos

$$\breve{C}(i) = \{(c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n) \mid (c_1, \ldots, c_{i-1}, 0, c_{i+1}, \ldots, c_n) \in C\}$$

y lo llamaremos la **reducción** de C en la $i-$ésima coordenada.

(b) Para $1 \leq i \leq n$. Definimos

$$\overset{\circ}{C}(i) = \{(c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n) \mid \exists c_i \in K (c_1, \ldots, c_n) \in C\}$$

y lo llamaremos la **perforación** de C en la $i-$ésima coordenada.

1.4.4 Ejemplo. Sea C el $[4, 2]-$código binario dado por

$$C = \{0000, 1011, 0110, 1101\}.$$

Entonces

1. $\overset{\circ}{C}(1) = \{(000, 011, 110, 101)\}$
2. $\overset{\circ}{C}(2) = \{(000, 111, 010, 101)\}$
3. $\breve{C}(1) = \{(000, 110)\}$
4. $\breve{C}(2) = \{(000, 111)\}$

1.4.5 Teorema. Sea C un $[n, k, d]-$código sobre un cuerpo finito K, con $|K| = q, n \geq 2$ y $k \geq 1$. Entonces

(a) $\breve{C}$ y $\overset{\circ}{C}$ son códigos lineales de longitud $n-1$ y $\breve{C} \subseteq \overset{\circ}{C}$.

(b) $k - 1 \leq \dim_K \breve{C}(i) \leq \dim_K \overset{\circ}{C}(i) \leq k$.

(c) Si $\breve{C} \neq \{0\}$, entonces $d(\breve{C}(i)) \geq d(\overset{\circ}{C}(i))$.

(d) $\dim_K \breve{C}(i) = k - 1$ si y solo si $(c_1, \ldots, c_n) \in C$, con $c_i \neq 0$

(e) $\dim_K \overset{\circ}{C}(i) = k - 1$ si y solo si $(0, \ldots, 0, c_i, 0, \ldots, 0) \in C$, con $c_i \neq 0$. En particular, si $d \geq 2$, entonces $\dim_K C^{\circ}(i) = k$.

f $\breve{C} = \overset{\circ}{C}$ si y solo si existe $(0, \ldots, 0, c_i, 0, \ldots, 0) \in C$, con $c_i \neq 0$ o para todo $(c_1, \ldots, c_n) \in C$ se verifica que $c_i = 0$.

(g) Si $\breve{C} \neq \{0\}$, entonces $d(\breve{C}(i)) \geq d$.

(h) Si $\dim_K \overset{\circ}{C}(i) = k$, entonces $d - 1 \leq d(\overset{\circ}{C}(i)) \leq d$.

Demostración.

(a) Se sigue inmediatamente de la definición 1.4.3.

(b) De (a) se sigue que

$$\dim_K \breve{C}(i) \leq \dim_K \overset{\circ}{C}(i) \leq k.$$

Si $c_i = 0$, para todo $c = (c_1, \ldots, c_n) \in C$, entonces

$$\dim_K \breve{C}(i) = \dim_K \overset{\circ}{C}(i) = k.$$

Sea $v_1 = (c_1, \ldots, c_n) \in C$, con $c_i \neq 0$. Entonces existe una base para C que contiene a v_1, digamos $B = (v_1, \ldots, v_k)$. Además existen escalares $\lambda_1, \ldots, \lambda_k \in K$ tales que cada uno de los vectores

$$v_2 - \lambda_2 v_1, \ldots, v_k - \lambda_k v_1$$

tiene un cero en la i-ésima posición. Note lo siguiente

$$c_2(v_2 - \lambda_2 v_1), \ldots, c_k(v_k - \lambda_k v_1) = 0$$
$$(-c_2\lambda_2 - \cdots - c_k\lambda_k)v_1 + c_2 v_2 + \cdots + c_k v_k = 0.$$

Como $v_1, \ldots, v_k$ son elementos de la base, se tiene que

$$c_2\lambda_2 = \cdots = c_k\lambda_k = 0.$$

Por lo tanto el conjunto

$$\{v_2 - \lambda_2 v_1, \ldots, v_k - \lambda_k v_1\}$$

es linealmente independiente. Con lo que se tiene que

$$\dim_K \breve{C}(i) \geq k - 1.$$

(c) Se sigue de (a).

(d) Usando (a) tenemos

$$\begin{aligned} \dim_K \breve{C}(i) = k - 1 &\Leftrightarrow \breve{C}(i) < C \text{ esto se sigue de (b)} \\ &\Leftrightarrow |\breve{C}(i)| < |C| \\ &\Leftrightarrow \exists (c_1, \ldots, c_n) \in C, \text{ con } c_i \neq 0. \end{aligned}$$

(e) Consideremos la proyección sobre la $i-$ésima coordenada

$$\pi : C \longrightarrow \overset{\circ}{C}(i)$$

definida por

$$\pi(c_1, \ldots, c_n) := (c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n).$$

Claramente π es un epimorfismo y además

$$\ker(\pi) = C \cap \{(0, \ldots, 0, c_i, 0, \ldots, 0) \mid c_i \in K\}.$$

Del primer teorema de isomorfía se sigue que

$$k - \dim_K \ker(\pi) = \dim_K C - \dim_K \ker(\pi) = \dim_K \breve{C}(i).$$

Entonces

$$\begin{aligned} \dim_K \breve{C}(i) = k-1 &\Leftrightarrow \dim_K \ker(\pi) = 1 \\ &\Leftrightarrow \exists (0,\ldots,0,c_i,0,\ldots,0) \in C \text{ con } c_i \neq 0. \end{aligned}$$

Si en particular, si $d \geq 2$, entonces no existe $(0,\ldots,0,c_i,0,\ldots,0) \in C$ con $c_i \neq 0$. Por lo tanto de (b) se sigue que $\dim_K \overset{\circ}{C}(i) = k$.

(f) Dado que $\breve{C} \subseteq \overset{\circ}{C}$, se tiene que

$$\begin{aligned} \breve{C} = \overset{\circ}{C} = k-1 &\Leftrightarrow \dim_K \breve{C}(i) = \overset{\circ}{C}(i) \\ &\Leftrightarrow \dim_K \breve{C}(i) = k \ \vee \ \dim \overset{\circ}{C}(i) = k-1. \end{aligned}$$

El resto se sigue de (d) y (e).

(g) Sea $0 \neq c = (c_1,\ldots,c_{i-1},c_{i+1},\ldots,c_n) \in \breve{C}$ con $\omega t(c) = d(\breve{C}(i))$. Entonces $c' = (c_1,\ldots,c_{i-1},0,c_{i+1},\ldots,c_n) \in C$ y se tiene que

$$d \leq \omega t(c') = \omega t(c) = d(\breve{C}(i)).$$

(h) Dado $k \geq 1$, se tiene que $C \neq \{0\}$. Sea $0 \neq c = (c_1,\ldots,c_n) \in C$ con $\omega t(c) = d$. Entonces $c' = (c_1,\ldots,c_{i-1},c_{i+1},\ldots,c_n) \in \overset{\circ}{C}(i)$ y $c' \neq 0$, ya que de lo contrario

$$C \cap \{(0,\ldots,0,c_i,0,\ldots,0) \mid c_i \in K\} \neq 0,$$

lo cual contradice el hecho que $\dim_K \overset{\circ}{C} = k$. Entonces

$$d(\overset{\circ}{C}(i)) \leq \omega t(c') \leq \omega t(c) = d. \tag{1.2}$$

Por otro lado, sea $0 \neq x' = (x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) \in \overset{\circ}{C}(i)$ con $\omega t(x') = d(\overset{\circ}{C}(i))$. Entonces existe

$$x = (x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_n) \in C,$$

con $x \neq 0$ y se sigue que

$$d \leq \omega t(x) \leq \omega t(x') + 1 = d(\overset{\circ}{C}(i)) + 1. \tag{1.3}$$

De (1.2) y (1.3) se sigue la afirmación. □

1.5. Dualidad

Sean K un cuerpo finito, $n \in \mathbb{N}$ y definimos

$$(\cdot \mid \cdot) : K^n \times K^n \longrightarrow K$$

mediante

$$(u \mid v) := \sum_{j=1}^{n} u_j v_j,$$

donde $u = (u_1, u_2, \ldots, u_n)$ y $v = (v_1, v_2, \ldots, v_n)$. Si $(u \mid v) = 0$, entonces decimos que u y v son ortogonales.

1.5.1 Ejemplo. Sea $u = 1101$ y $v = 1111$ entonces

$$(u \mid v) = 1 + 1 + 0 + 1 = 1$$

1.5.2 Definición. Sea C un código lineal $[n, k]$-código sobre un cuerpo finito K.

(a) El **código dual** de C, notado con $C^{\perp}$ se define como:

$$C^{\perp} = \{u \in K^n \mid (u \mid c) = 0, \forall c \in \mathrm{C}\}.$$

(b) C se denomina **auto-dual** si, $C = C^{\perp}$.

(c) C se denomina **auto-ortogonal** si, $C \subseteq C^{\perp}$.

1.5.3 Ejemplo.

(a) Sea $C = \{0000, 0011, 1100, 1111\}$, note que $C = C^{\perp}$.

(b) Sea $C = \{000, 011, 110, 101\}$, note que $C^{\perp} = \{000, 111\}$.

1.5.4 Teorema. Sea C un código lineal $[n, k]$-código sobre un cuerpo finito K. $|K| = q$. Entonces

(a) $C^{\perp}$ es un $[n, n-k]$-código sobre K.

(b) $(C^{\perp})^{\perp} = C$.

(c) G es una matriz generadora de C si y solo si G es una matriz de control de $C^{\perp}$.

(d) H es una matriz control de C si y solo si H es una matriz generadora de $C^{\perp}$.

(e) Si $(I_k \mid A)$ es una matriz generadora de C (en forma estándar), entonces $(-A \mid I_{n-k})$ es una matriz generadora de $C^{\perp}$, por lo tanto una matriz de control de C.

(f) Si C es auto-dual, entonces $n = 2k$. En particular, todo código auto-dual tiene longitud par.

Demostración.

(a) Sea G una matriz generadora para C, cuyas filas están dadas por

$$f_j = (g_{j1}, g_{j2}, \ldots, g_{jn}), j = 1, \ldots, k.$$

Entonces

$$\begin{aligned} x \in C^{\perp} &\Leftrightarrow (x \mid c) = 0, \text{ para todo } c \in C. \\ &\Leftrightarrow (x \mid v_j) = 0, \text{ para una base } B = (v_1, \ldots, v_k) \text{ de } C. \\ &\Leftrightarrow (x \mid f_j) = 0. \\ &\Leftrightarrow Gx^t = 0. \end{aligned}$$

Con lo cual se demuestra que G es una matriz de control de $C^{\perp}$. Del algebra lineal sabemos que

$$n = \text{Rang } G + \dim_k ker(G).$$

Es decir,

$$n = \text{Rang } G + \dim_k C^{\perp}.$$

De donde se sigue que $\dim_k C^{\perp} = n - k$.

(b) Es claro que $C \subseteq (C^{\perp})^{\perp}$. En efecto, si $x \in C$ y $v \in C^{\perp}$, entonces $(x \mid v) = 0$. Por lo tanto $x \in (C^{\perp})^{\perp}$.

Del inciso (a) se sigue que

$$\dim_k (C^{\perp})^{\perp} = n - (n - k) = k = \dim_k C.$$

Con lo cual se tiene la igualdad.

(c) En (a) se mostró ya la implicación de izquierda a derecha.

Recíprocamente, sea G una matriz de control de $C^{\perp}$. Dado que $C^{\perp}$ es un $[n, n-k]$-código sobre K, se verifica que toda matriz generadora de $C^{\perp}$ tiene $n-k$ filas. Por lo tanto cualquier matriz de control de $C^{\perp}$ tiene $n-(n-k)=k$ filas. Supongamos entonces que las filas de G están dadas por

$$f_j = (g_{j1}, \ldots, g_{jn}),\ j = 1, \ldots, k.$$

Entonces $(y \mid f_j) = 0$, para todo $y \in C^{\perp}$ y todo $j = 1, \ldots, k$. Es decir, $f_j \in (C^{\perp})^{\perp} = C$, para todo $y \in C^{\perp}$ y $j = 1, \ldots, k$.
Dado que las filas de G son linealmente independientes y además

$$\dim_k Cn - (n-k) = k,$$

con k el número de filas de G, se tiene que G es una matriz generadora de C.

(d) Usando (b) y (c) tenemos:

H es una matriz generadora de $C^{\perp}$ si y solo si H es una matriz de control de $(C^{\perp})^{\perp} = C$.

(e) Se verifica que:

$$(-A^t \mid I_{n-k})(I_k \mid A)^t = (-A^t \mid I_{n-k}) \begin{pmatrix} I_k \\ A^t \end{pmatrix} = -A^t I_k + I_{n-k} A^t = 0.$$

Se sigue entonces que $(-A^t \mid I_{n-k})$ es una matriz de control de C, y consecuentemente una matriz generadora de $C^{\perp}$.

(f) Si C es auto-dual, entonces C y $C^{\perp}$ tienen la misma dimensión sobre K. Por lo tanto $n-k=k$ y con eso se tiene la afirmación.□

1.5.5 Lema. Sea $C + D = \{u + v \in \mathrm{C} \mid u \in C, v \in D\}$, entonces

$$(C+D)^{\perp} = C^{\perp} \cap D^{\perp}$$

Demostración. Mostremos entonces la doble contenencia

(a) Sea $x \in (C+D)^{\perp}$, entonces $(x \mid u+v) = 0$, para todo $u+v \in (C+D)$, donde $u \in C, v \in D$ puesto que C y D son subespacios de $\mathbb{F}_q^n$, tenemos que:

$$\begin{aligned} 0 \in D \wedge 0 \in C &\Rightarrow u+0 \in (C+D) \ \wedge \ 0+v \in (C+D) \\ &\Rightarrow (x \mid u+0) = 0 \ \wedge \ (x \mid 0+v) = 0 \\ &\Rightarrow (x \mid u) = 0 \ \wedge \ (x \mid v) = 0 \\ &\Rightarrow x \in C^{\perp} \ \wedge \ x \in D^{\perp} \\ &\Rightarrow x \in C^{\perp} \cap D^{\perp} \end{aligned}$$

Por lo tanto $(C+D)^{\perp} \subseteq C^{\perp} \cap D^{\perp}$.

(b) Sea $x \in C^{\perp} \cap D^{\perp}$, entonces tenemos que:

$$\begin{aligned} x \in C^{\perp} \wedge x \in D^{\perp} &\Rightarrow \{(x \mid u) = 0, \forall u \in C\} \wedge \{(x \mid v) = 0, \forall \ v \in D\} \\ &\Rightarrow (x \mid u) + (x \mid u), \ \forall \ u+v \in (C+D) \\ &\Rightarrow (x \mid u+v) = 0, \ \forall \ u+v \in (C+D) \\ &\Rightarrow x \in (C+D)^{\perp} \end{aligned}$$

Por lo tanto

$$C^{\perp} \cap D^{\perp} \subseteq (C+D)^{\perp}.$$

Así que por (a) y (b) tenemos que

$$(C+D)^{\perp} = C^{\perp} \cap D^{\perp}. \square$$

1.6. Códigos de peso constante

Es posible distinguir una familia importante en esta clase de códigos, los cuales reciben el nombre de su creador el Matemático estadounidense Richard Hamming, el cual empezó a reflexionar sobre la posibilidad de diseñar códigos correctores de errores para usos prácticos; así que en 1948 surgen este tipo de códigos conocidos desde entonces como ***códigos de Hamming.*** El siguiente teorema demuestra la existencia de dichos códigos.

1.6.1 Teorema. Sea K un cuerpo finito, con $|K| = q$ y sean $k \in \mathbb{N}$ con $k \geq 2$ y $n = \frac{q^k-1}{q-1}$. Entonces existe un $[n, n-k, 3]$-código sobre K, el cual llamaremos ***código de Hamming.***

Demostración. Todo vector no nulo $v \in K^n$ define una recta que pasa por el vector cero. Es decir, define un subespacio vectorial de dimensión uno, generado por u. Esta recta esta dada por

$$\langle u \rangle = \{ku \mid k \in K\}.$$

El espacio vectorial K^k tiene $q^k - 1$ vectores no nulos y dado un vector no nulo $u \in K^k$ existen $q - 1$ múltiplos escalares no nulos de u.

Dado que cualquier vector de la forma ku, con $k \in K^\times$ define la misma recta de u, se sigue que existen exactamente

$$n = \frac{q^k - 1}{q - 1}$$

rectas distintas en K^k.

Sean $\langle h_1\rangle, \ldots, \langle h_n\rangle$ rectas y sea $H \in \mathrm{Mat}(k \times n, K)$ la matriz cuyas columnas son precisamente los vectores h_j. Esto es, $H = (h_1 \ldots h_n)$.

El código C con matriz de control H, es decir,

$$C = \{c \in K^n \mid Hc^t = 0\},$$

se denomina un ***código de Hamming***. Las columnas $h_1, \ldots, h_k$ puede elegirse de tal manera que $B = (h_1, \ldots, h_k)$ sea una base para K^k. Entonces Rang $(H) = K$ y se tiene que $\dim(C) = n - k$. Es decir, $|C| = q^{n-k}$.

Por otro lado cualquier par de columnas de H son linealmente independientes, pero de manera adecuada tres columnas resultan linealmente dependiente. Luego $\omega t(C) = 3$, como $\omega t(C) = d(C)$, entonces C tiene los parámetros $[n, n - k, 3]$.

En la construcción de C hemos elegido aleatoriamente, de un lado los representantes h_j en los subespacios vectoriales de dimensión uno y por otro lado, la ubicación de los h_j en la matriz H. Otras elecciones no suministran resultados esencialmente distintos, por ser códigos equivalentes. Esto nos permite hablar sin ambigüedades del $[\frac{q^k-1}{q-1}, n-k, 3]$-código de Hamming sobre K. □

En adelante usaremos la notación $\mathrm{Ham}_q(k)$ para referirnos a este código.

1.6.2 Ejemplo. Sea C un $[7, 4]$-código binario de Hamming. La matriz de control es

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

En los códigos de Hamming dado que $d = 3$ y en virtud del teorema 1.1.17 se tiene que este tipo de códigos son únicamente 1-correctores.

Códigos Simplex

El código simplex es el código dual del código de Hamming. Lo notaremos con $\mathrm{sim}_q(k)$.

1.6.3 Lema. Sea K un cuerpo finito, con $|K| = q$.

(*a*) Si $0 \neq c \in \mathrm{sim}_q(k)$, entonces $\omega t(c) = q^{k-1}$. Es decir, todos los codewords no nulos del código Simplex tienen el mismo peso.

(*b*) $\mathrm{sim}_q(k)$ es un $[\frac{q^k-1}{q-1}, k, q^{k-1}]-$código sobre K.

Demostración. Sea H una matriz de control de $\mathrm{Ham}_q(k)$, con filas $f_1, \ldots, f_k$. Sea además $0 \neq c = (c_1, \ldots, c_n) \in \mathrm{sim}_q(k)$. Dado que H es una matriz generadora de $\mathrm{sim}_q(k)$ se tiene que $B = (f_1, \ldots, f_k)$ es una base para $\mathrm{sim}_q(k)$. Por lo tanto

$$c = \sum_{j=1}^{k} a_j f_j = \sum_{j=1}^{k} a_j (f_{j1}, \ldots, f_{jk}),\ a_j \in K.$$

Sea $h_i := (f_{1i}, \ldots, f_{ki})^t$ la i-ésima columna de H, $a := (a_1, \ldots, a_k) \in K^k$ y definamos el conjunto $U(a)$ de la siguiente manera:

$$U(a) := \{(b_1, \ldots, b_k)^t \mid b_j \in K, \sum_{j=1}^{k} a_j b_j = 0\} \subseteq K^k.$$

Se verifica inmediatamente que $U(a) = \langle a \rangle^{\perp}$. Por lo tanto

$$\dim_K U(a) = \dim_K \langle a \rangle^{\perp} = k - \dim_K \langle a \rangle = k - 1.$$

En consecuencia $U(a)$ tiene $\frac{q^{k-1}-1}{q-1}$ columnas h_i de H. Note que

$$c_i = 0 \Leftrightarrow \sum_{j=1}^{k} a_j f_{ji} = 0 \Leftrightarrow (f_{1i}, \ldots, f_{ki})^t \in U(a).$$

Esto demuestra que en c hay exactamente $\frac{q^{k-1}-1}{q-1}$ ceros. Entonces

$$\omega t(c) = n - \frac{q^{k-1}-1}{q-1} = \frac{q^k-1}{q-1} - \frac{q^{k-1}-1}{q-1} = q^{k-1}.$$

Por lo tanto un código Simplex tiene parámetros $[\frac{q^k-1}{q-1}, k, q^{k-1}] = [n, k, q^{k-1}]$. Más aún, para todo $c \in \mathrm{sim}_q(k)$ no nulo, se tiene que $\omega t(c) = q^{k-1}$. □

1.6.4 Ejemplo. Consideremos un código Simplex binario con parámetros $[2^k - 1, k, 2^{k-1}]$. Para $k = 2$, el $\mathrm{sim}_2(2)$ es un $[3, 2, 2]$-código generado por

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

Luego $\mathrm{sim}_2(2) = \{000, 011, 101, 110\}$ es claro que $d_2 = 2 = 2^{2-1}$.

Capítulo 2

Códigos de red lineales

Antes de iniciar este capitulo haremos algunas observaciones con relación a la notación que se implementará en adelante.

Sea $q \geq 2$ potencia de un número primo. Asumiremos a menos que se indique lo contrario que todos los vectores y matrices considerados tienen entradas en el cuerpo finito $\mathbb{F}_q$. Utilizamos $\mathrm{Mat}(n \times m, \mathbb{F}_q)$ para denotar el conjunto de todas las matrices de tamaño $n \times m$ sobre $\mathbb{F}_q$, cada vez que tengamos $\mathrm{Mat}(n \times 1, \mathbb{F}_q)$ lo entenderemos como $\mathbb{F}_q^n$. En particular, $v \in \mathbb{F}_q^n$ denota un vector columna y $v \in \mathbb{F}_q^{1 \times m}$ denota un vector fila. Si A es una matriz, entonces el símbolo A_i denota la i-ésima fila o la columna i-ésima de A; la distinción siempre estará clara a partir de la forma en la que A se define. En cualquier caso, el símbolo A_{ij} siempre se refiere a la entrada en la fila i y la j-ésima columna de A. Establecemos también que $I = I_n$, es la matriz identidad de orden n, cada vez que se use la notación I_i se entenderá como la i-ésima columna de I.

2.1. Introducción.

Los códigos de red presentan una ampliación en la visión clásica de la teoría de códigos, y se espera que sean una alternativa de transmisión de mensajes mas rápida y segura. En el capitulo 1, se habló de que el papel de la teoría de códigos es permitir la transmisión de información libre de error. Es sumamente interesante tratar de replicar los principales resultados de la teoría clásica en la teoría de códigos de red.

En la comunicación de red la multidifusión[1] de información hace necesario comprimir la información en el nodo de origen organizando y enrutando los paquetes de datos para luego enviarlos al nodo receptor a través de los nodos intermedios en la red. En el caso en que haya más de un nodo receptor, la información debe ser replicada en ciertos nodos intermedios de modo que cada nodo receptor puede recibir una copia de la información. Este método de transmisión de información en una red se conoce generalmente como ***store-and-forward or routing*** (almacenamiento-y-reenvío o enrutamiento). De hecho, casi todas las redes de ordenadores construidas en las últimas décadas se basan en este principio, donde los enrutadores se han desplegado en los nodos intermedios para transmitir un paquete de datos desde un canal de entrada hasta un canal de salida sin procesar el contenido de los datos. En la teoría de códigos de red, los nodos intermedios almacenan y transmiten la información entrante y estos nodos a su vez tienen la opción de mezclar dicha información antes de transmitirla.

La codificación en general, debe ser realizada en los nodos intermedios con el fin de tener una tasa de flujo de información mayor. Esta es una característica principal de la codificación en red. Usaremos un grafo dirigido finito para representar la red de comunicación, recordemos entonces que un grafo dirigido es una dupla ordenada $(V(G), E(G))$, donde $V(G)$ es un conjunto finito y no vacío de vértices y $E(G)$ es el conjunto de arcos, un arco $e \in E(G)$ es un par ordenado (u, v) de vértices, u y v; el arco con origen en u y extremo en v.

Un ejemplo bajo el contexto de la teoría de códigos de red se dará a continuación, cabe mencionar que cada vez que hablemos de un nodo en la red este corresponderá a un vértice en el grafo, mientras que un canal de comunicación corresponderá a un arco del grafo, por lo que no haremos distinción entre un nodo y un vértice, ni vamos a distinguir un canal de un arco. En el grafo, un nodo es representado por un circulo.

Para referirnos a un canal en particular haremos uso de los nodos, haciendo mención de los nodos implicados, por ejemplo en la figura 2.1 suponga que quiere mencionar el canal que va del nodo $T_1 \in V(G)$ al nodo $I \in V(G)$, ese canal lo notaremos entonces como $(T_1, I) \in E(G)$.

[1] multidifusión:= transmitir información desde un nodo de origen a un conjunto especifico de nodos.

Consideremos el siguiente ejemplo de una ¨red mariposa¨.

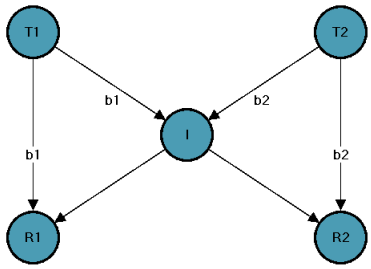

Figura 2.1: Red mariposa con capacidad de 1bit.

En esta red, dos bits b_1 y b_2 son enviados simultáneamente por T_1 y T_2 respectivamente, se requiere que R_1 y R_2 reciban ambos bits.

Note que al enviar b_1 en el canal (T_1, I) y b_2 por el canal (T_2, I) llegan a un único nodo de salida I. El nodo intermedio I no puede mandar simultáneamente el bit b_1 y el b_2, entonces debemos escoger uno de los dos bits, supongamos entonces que escogemos y enviamos el bit b_1 (ver figura 2.2), entonces dos copias de b_1 son recibidas y esto hace que b_2 no pueda ser recuperado o que sea recibido con un retraso de tiempo.

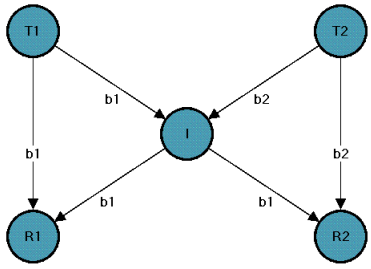

Figura 2.2: Red mariposa (esquema clásico).

Sin embargo en la codificación en red es posible que el nodo I combine b_1 y b_2 y transmita mediante los canales (I, R_1) y (I, R_2) el bit $b_1 + b_2$, donde $+$ denota una adición modulo 2; entonces ambos receptores R_1 y R_2, obtienen simultáneamente la información de ambos bits pero en este caso si R_1 recibe

a b_1 y a $b_1 + b_2$, b_2 puede ser recuperado, por ejemplo: R_1 calcula b_2 de:

$$b_1 + (b_1 + b_2) = (b_1 + b_1) + b_2 = 0 + b_2 \quad = \quad b_2.$$

El anterior ejemplo establece la ventaja de la codificación en red sobre el

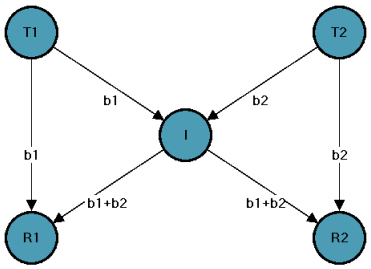

Figura 2.3: Red mariposa (esquema de codificación en red).

encaminamiento con respecto a la decodificación y a la tasa de flujo de información.Ahora procederemos a formalizar definiciones; algunas de las cuales tienen su equivalente en la teoría de códigos clásica.

El modelo del canal operador

Comencemos por la formulación del problema en el caso de una única difusión ***(unicast)***, es decir, la comunicación entre un único emisor y un único receptor. La generalización de la multidifusión ***(multicast)*** es sencilla.

En la codificación de red lineal aleatoria, el transmisor envía una serie de paquetes de longitud fija en la red, cada uno de los cuales puede considerarse como un vector fila de longitud N sobre un campo finito $\mathbb{F}_q$. Estos paquetes se propagan por la red, pasando posiblemente a través de una serie de nodos intermedios entre el transmisor y el receptor. Siempre que un nodo intermedio tiene la oportunidad de enviar un paquete, se crea una combinación lineal sobre $\mathbb{F}_q$ de los paquetes que dispone y transmite esta combinación aleatoriamente. Finalmente, el receptor recibe tales paquetes generados aleatoriamente y trata de decidir cual fue el conjunto de paquetes enviados por medio de la red. Aquí no hay suposición alguna de que la red opera sincrónicamente o sin retraso o que la red es acíclica. El conjunto de paquetes transmitidos con éxito en una generación induce un multigrafo dirigido, es decir, un grafo en el que hay pares de vértices unidos por más de una arista, en la que las aristas o bordes denotan transmisiones exitosas de paquetes. La tasa de transmisión de

información (paquetes por generación) entre el emisor y el receptor esta acotado superiormente por el menor corte ***(min-cut)*** entre los nodos existentes, donde el mínimo corte corresponde a la capacidad mínima de corte efectuada en un grafo, teniendo en cuenta que un corte en un grafo es la eliminación de arcos que separan completamente el nodo de origen del nodo de destino, de forma equivalente esto puede entenderse como el mínimo número de borraduras de la arista en el grafo que causaría la separación entre el emisor y el receptor.
Sea $\{p_1, \cdots, p_M\}$, $p_i \in \mathbb{F}_q^N$, denotado como el conjunto de vectores enviados. En el caso en que se este libre de errores, el receptor recibe paquetes y_j, $j = 1, \cdots, L$ donde cada y_j es formado de la siguiente manera

$$y_j = \sum_{i=1}^{M} h_{j,i} p_i,$$

con coeficientes aleatorios $h_{j,i} \in \mathbb{F}_q$. Para ilustrarlo, supongamos que $h_{j,i} = 1$ y $P = \{P_1, P_2, P_3\}$, cuya codificación se representa en el siguiente grafo.

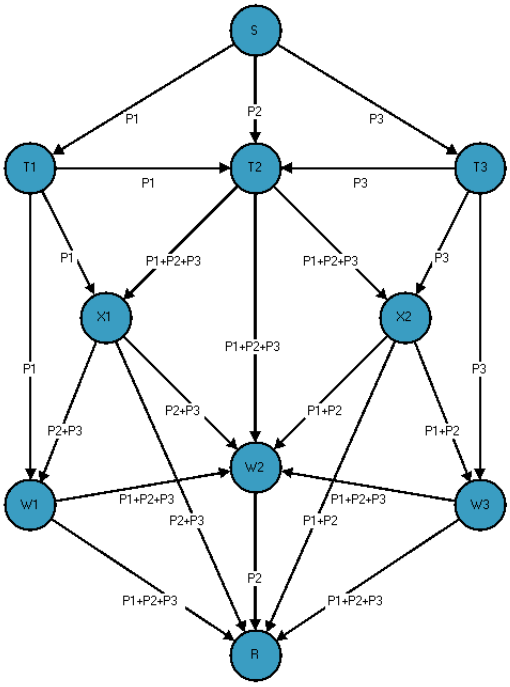

Figura 2.4: Esquema de codificación en red libre de error.

Obteniendo así que:

$$\begin{aligned} y_1 &= P_1 + P_2 + P_3 \\ y_2 &= P_2 + P_3 \\ y_3 &= P_2 \\ y_4 &= P_1 + P_2 \\ y_5 &= P_1 + P_2 + P_3. \end{aligned}$$

Note que L no es fijo y el receptor normalmente recoge tantos paquetes como le sea posible. Sin embargo, como se señaló anteriormente las propiedades de la red tales como la del mínimo corte entre el transmisor y el receptor pueden influir en la distribución conjunta de la $h_{i,j}$ y, en algunos puntos no habrá ningún beneficio a partir de la recopilación de información adicional y redundante.

Si optamos por considerar el envió de paquetes erróneos T, este modelo se amplia para incluir paquetes de error e_t, $t = 1, \cdots, T$, donde donde nuevamente $g_{j,t} \in \mathbb{F}_q$ son coeficientes aleatorios y desconocidos, entonces

$$y_j = \sum_{i=1}^{M} h_{j,i} p_i + \sum_{i=1}^{T} g_{j,t} e_t,$$

para ilustrarlo seguiremos considerando el anterior ejemplo.

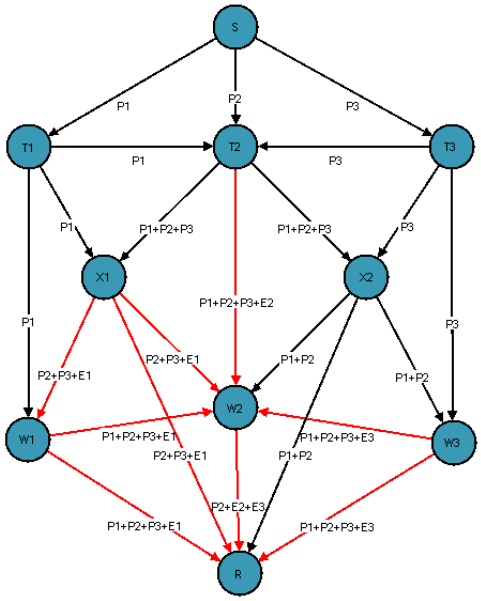

Figura 2.5: Esquema de codificación en red con error.

Se tiene entonces que

$$\begin{aligned} y_1 &= P_1 + P_2 + P_3 + E_1 \\ y_2 &= P_2 + P_3 + E_1 \\ y_3 &= P_2 + E_3 \\ y_4 &= P_1 + P_2 \\ y_5 &= P_1 + P_2 + P_3 + E_3, \end{aligned}$$

es importante tener en cuenta que estos paquetes erróneos se pueden enviar en cualquier lugar dentro de la red y pueden causar un error extendido (fenómeno de propagación del error en la red). En particular, si $g_{j,1} \neq 0$ para todo j, un único error de paquete e_1 tiene el potencial para afectar todos y cada uno de los paquetes recibidos.

En forma matricial, el modelo de transmisión puede ser escrito como

$$y = Hp + Ge, \tag{2.1}$$

donde $H \in \mathrm{Mat}(L \times M, \mathbb{F}_q)$, $G \in \mathrm{Mat}(L \times T, \mathbb{F}_q)$, $p \in \mathrm{Mat}(M \times N, \mathbb{F}_q)$, $y \in \mathrm{Mat}(N \times L, \mathbb{F}_q)$ y $e \in \mathrm{Mat}(T \times N, \mathbb{F}_q)$, en donde las filas de las matrices p, y y e representan los vectores de transmisión, los vectores recibidos y los vectores de error respectivamente.

En este punto dado que H es escogida aleatoriamente, podemos preguntarnos, ¿qué propiedad de la secuencia de envío de paquetes permanece invariante en el canal descrito por (2.1), incluso en la ausencia de ruido ($e = 0$)?. Como H es una matriz aleatoria todo lo que es fijado por el producto de Hp es el espacio de filas de p. En efecto, en cuanto a lo que el receptor se refiere, cualquiera de los posibles conjuntos en el espacio serán equivalentes. Esto nos lleva por lo tanto, a considerar la transmisión de información a través de la no elección de p, sino más bien por la elección del espacio vectorial generado por las filas de p. Esta simple observación es el centro de los modelos de canales y de las estrategias de transmisión consideradas en este trabajo.

Con relación al espacio vectorial seleccionado por el transmisor, el único efecto perjudicial que una multiplicación con H puede tener es que Hp puede tener un menor rango que p, debido a borraduras, paquetes de borraduras o a un mínimo corte insuficiente, en cuyo caso Hp genera un subespacio de el espacio generado por las filas de p, es decir, que el objetivo será identificar p a partir de su generado, por tal motivo ahora se consideran subespacios como elementos de los códigos de red.

2.2. Los parámetros de un código de red lineal

Sea K $= \mathbb{F}_q$ y W un espacio vectorial finito dimensional sobre un cuerpo K y $dim(W) = N$.

2.2.1 Definición. Un ***canal operador*** para el espacio ambiente W es un canal con alfabeto de entrada y alfabeto de salida $\mathbb{P}(W)$, donde

$$\mathbb{P}(W) = \{V \mid V \leq W\}.$$

Es decir $\mathbb{P}(W)$ es el conjunto de todos los subespacios de W, a menudo $\mathbb{P}(W)$ también es llamado espacio proyectivo de W.

En particular, si V es la entrada de un canal operador y U es la salida, entonces como ya se menciono en la observación anterior U puede ser expresado como:

$$U = \mathcal{H}_k(V) \oplus E$$

donde $k = \dim(U \cap V)$ y E es el espacio error. Se dice además que el canal comete: $\varrho = dimV/(U \cap V)$ borraduras y $t = \dim(E)$ errores.

Es preciso tener en cuenta que hemos elegido como modelo el espacio de error E como la intersección trivial con el subespacio transmitido V, por lo tanto la elección de la e no es independiente de V. Sin embargo, si tuviéramos que modelar el espacio recibido como $U = \mathcal{H}_k(V)+E$ por un espacio arbitrario de error E, entonces, ya que siempre se descompone E para algunos espacios E' como $E = (E \cap V) \oplus E'$, se pueden conseguir $U = \mathcal{H}_k(V) + (E \cap V) \oplus E' = \mathcal{H}'_k(V) \oplus E'$ para algún $k' > k$. En resumen, un canal operador toma de un espacio vectorial y pone a cabo otro espacio vectorial, posiblemente con borraduras, es decir, supresión de vectores desde el espacio de transmisión o con errores, en ese caso con vectores adicionales a los vectores de transmisión al espacio.

2.2.2 Definición. Un código lineal de red C es un subconjunto no vacío de $\mathbb{P}(W)$. Cada elemento de C es llamado codeword.

Definimos la siguiente función de distancia

$$d : \mathbb{P}(W) \times \mathbb{P}(W) \longrightarrow \mathbb{Z}^+$$

por:

$$d(U, V) := \dim(U + V) - \dim(U \cap V),$$

como $d(U+V) := \dim U + \dim V - \dim(U \cap V)$, la anterior definición puede ser también escrita como se indica a continuación:

$$\begin{aligned} d(U,V) &= \dim(U+V) - \dim(U \cap V) \\ &= \dim U + \dim V - \dim(U \cap V) - \dim(U \cap V) \\ &= \dim U + \dim V - 2\dim(U \cap V). \end{aligned}$$

El siguiente lema es de suma importancia para el diseño de códigos en el canal operador que previamente definimos.

2.2.3 Lema. la distancia d define una métrica sobre $\mathbb{P}(W)$, esto es para $U, V, X \in \mathbb{P}(W)$ tenemos:

(a) $d(U,V) \geq 0$ y $d(U,V) = 0$ si y solo si $U = V$.

(b) $d(U,V) = d(V,U)$.

(c) $d(U,V) \leq d(U,X) + d(X,V)$.

Demostración. La demostración de los dos primeros incisos es inmediata, así que demostraremos solo el último de ellos.
Veamos inicialmente que:

$$\begin{aligned} \frac{1}{2}\left[d(U,V) - d(U,X) - d(X,V)\right] &= \frac{1}{2}\left[\dim U + \dim V - 2\dim(U \cap V)\right. \\ &\quad - \dim U - \dim X + 2\dim(U \cap X) \\ &\quad \left. - \dim X - \dim V + 2\dim(X \cap V)\right] \\ &= -\dim(U \cap V) - \dim X \\ &\quad + \dim(U \cap X) + \dim(X \cap V) \\ &= -\dim(U \cap V) - \dim X \\ &\quad + \dim(U \cap X) + \dim(X \cap V) \\ &\quad - \dim(U \cap X \cap X \cap V) \\ &\quad + \dim(U \cap X \cap X \cap V) \\ &= [\dim(U \cap X + X \cap V) - \dim X] \\ &\quad + [\dim(U \cap V \cap X) - \dim(U \cap V)], \end{aligned}$$

donde el primer sumando de la ultima igualdad es menor o igual que cero dado que $U \cap X + X \cap V \subseteq X$ y en el segundo sumando también se obtiene que es menor o igual que cero, ya que $U \cap V \cap X \subseteq U \cap V$. Por lo tanto podemos concluir que

$$d(U,V) - d(U,X) - d(X,V) \leq 0. \quad \square$$

2.2.4 Definición. El número de elementos de un código C, donde $C \subseteq \mathbb{P}(W)$ se denota por $|C|$. La distancia mínima de C es definida como:

$$D(C) := \text{mín}\{d(U,V) \mid U,V \in C,\ U \neq V\}.$$

La máxima dimensión de C es definido por:

$$\ell(C) := \text{máx}\{\dim(V) \mid V \in C\}.$$

Si la dimensión de cada codeword de C es constante, entonces se dice que C es un código de ***dimensión constante***

2.2.5 Definición. Sea $C \subseteq \mathbb{P}(W)$. Si D es la distancia mínima de C, entonces llamamos a C un $[N, \ell(C), |C|, D]_q$-código sobre $\mathbb{F}_q$.

Note que la anterior definición es análoga a la notación usada para describir los parámetros de un código lineal en la teoría clásica de códigos, con $[n,k,d]$-código donde la longitud del código C era notada como n, k hacia referencia a la dimensión y para la distancia mínima usábamos d.

Corrección de errores y borraduras

Un decodificador de distancia mínima para un código C, toma $V \in C$ lo más cercano posible a $U \in \mathbb{P}(W)$ con respecto a la métrica d, es decir, si $V \in C$ se cumple que $d(U,V) \leq d(U,V')$, para todo $V' \in C$. Claramente V no es único. Este proceso es conocido como decodificación mediante el vecino más cercano o decodificación mediante el vecino más proximo.

La importancia de la distancia mínima $D(C)$ para el código $C \subseteq \mathbb{P}(W)$ es dada en el proximo teorema, el cual proporciona la capacidad de combinación de errores y capacidad de corrección de borraduras bajo la decodificación de la distancia mínima. Antes definamos $(x)_+$, como $(x)_+ := \text{máx}\{0, x\}$.

2.2.6 Definición. Sea C un código. Supongamos que $V \in \mathbb{P}(W)$ es transmitido a través de un canal operador y $U \in \mathbb{P}(W)$ es recibido se dice que el canal comete

$$\begin{aligned} t &= \dim U/(U \cap V) \text{ errores y} \\ \varrho &= \dim V/(U \cap V) \text{ borraduras.} \end{aligned}$$

2.2.7 Teorema. Sea $C \subseteq \mathbb{P}(w)$ un código. Supongamos que $V \in C$ es transmitido a través de un canal operador y

$$U = \mathcal{H}_k(V) \oplus E$$

es recibido, con $t = \dim(E)$ errores, sea $\varrho = (\ell(C) - k)_+$, denotado como el máximo número de borraduras producidos por el canal. Si

$$2(t + \varrho) < D(C),$$

entonces un decodificador sera capaz de obtener V a partir de U.

Demostración. Dado que U puede ser expresado de la siguiente manera,

$$U = \mathcal{H}_k(V) \oplus E \text{ con } k \leq \dim V.$$

Note que

$$\begin{aligned} d(V, \mathcal{H}_k(V)) &= \dim V + \dim_k(V) - 2\dim \mathcal{H}_k(V) \\ &= \dim V - \dim \mathcal{H}_k(V) \\ &= \dim V - \dim(V \cap U) \\ &= \dim V/(V \cap U) \leq \varrho \end{aligned}$$

y también que

$$\begin{aligned} d(U, \mathcal{H}_k(V)) &= \dim U + \dim \mathcal{H}_k(V) - 2\dim \mathcal{H}_k(V) \\ &= \dim U - \dim \mathcal{H}_k(V) \\ &= \dim U - \dim(V \cap U) \\ &= \dim U/(V \cap U) = t. \end{aligned}$$

Entonces por la desigualdad triangular y lo mostrado previamente se tiene lo siguiente que

$$d(U, V) = d(V, \mathcal{H}_k(V)) + d(U, \mathcal{H}_k(V)) \leq \varrho + t.$$

Si $X \neq V$ y $X, V \in C$, entonces

$$D(C) \leq d(V, X) \leq d(V, U) + d(U, X),$$

Por lo tanto

$$d(U, X) \leq D(C) - d(V, U) = D(C) - (\varrho + t)$$

como $2(\varrho + t) \leq D(C)$, entonces

$$t + \varrho < \frac{D(C)}{2},$$

con lo que se concluye que $d(U, X) > t + \varrho \geq d(U, V)$. □

2.2.8 Observación. Si C es un código clásico con distancia mínima

$$D(C) > 2e + a.$$

Entonces C puede corregir e errores y a borraduras. Esto significa que las borraduras cuestan menos que los errores en contraste con los códigos de red este costo es igual para errores y borraduras.

Ahora si estuviéramos frente al caso en que la red no produjera errores, es decir, si el espacio error $E = \{0\}$, entonces se obtendría el siguiente corolario, consecuencia inmediata del teorema anterior.

2.2.9 Corolario Sea $C \subseteq \mathbb{P}(w)$ un código. Supongamos que $V \in C$ es transmitido a través de un canal operador. Si

$$U = \mathcal{H}_{\dim(W)}(V) \oplus E = V \oplus E$$

es recibido, y si $2t < D(C)$, donde $\dim(E) = t$, entonces un decodificador sera capaz de obtener V de U. Similarmente si $2\varrho < D(C)$, con $\varrho = (\dim(C) - k)_+$, entonces una decodificación por mínima distancia para C produce a V.

En otras palabras, la primera parte del corolario declara que en la ausencia de borraduras un decodificador de distancia mínima únicamente corrige errores por encima de la dimension

$$\left\lfloor \frac{D(C) - 1}{2} \right\rfloor,$$

análogo a la habitual situación de corregir un error.

2.3. Códigos de dimension constante.

En el contexto de la codificación de red, es natural considerar códigos en los cuales todos los codewords tengan la misma dimensión, este tipo de códigos son análogos a los códigos de peso constante en el espacio de Hamming (en el que cada codeword tiene un mismo peso). Además el conocimiento de la dimensión del codeword puede ser usada por el decodificador para iniciar el proceso de decodificación.

Una definición formal es la siguiente:

2.3.1 Definición. Sea $\mathbb{P}(W,\ell) := \{V \mid V \in \mathbb{P}(W),\ \dim V = \ell\}$ y $C \neq \emptyset$ un subconjunto de $\mathbb{P}(W)$, C es llamado un código de red de dimensión constante si $C \subseteq \mathbb{P}(W,\ell)$, donde el conjunto $\mathbb{P}(W,\ell)$ es llamado la $\ell-$**grassmaniana**.

2.3.2 Observación.

(a) Si $C \subseteq \mathbb{P}(W,\ell)$, entonces $2 \mid d(U,V)$ para todo $U,V \in C$. En efecto

$$\begin{aligned} d(U,V) &= \dim U + \dim V - 2\dim(U \cap V) \\ &= \ell + \ell - 2\dim(U \cap V) \\ &= 2\ell - 2\dim(U \cap V) \\ &= 2(\ell - \dim(U \cap V)). \end{aligned}$$

(b) Si $C \subseteq \mathbb{P}(W,\ell)$, entonces $D(C) \leq 2\ell$ o $D(C) = 2\ell$ si y sólo si $U \cap V = \{0\}$ para todo $U,V \in C$ con $U \neq V$, observe que esto es una consecuencia inmediata de (a).

2.3.3 Ejemplo. Sean $W \leq \mathbb{F}_q^n$ y $C = \{U_i \mid i = 1,2,\ldots,|C|\} \subseteq \mathbb{P}(W,\ell)$ con matriz generadora

$$G(U_i) = (I \mid A_i),$$

donde $I \in \mathrm{Mat}(\ell \times \ell, \mathbb{F}_q)$ y la matriz $A \in \mathrm{Mat}(\ell \times (N-\ell), \mathbb{F}_q)$. Es fácil ver que cada $G(U_i)$ genera $G(U_i) \neq G(U_j)$, $i \neq j$ con $i,j \in \{1,\ldots,|C|\}$ que al intersecarlos generan espacios de dimension a lo mas $\ell - 1$, por lo tanto la distancia mínima del código es

$$2\ell - 2(\ell - 1) = 2.$$

Esté ejemplo corresponde a un $[N,\ell,\ell(N-\ell),2]$-código.

2.3.4 Teorema. Sean $W \leq \mathbb{F}_q^n$ y $C \subseteq \mathbb{P}(W,\ell)$ y $D(C) = 2\ell$, entonces

$$|C| \leq \frac{q^N - 1}{q^\ell - 1}.$$

En particular, si se da la igualdad $|C| = \dfrac{q^N - 1}{q^\ell - 1}$, entonces $\ell \mid N$.

Demostración. Si $C = \{V_i \mid i = 1,2,\ldots,r\}$, entonces $\bigcup_{i=1}^r V_i^\times \subseteq W^\times$, donde $V_i^\times = V_i \backslash \{0\}$ y $W^\times = W \backslash \{0\}$. Por lo tanto las posibles combinaciones

sin el elemento 0 son cada subespacio de $q^\ell - 1$ elementos.

$$\begin{aligned} r(q^\ell - 1) &\leq q^N - 1 \\ &\Leftrightarrow \quad r \leq \frac{q^N - 1}{q^\ell - 1} \\ &\Leftrightarrow \quad |C| \leq \frac{q^N - 1}{q^\ell - 1}. \end{aligned}$$

Ahora si, $|C| = \dfrac{q^N - 1}{q^\ell - 1} \in \mathbb{N}$, entonces de la teoría de números, $\ell \mid N$. □

2.3.5 Ejemplo. (Código de extensión) Sea $\ell \mid N$, entonces

$$L = \mathbb{F}_{q^\ell} \leq E = \mathbb{F}_{q^N} = W,$$

por lo tanto

$$L^\times = \mathbb{F}_{q^\ell}^\times \leq E^\times = \mathbb{F}_{q^N}^\times = W.$$

Sea $E^\times = \dot{\bigcup}_{i=1}^{r} a_i L^\times$, donde $\dfrac{q^N - 1}{q^\ell - 1}$ y $a_1, a_2, \ldots, a_r$ es un conjunto transversal de $L^\times$ sobre $E^\times$, claramente, $V_i = a_i L^\times \cup \{0\}$ para $i \neq j$, luego

$$C = \{V_i \mid i = 1, \ldots, r\} \subseteq \mathbb{P}(W, \ell),$$

$D(C) = 2\ell$ y $|C| = \dfrac{q^N - 1}{q^\ell - 1}$, donde C alcanza la cota dada en el teorema anterior.

2.4. Dualidad.

Si fijamos una base para W, entonces del hecho de que W sea un espacio vectorial N-dimensional sobre $\mathbb{F}_q$, los elementos de W pueden ser representados por N-tuplas de $\mathbb{F}_q$ con valores de coordenadas respecto a la base fijada. Por lo tanto tomamos el producto interno usual entre los vectores $u = (u_1, \ldots, u_N)$ y $v = (v_1, \ldots, v_N)$ como $(u, v) = \sum_{i=1}^{N} u_i v_i$.

2.4.1 Definición. Sea $U \in \mathbb{P}(W)$, entonces el subespacio ortogonal

$$U^\perp := \{v \in W : (u, v) = 0 \text{ para todo } u \in U\}$$

es un espacio de dimensión $N - k$.

Como bien sabemos, para cualquier subespacios U y V de W se tiene que: $(U^\perp)^\perp = U$, además $(U+V)^\perp = U^\perp \cap V^\perp$ y $(U \cap V)^\perp = U^\perp + V^\perp$. Teniendo esto en cuenta procedemos a enunciar el siguiente lema, el cual establece que la distancia entre los subespacios U y V está perfectamente reflejada entre la distancia de los subespacios ortogonales $U^\perp$ y $V^\perp$.

2.4.2 Lema. $d(C) = d(C^\perp)$

Demostración. Sea $U, V \in C$

$$\begin{aligned} d(U^\perp, V^\perp) &= \dim U^\perp + \dim V^\perp - 2\dim(U^\perp \cap V^\perp) \\ &= \dim U^\perp + \dim V^\perp - 2\dim((U+V)^\perp) \\ &= \dim U^\perp + \dim V^\perp - 2(N - \dim(U+V)) \\ &= N - \dim U + N - \dim V - 2N + 2\dim(U+V) \\ &= -\dim U - \dim V + 2(\dim U + \dim V - \dim(U \cap V)) \\ &= \dim U + \dim V - 2\dim(U \cap V) \\ &= d(U,V). \quad \square \end{aligned}$$

2.4.3 Observación. En virtud del lema anterior, tenemos que

$$D(C) = D(C^\perp).$$

Si C es un código de dimensión constante y C es un $[N, \ell(C), |C|, D]_q$-código, entonces $C^\perp$ es un $[N, N-\ell(C), |C|, D]_q$-código con dimensión constate.

En la siguiente sección se se dan a conocer otras cotas para códigos, las cuales son de nuestro interés.

2.5. Algunas cotas para los códigos de red lineales.

Para hablar de construcción de códigos de dimension constante, debemos comenzar entonces esta sección introduciendo algunas notaciones que podrían ser relevantes para paquetes en $\mathbb{P}(W, \ell)$, donde W es el espacio ambiente de dimensión N sobre $K = \mathbb{F}_q$.

El q-ésimo ***coeficiente Gaussiano*** es definido por $\ell, n \in \mathbb{Z}^+$ con $\ell \leq n$, de la siguiente manera:

$$\begin{aligned} \begin{bmatrix} n \\ \ell \end{bmatrix}_q &:= \frac{(q^n-1)(q^{n-1}-1)\cdots(q^{n-\ell+1}-1)}{(q^\ell-1)(q^{\ell-1}-1)\cdots(q-1)} \\ &= \prod_{i=0}^{\ell-1} \frac{q^{n-i}-1}{q^{\ell-i}-1}, \end{aligned}$$

donde el producto vació es obtenido cuando $\ell = 0$ y es interpretado como 1.

El coeficiente Gaussiano $\begin{bmatrix} n \\ \ell \end{bmatrix}_q$ nos permite calcular el número de subespacios distintos ℓ-dimensionales de un n-dimensional espacio vectorial sobre $\mathbb{F}_q$.

El comportamiento del coeficiente Gaussiano para $q > 1$ esta dado por el siguiente lema.

2.5.1 Lema. Para todo $0 < \ell < n$, se cumple que:

$$1 < q^{-\ell(n-\ell)} \begin{bmatrix} n \\ \ell \end{bmatrix}_q < 4$$

Demostración. La cantidad $q^{\ell(n-\ell)}$ es el número de subespacios ℓ-dimensionales determinados por las filas de matrices de la forma $\begin{pmatrix} I_\ell \mid X \end{pmatrix}$, donde $I \in \mathrm{Mat}(\ell \times \ell, \mathbb{F}_q)$ y $X \in \mathrm{Mat}(\ell \times (n-\ell), \mathbb{F}_q)$. Como $\ell > 0$ este número es más pequeño que el número de subespacios ℓ-dimensionales de $\mathbb{F}_q^n$, es decir,

$$q^{\ell(n-\ell)} < |\mathbb{P}(W, \ell)| = \begin{bmatrix} n \\ \ell \end{bmatrix}_q,$$

entonces $1 < q^{\ell(n-\ell)} \begin{bmatrix} n \\ \ell \end{bmatrix}_q$. Para el lado derecho de la desigualdad observe que $\begin{bmatrix} n \\ \ell \end{bmatrix}_q$ puede ser escrito como se indica a continuación:

$$\begin{aligned}
\begin{bmatrix} n \\ \ell \end{bmatrix}_q &= \frac{(q^n-1)(q^{n-1}-1)\cdots(q^{n-\ell+1}-1)}{(q^\ell-1)(q^{\ell-1}-1)\cdots(q-1)} \\
&= \frac{q^n(1-q^{-n})q^{n-1}(1-q^{-n+1})\cdots q^{n-\ell+1}(1-q^{-n+\ell-1})}{q^\ell(1-q^{-\ell})q^{\ell-1}(1-q^{-\ell+1})\cdots q(1-q^{-1})} \\
&= q^{\ell(n-\ell)} \frac{(1-q^{-n})(1-q^{-n+1})\cdots(1-q^{-n+\ell-1})}{(1-q^{-\ell})(1-q^{-\ell+1})\cdots(1-q^{-1})} \\
&= q^{\ell(n-\ell)} \frac{1}{(1-q^{-\ell})\cdots(1-q^{-1})} \\
&< q^{\ell(n-\ell)} \prod_{j=1}^{\infty} \frac{1}{(1-q^{-j})}.
\end{aligned}$$

Consideremos el caso $f(\frac{1}{q})$ para $q \geq 2$, en la función

$$f(x) = \prod_{j=1}^{\infty} \frac{1}{(1-x^j)}$$

la cual es la función generadora de división de enteros [5], entonces

$$\begin{aligned} f(\tfrac{1}{q}) &= \prod_{j=1}^{\infty} \frac{1}{(1-q^{-j})} \\ &= \prod_{j=1}^{\infty} \frac{q^j}{(q^j-1)} \\ &< \prod_{j=1}^{\infty} \frac{2^j}{(2^j-1)} \\ &= \prod_{j=1}^{\infty} \frac{1}{1-(2^{-j})} \\ &= \frac{1}{Q_0} \\ &< 4, \end{aligned}$$

donde $Q_0 \approx 0,288788095$, ver [8]. □

Como se desea establecer algunas otras cotas necesitamos tener una noción mas concreta de lo que es una esfera, la cual se define formalmente como se muestra a continuación.

2.5.2 Definición. La esfera $S(V, \ell, t)$ de radio t centrada en un espacio V en $\mathbb{P}(W, \ell)$ es definida como el conjunto de todos los subespacios U que satisfacen la siguiente desigualdad $d(U, V) \leq 2t$

$$S(V, \ell, t) = \{U \in \mathbb{P}(W, \ell) : d(U, V) \leq 2t\}.$$

Note que nos referimos a la definición del radio en términos de la distancia del grafo en el grafo Grassmaniano. El radio no puede tomar un valor entero negativo.

2.5.3 Teorema. El número de espacios en $S(V, \ell, t)$ es independiente de V e igual a

$$|S(V, \ell, t)| = \sum_{i=0}^{t} q^{i^2} \begin{bmatrix} \ell \\ i \end{bmatrix} \begin{bmatrix} N-\ell \\ i \end{bmatrix}$$

para $t \leq \ell$.

Demostración. La afirmación de que $S(V, \ell, t)$ es independiente de V se deduce del hecho de que $P(W, \ell)$ constituye un gráfico distancia regular ver [9], es decir, grafo donde cada vértice tiene el mismo grado o número de

aristas adyacentes.

Damos una expresión para el número de espacios U que se intersecan en V en un subespacio $(\ell - i)$-dimensional. Podemos elegir el subespacio $(\ell - i)$-dimensional de intersección en $\begin{bmatrix} \ell \\ \ell - i \end{bmatrix} = \begin{bmatrix} \ell \\ i \end{bmatrix}$ maneras. Una vez hecho esto podemos completar el subespacio en

$$\frac{(q^N - q^\ell)(q^N - q^{\ell+1}) \cdots (q^N - q^{\ell+i-1})}{(q^\ell - q^{\ell-i})(q^\ell - q^{\ell-i+1}) \cdots (q^\ell - q^{\ell-1})} = q^{i^2} \begin{bmatrix} N - \ell \\ i \end{bmatrix}$$

maneras. Así, la cardinalidad de un recubrimiento de espacios a una distancia $2i$ alrededor de V es igual a $q^{i^2} \begin{bmatrix} N - \ell \\ i \end{bmatrix} \begin{bmatrix} \ell \\ i \end{bmatrix}$. Sumando la cardinalidad del recubrimiento se da la afirmación del teorema. Tenga en cuenta que

$$|S(V, \ell, t)| = |S(V, N - \ell, t)|,$$

como se esperaba a partir de la definición anterior. □

2.5.4 Teorema. Sea $C \subseteq \mathbb{P}(W, \ell)$ tal que $D(C) \geq 2t$, y sea $s = \lfloor \frac{t-1}{2} \rfloor$. El tamaño de C debe satisfacer que.

$$\begin{aligned} |C| &\leq \frac{|\mathbb{P}(W, \ell)|}{|S(V, \ell, s)|} \\ &= \frac{\begin{bmatrix} N \\ \ell \end{bmatrix}}{|S(V, \ell, s)|} \\ &< \frac{\begin{bmatrix} N \\ \ell \end{bmatrix}}{q^{s^2} \begin{bmatrix} N - \ell \\ s \end{bmatrix} \begin{bmatrix} \ell \\ s \end{bmatrix}} \\ &< 4\, q^{(\ell - s)(N - s - \ell)}. \end{aligned}$$

Por el contrario, existe un código C' con distancia $D(C) \geq 2t$ tal que $|C'|$

es reducido por

$$\begin{aligned} |C'| &\geq \frac{|\mathbb{P}(W,\ell)|}{|S(V,\ell,t-1)|} \\ &= \frac{\begin{bmatrix} N \\ \ell \end{bmatrix}}{|S(V,\ell,t-1)|} \\ &> \frac{\begin{bmatrix} N \\ \ell \end{bmatrix}}{(t-1)q^{(t-1)^2}\begin{bmatrix} N-\ell \\ t-1 \end{bmatrix}\begin{bmatrix} \ell \\ t \end{bmatrix}} \\ &> \frac{1}{16}\, q^{(\ell-t+1)(N-t+1)}. \end{aligned}$$

Demostración. Dado que

$$|S(V,\ell,t)| = \sum_{i=0}^{t} q^{i^2} \begin{bmatrix} m \\ \ell \end{bmatrix}\begin{bmatrix} m-\ell \\ i \end{bmatrix}$$

para $t \leq \ell$ en $\mathbb{P}(W,\ell)$, la cota inferior y superior es justamente la familia de paquetes de recubrimientos para códigos en gráficos de distancia regular.□

Como en el caso de la cota del esquema de Hamming, la cota superior no es tan buena, ya que es fácil ver que δ no puede ser mayor que 1. Necesitamos derivar una cota tipo Singleton para paquetes en el grafo de Grassmann.

Cota de Singleton.

Empecemos definiendo una operación de perforación adecuada en los códigos. Supongamos que C es una colección de espacios en $\mathbb{P}(W,\ell)$, donde W es de dimensión N. Sea W' cualquier subespacio de W de dimensión $N-1$. Un código perforador C' se obtiene a partir de C mediante la sustitución de cada espacio $V \in C$ por $V' = H_{\ell-1}(V \cap W')$ donde $H_{\ell-1}$ denota las borraduras definida anteriormente. En otras palabras, V se sustituye por $V \cap W'$ si $V \cap W'$ tiene dimensión $\ell-1$, de lo contrario V se sustituye por un subespacio $(\ell-1)$-dimensional de V. Aunque esta operación de perforación no resulta en un único código, un código perforado lo denotamos como $C|W'$. Dando lugar al siguiente teorema.

2.5.5 Lema. Si $C \subseteq \mathbb{P}(W,\ell)$ con parámetros $[N,\ell,|C|,D]_q$ con $D > 2$ y W' es un subespacio $(N-1)-$dimensional de W, entonces $C' = C|_{W'}$ tiene parámetros $[N-1,\ell-1,|C|,D']_q$, donde $D' \geq D-2$.

Demostración. Claramente, todos los espacios de C' son subespacios de dimensión $\ell - 1$ en W'. Entonces tenemos que probar que $|C| = |C'|$ y que $D' \geq D - 2$. Sea U y V dos codewords de C, y supongamos que $U' = \mathcal{H}_{\ell-1}(U \cap W')$ y $V' = \mathcal{H}_{\ell-1}(V \cap W')$ son los códigos correspondientes en C'. Por lo tanto $U \subseteq U'$ y $V' \subseteq V$, entonces tenemos que $U' \cap V' \subseteq U \cap V$, luego

$$2\dim(U' \cap V') \leq 2\dim(U \cap V) \leq 2\ell - D$$

por lo tanto

$$D \leq d(U, V) = 2\ell - 2(U \cap V),$$

entonces en C' se tiene que

$$\begin{aligned} d(U', V') &= \dim(U') + \dim(V') - 2\dim(U' \cap V') \\ &= 2(\ell - 1) - 2\dim(U' \cap V') \\ &\geq 2\ell - 2 - (2\ell - D) \\ &= D - 2. \end{aligned}$$

Dado que $D > 2$, $d(U' \cap V') > 0$, como U' y V' son distintos, lo cual muestra que C' tiene tantos codewords como C. □

2.5.6 Teorema. [Singleton para códigos de dimension constante.] Sea $C \subseteq \mathbb{P}(W, \ell)$ con parámetros $[N, \ell, |C|, D]_q$ entonces

$$|C| \leq \begin{bmatrix} N - (D-2)/2 \\ \text{máx}\{\ell, N - \ell\} \end{bmatrix}_q$$

Demostración. Por 2.5.5 el código C puede ser perforado exactamente $\dfrac{D-2}{2}$ veces. Dado que $D - 2\dfrac{D-2}{2} = 2$. Escogemos un código $C \subseteq \mathbb{P}(W', \ell')$, donde la $\dim W' = N - \dfrac{D-2}{2}$ y $\ell' = \ell - \dfrac{D-2}{2}$.

$$\begin{aligned} |C| = |C| &\leq |\mathbb{P}(W')| \\ &= \begin{bmatrix} N - (D-2)/2 \\ \ell - (D-2)/2 \end{bmatrix}_q \\ &= \begin{bmatrix} N - (D-2)/2 \\ N - \ell \end{bmatrix}_q \\ &= a. \end{aligned}$$

Reemplazamos C por $C^{\perp}$

$$\begin{aligned} |C| &= |C^{\perp}| \\ &\leq \begin{bmatrix} N-(D-2)/2 \\ \ell \end{bmatrix}_q \\ &= b. \end{aligned}$$

Note que $a < b$ si y solo si $\ell < N - \ell$.□

2.5.7 Teorema. Sea $C \subseteq \mathbb{P}(W, \ell)$ un código de dimensión constante con $\dim W = N$ y $d(C) = 2d > 2$, entonces

$$|C| \leq 4\ q^{\text{máx}\{\ell, N-\ell\}(\text{mín}\{\ell, N-\ell\}-d+1)}. \tag{2.2}$$

Demostración.
Por el teorema 2.5.6 se tiene que

$$|C| \leq \begin{bmatrix} N-(D-2)/2 \\ \text{máx}\{\ell, N-\ell\} \end{bmatrix}_q$$

y de acuerdo con el lema 2.5.1 se obtiene lo siguiente

$$\begin{aligned} \left[\begin{array}{c} N-d+1 \\ \text{máx}\{\ell, N-\ell\} \end{array} \right]_q &< 4\ q^{\text{máx}\{\ell, N-\ell\}(N-d+1-\text{máx}\{\ell, N-\ell\})} \\ &= 4\ q^{\text{máx}\{\ell, N-\ell\}(\text{mín}\{\ell, N-\ell\}-d+1)}. \quad \square \end{aligned}$$

Sea $A_q[N, 2d, \ell]$ el número máximo de codewords en un código de dimensión constante con distancia mínima $2d$, donde

$$A_q[N, 2d, \ell] \leq \left[\begin{array}{c} N-d+1 \\ \text{máx}\{\ell, N-\ell\} \end{array} \right]_q < 4\ q^{\text{máx}\{\ell, N-\ell\}(\text{mín}\{\ell, N-\ell\}-d+1)}$$

Capítulo 3

Códigos de red y la métrica del rango.

3.1. La métrica del rango

Uno de los objetivos de este capitulo es el estudio de la métrica del rango, la cual fue introducida por E. Gabidulin[1] en 1985, iniciaremos recordando en breve algunos preliminares necesarios para desarrollar dicha teoría.

I_n, denotará la matriz identidad de orden n, entonces la notación I_i hará referencia a la i-ésima columna de I. En general si $U \subseteq \{1, \ldots, N\}$, entonces $I_U = [I_i, i \in U]$ denotará la submatriz de I que consiste en las columnas indexadas por U.

El espacio vectorial generado por un conjunto de vectores $v_1, \ldots, v_k$ se denota por $\langle v_1, \ldots, v_k \rangle$.

Sea $X \in \mathrm{Mat}(n \times m, \mathbb{F}_q)$, $\langle X \rangle$ denota el espacio vectorial generado por las filas de X, Rang X denota el rango de la matriz X y el peso de X, es decir el número de filas no nulas de X se denota por $\omega t(X)$.

[1]Ernst Mukhamedovich Gabidulin nació en Kok-Jangak, Kirguistán, la URSS, el 4 de junio de 1937. Profesor en el departamento de ingeniería de radio MIPT y jefe de este departamento. Sus intereses de investigación incluyen la teoría de códigos, técnicas de codificación, comunicaciones, redes de comunicación, la teoría de Shannon, códigos fuente, la criptografía y codificación de la fuente entre otras.

3.1.1 Definición. Si $C \subseteq \text{Mat}(n \times m, \mathbb{F}_q)$ y $C \neq \emptyset$, entonces C es llamado ***código matricial.***

Una medida de distancia natural y útil entre los elementos de $\text{Mat}(n \times m, \mathbb{F}_q)$ se presenta en la siguiente definición.

3.1.2 Definición. Sean $X, Y \in \text{Mat}(n \times m, \mathbb{F}_q)$, se define ***la distancia del rango*** entre X y Y así:

$$d_R(X, Y) := \text{Rang } (Y - X).$$

3.1.3 Ejemplo. Sean $X, Y \in \text{Mat}(4, \mathbb{F}_2)$

$$X = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \; Y = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix},$$

el rango para la matriz X claramente es 2, note también que

$$d_R(X, Y) = \text{Rang } (Y - X) = 3,$$

donde

$$Y - X = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

3.1.4 Lema. La distancia del rango d_R es una métrica. Es decir, para todo $X, Y, Z \in \text{Mat}(n \times m, \mathbb{F}_q)$ se verifican

(a) $d_R(X, Y) \geq 0$ y $d_R(X, Y) = 0$ si y solo si $X = Y$.

(b) $d_R(X, Y) = d_R(Y, X)$.

(c) $d_R(X, Y) \leq d_R(X, Z) + d_R(Z, Y)$.

Demostración. Sean $X, Y, Z \in \text{Mat}(n \times m, \mathbb{F}_q)$

(a) Dado que

$$\text{Rang } (Y - X) \in \mathbb{N}_0,$$

se tiene que $\text{Rang } (Y - X) \geq 0$. Por otro lado si,

$$\begin{aligned} Y = X &\Rightarrow Y - X = 0 \\ &\Rightarrow \text{Rang } (Y - X) = 0. \\ &\Rightarrow d_R(X, Y) = 0. \end{aligned}$$

Recíprocamente si Rang $(Y - X) = 0$, entonces $Y - X = 0$ por lo tanto $X = Y$.

(b)

$$\begin{aligned} d_R(X,Y) &= \text{Rang } (Y-X) \\ &= \text{Rang } (X-Y) \\ &= d_R(Y,X). \end{aligned}$$

(c)

$$\begin{aligned} d_R(X,Y) &= \text{Rang } (Y-X) \\ &= \text{Rang } (Y-Z+Z+X) \\ &\leq \text{Rang } (Y-Z) + \text{Rang } (Z-X) \\ &= d_R(X,Z) + d_R(Z,Y). \quad \square \end{aligned}$$

En el contexto de la métrica del rango, un código matricial es denominado ***código con la métrica del rango***, también conocidos como ***MR-códigos***.

3.1.5 Definición. Sea $C \subseteq \text{Mat}(n \times m, \mathbb{F}_q)$, ***la distancia mínima*** de un código C con métrica del rango se define de la siguiente manera:

$$\bar{d} = D_R(C) := \text{mín}\{d_R(X,Y) \mid X, Y \in C, X \neq Y\}.$$

3.1.6 Observación. Sea $C \subseteq \text{Mat}(n \times m, \mathbb{F}_q)$. Asociado a C está su correspondiente código transpuesto $C^T \subseteq \mathbb{F}_q^{m \times n}$, cuyos codewords se obtienen transponiendo los codewords de C. Es decir,

$$C^T = \{X^T \mid X \in C\}.$$

Note que

$$|C^T| = |C| \text{ y } D_R(C^T) = D_R(C).$$

Existe una amplia teoría de códigos con la métrica del rango que es análoga a la teoría clásica de códigos en la métrica de Hamming. En particular se menciona la existencia de una cota de Singleton la cual establece lo siguiente.

3.1.7 Teorema. (Cota de Singleton) Sea $C \subseteq \text{Mat}(n \times m, \mathbb{F}_q)$ un MR-código con distancia mínima $\bar{d}$. Entonces

$$\begin{aligned} \log_q |C| &\leq \text{mín}\{n(m - \bar{d} + 1), m(n - \bar{d} + 1)\} \\ &= \text{máx}\{n, m\}(\text{mín}\{n, m\} - \bar{d} + 1). \end{aligned}$$

Demostración. Consideremos la siguiente proyección

$$\pi : C \longrightarrow \mathrm{Mat}(n \times (m - \bar{d} + 1), \mathbb{F}_q)$$

definida de tal forma que a cada $X \in \mathrm{Mat}(n \times m, \mathbb{F}_q)$, le asigna $\pi(X) \in \mathrm{Mat}(n \times (m - \bar{d} + 1), \mathbb{F}_q)$, que consiste en borrarle a la matriz X las ultimas $\bar{d} - 1$ columnas.

Supongamos que $\pi(X) = \pi(Y)$, entonces Rang $(X - Y) \leq \bar{d} - 1$ y por la definición de d_R se tiene que

$$d_R(X, Y) = \mathrm{Rang}\ (X - Y) = 0.$$

Por lo tanto $X = Y$ con lo que concluimos que π es inyectiva. En consecuencia

$$|C| \leq |\mathrm{Mat}(n \times (m - \bar{d} + 1), \mathbb{F}_q)| = q^{n(m-\bar{d}+1)}.$$

Es decir que

$$\log_q |C| \leq n(m - \bar{d} + 1).$$

Análogamente si omitimos las ultimas $\bar{d} - 1$ filas, se obtiene que:

$$\log_q |C| \leq m(n - \bar{d} + 1).$$

Por lo tanto

$$\log_q |C| \leq \mathrm{mín}\{n(m - \bar{d} + 1), m(n - \bar{d} + 1)\}.$$

Note que si $m \geq n$, entonces $n(m-\bar{d}+1) \geq m(n-\bar{d}+1)$, y si $m \leq n$, entonces $n(m - d + 1) \leq m(n - d + 1)$, por lo tanto

$$\mathrm{mín}\{n(m - \bar{d} + 1), m(n - \bar{d} + 1)\} = \mathrm{máx}\{n, m\}(\mathrm{mín}\{n, m\} - \bar{d} + 1). \quad \square$$

Los códigos que alcanzan esta cota se denominan MRD-códigos.

3.1.8 Definición. (MRD-código) Sea $C \subseteq \mathrm{Mat}(n \times m, \mathbb{F}_q)$ con distancia mínima $\bar{d}$ y C un MR-código, decimos que C es un código de distancia de rango máximo o MRD-código si:

(a) $\bar{d} = \mathrm{mín}\{n, m\} - k + 1$

(b) $|C| = q^{\mathrm{máx}\{n,m\} \cdot k}$

Ejemplos de estos códigos fueron presentados por Gabidulin en [11].

3.2. Control de error en códigos de red.

3.2.1 Definición. Sea f es una función cualquiera, entonces el ***argmin*** de f en x se define de la siguiente manera:

$$\underset{x\in \text{ Dom}f}{\text{argmin}} f(x) := \{x \mid f(y) \geq f(x), \forall y \in \text{Dom}f\}$$

Un decodificador de mínima distancia para un código $C \subseteq \text{Mat}(n \times m, \mathbb{F}_q)$ con la métrica del rango toma un $S \in \text{Mat}(n \times m, \mathbb{F}_q)$ y retorna el codeword $X \in C$ que está mas cerca de S con respecto a la métrica del rango; esto es

$$X = \underset{Y\in\ C}{\text{argmin}}\ \text{Rang}\ (S - Y). \tag{3.1}$$

Note que si $d_R(S, R) < \frac{1}{2}\ D_R(C)$, entonces un decodificador de distancia mínima garantiza que retornará siempre $Y = X$. El problema (3.1) se conoce como el problema tradicional de decodificación del rango.

En capítulos previos se abordaron temas relacionados con la decodificación del síndrome, el canal operador, la decodificación y la distancia mínima; en esta sección se tiene como objetivo hacer las mismas consideraciones pero desde un enfoque matricial. Para hacer eso, simplemente reemplazamos a W un espacio vectorial N-dimensional sobre $\mathbb{F}_q$ por $\text{Mat}(n \times m, \mathbb{F}_q)$ en donde $C \subseteq \text{Mat}(n \times m, \mathbb{F}_q)$ un código matricial y la transmisión de información es análoga a la descrita en el capítulo 2, en donde el nodo fuente elige una matriz $X \in C$ para transmitirla sobre el canal, es claro que la transmisión de X es equivalente a la transmisión de $\langle X\rangle$ usar a C el código matricial es equivalente a usar el código de subespacios

$$\langle C\rangle := \{\langle X\rangle \mid X \in C\}.$$

Tras la recepción de Y, el nodo intenta inferir la matriz transmitida usando la regla de decodificación de distancia mínima

$$\hat{X} = \underset{X\in\ C}{\text{argmin}}\ d(\langle X\rangle, \langle Y\rangle). \tag{3.2}$$

$$\hat{x} = \underset{x\in C}{\text{argmin}}\ \text{Rang}\ (y - \hat{A}x). \tag{3.3}$$

Recordemos que estaremos frente a una decodificación exitosa siempre que

$$d(\langle X\rangle, \langle Y\rangle) < \frac{D(C)}{2}\ .$$

El modelo del canal

Esta sección tiene como objetivo el análisis del modelo de canal para códigos matriciales, el cual funciona de forma similar al canal operador explicado en capítulo dos para códigos de dimensión constante, recordemos que para este caso los paquetes de entrada eran vectores filas así como los paquetes recibidos, en el caso de códigos matriciales los paquetes enviados y recibidos son matrices, pero el envió de estos se hace enviando cada una de las filas de la matriz, es decir, tomando cada una de estas como un vector fila lo cual genera el modelo de canal operador explicado en capitulo dos, sin embargo haremos nuevamente algunas especificaciones de nuestro nuevo modelo de canal para mayor claridad.

Sea $X \in \mathrm{Mat}(n \times M, \mathbb{F}_q)$ una matriz cuyas filas son los paquetes transmitidos $X_1, \ldots, X_n$ y $Y \in \mathrm{Mat}(N \times M, \mathbb{F}_q)$ una matriz cuyas filas son los paquetes recibidos $Y_1, \ldots, Y_N$. Dado que todas las operaciones de paquetes son lineales sobre $\mathbb{F}_q$, entonces, independientemente de la topología de red, los paquetes transmitidos X y los paquetes recibidos Y se relacionan de la siguiente manera

$$Y = AX, \tag{3.4}$$

donde $A \in \mathrm{Mat}(N \times n, \mathbb{F}_q)$ que corresponde a la transformación lineal global aplicada por la red.

Si optamos por considerar el envió de paquetes erróneos el modelo se extiende, consideramos entonces que los errores pueden ocurrir en cualquiera de los arcos de la red. Supongamos que los arcos de la red se indexan desde 1 a ℓ, donde Z_i denota el paquete de errores ocurridos en $i \in \{1, \ldots, \ell\}$. La aplicación de un error se modela como sigue. Se supone que, para cada enlace i, el nodo de transmisión en ese primer enlace crea un paquete prescrito $P_{\mathtt{in},i} \in \mathbb{F}_q^{1\times M}$ siguiendo el procedimiento descrito anteriormente. Entonces, un paquete de error $Z_i \in \mathbb{F}_q^{1\times M}$ es añadido a $P_{\mathtt{in},i}$, con el fin de producir el paquete de salida en este enlace, es decir,

$$P_{out,i} = P_{\mathtt{in},i} + Z_i.$$

Tenga en cuenta que cualquier paquete arbitrario $P_{out,i}$ puede estar formado simplemente por la elección de

$$Z_i = P_{out,i} - P_{\mathtt{in},i}.$$

Por lo tanto $Z \in \mathrm{Mat}(\ell \times M, \mathbb{F}_q)$ una matriz cuyas filas son los paquetes de error $Z_1, \ldots, Z_\ell$ y por linealidad de la red, podemos escribir

$$Y = AX + BZ, \tag{3.5}$$

donde $B \in \text{Mat}(N \times \ell, \mathbb{F}_q)$ es la matriz correspondiente a la transformación lineal global aplicada a $Z_1, \ldots, Z_\ell$ sobre la ruta de destino. Recordemos que si $Z_i = 0$ esto significa que no ocurrió ningún error en el enlace i lo que da lugar a la siguiente definición.

3.2.2 Definición. Sea $Z \in \text{Mat}(\ell \times M, \mathbb{F}_q)$ una matriz cuyas filas son los paquetes de error $Z_1, \ldots, Z_\ell$, el peso de Z denotado por $\omega t(Z)$ se define de la siguiente manera

$$\omega t(Z) := |\{i \mid Z_i \neq 0, i = 1, \ldots, \ell\}|,$$

es decir, el peso de la matriz Z es el número de filas de Z distintas cero, que son el número de paquetes erróneos inyectados en la red.

Lo que se desea ahora es establecer las condiciones para que un código sea eficiente y además que estos parámetros sean lo suficientemente generales para que no sea necesario tomar completamente la topología de la red en cuenta. Para ello hacemos las siguientes suposiciones considerando a $Y = AX + BZ$, donde X, Y, Z son matrices cuyas filas representan los paquetes transmitidos, recibidos y corruptos y A y B son las correspondientes matrices de transferencias inducidas por la codificación de red lineal. Los supuestos son:

(a) La deficiencia del rango de columna de la matriz de transferencia A nunca es mayor que ρ, es decir, el rango de Rang $A \geq n - \rho$, donde ρ define al igual que en capitulo 2 el máximo número de supresiones ocurridas en el canal.

(b) Los nodos transmisores pueden enviar simultáneamente a lo mas t paquetes corruptos, es decir, con $\omega t(Z) \leq t$.

El siguiente resultado caracteriza las garantías de rendimiento de un código de subespacio bajo nuestras suposiciones.

3.2.3 Teorema. Supongamos que Rang $A \geq n - \rho$ y $\omega t(Z) \leq t$. Entonces, la decodificación dada por (3.2) es exitosa siempre que $2t + \rho < \frac{D(C)}{2}$.

Demostración. Del corolario A.2.2 tenemos que

$$d(\langle AX \rangle, \langle Y \rangle) \leq 2\text{Rang } BZ \leq 2\text{Rang } Z \leq \omega t(Z) \leq 2t.$$

Si usamos (A.4) encontramos que

$$d(\langle X \rangle, \langle AX \rangle) = \text{Rang } X - \text{Rang } AX \leq n - \text{Rang } A \leq \rho.$$

Por la desigualdad del triángulo, tenemos

$$\begin{aligned} d(\langle X\rangle, \langle Y\rangle) &\leq d(\langle X\rangle, \langle AX\rangle) + (\langle AX\rangle \cap \langle Y\rangle) \\ &\leq \rho + 2t \\ &< \frac{D(C)}{2} \end{aligned}$$

lo que garantiza que la decodificación sea exitosa. □

El teorema 3.2.3 es análogo al teorema 2.2.7 el cual establece que la decodificación por mínima distancia tiene éxito si $2(\mu + \delta) < d(C)$, donde μ y δ son, respectivamente, el número de supresiones (proceso en el cual una columna L.I es reemplazada por una L.D) e inserciones (proceso inverso a la supresión) de dimensiones que ocurren en el canal operador descrito en capitulo 2.

Dado, que un paquete corrupto y enviado en una red por el mínimo corte ***(min-cut)*** puede intuitivamente sustituir efectivamente una dimensión del subespacio transmitido, vemos que t paquetes corruptos pueden causar t borraduras y t inserciones de dimension. Combinando con ρ posibles supresiones adicionales causadas por una deficiencia de rango de A, se tiene que $\delta = t$ y $\mu = t + \delta$. Por lo tanto,

$$\delta + \mu < \frac{d(C)}{2} .$$

Entonces

$$2t + \rho < \frac{d(C)}{2} .$$

En otras palabras, bajo la condición que los paquetes corruptos pueden ser inyectados en cualquiera de los enlaces de la red (esto debe asumirse para no tener en cuenta la topología de la red), la garantía de rendimiento de un decodificador de distancia mínima es esencialmente dada por el teorema 3.2.3.

Vale la pena mencionar que de acuerdo con los últimos resultados de [12], la decodificación por mínima distancia de subespacios puede no ser la regla de decodificación mas óptima cuando los subespacios en C tienen diferentes dimensiones. Sin embargo, nosotros nos concentraremos en códigos de dimensión constante y por lo tanto utilizaremos la regla de decodificación de mínima distancia (3.2). Iniciemos con la decodificación de un código matricial.

3.3. Decodificación basada en códigos matriciales

En esta sección demostramos cómo un código de dimensión constante puede ser construido a partir de un código con la métrica del rango. En particular, esta construcción nos permitirá obtener códigos de subespacios que posean algoritmos eficaces para codificación y decodificación.

Construcción por levantamiento.

A partir de ahora, asumiremos que $M \geq n$ y que $m = M - n$, donde $m > 0$.

3.3.1 Definición. El levantamiento de una matriz $X \in \mathrm{Mat}(n \times m, \mathbb{F}_q)$ se define como la matriz

$$I(X) = \langle (I_n \mid X) \rangle$$

que se obtiene a través del generado de la matriz que resulta de anteponer a la matriz X la matriz identidad I_n. De manera similar se define el ***levantamiento de un código matricial*** para $C \subseteq \mathrm{Mat}(n \times m, \mathbb{F}_q)$ como

$$I(C) := \{\langle (I_n \mid X) \rangle \mid X \in C\}.$$

Es decir se obtiene mediante el levantamiento de cada codeword x de C.

La definición 3.3.1 proporciona una aplicación inyectiva entre los códigos con métrica del rango y códigos de subespacios.

Dado un código matricial $C \subseteq \mathrm{Mat}(n \times m, \mathbb{F}_q)$ es siempre posible construir un código $I(C)$. El correspondiente código de subespacios $I(C)$ será siempre un código de dimensión constante, en el cual cada codeword tiene dimensión n. Note que $I(\mathrm{Mat}(n \times m, \mathbb{F}_q))$ pertenece a $\mathbb{P}(\mathbb{F}_q^{n+m})$ para todo $m > 0$, recuerde que $\mathbb{P}(\mathbb{F}_q^{n+m})$ es el conjunto de todos los subespacios de $\mathbb{F}_q^{n+m}$.

Aunque la construcción por levantamiento es una opción particular para la construcción de códigos de subespacios, también puede ser considerado como una generalización del método estándar para la codificación de red aleatoria [14], [13]. En este último, todas las matrices transmitidas tienen la forma

$$(I_n \mid X)$$

donde la matriz $X \in \mathrm{Mat}(n \times m, \mathbb{F}_q)$ corresponde a los datos que serán comunicados. En nuestro enfoque, cada matriz de transmisión es también de la forma

$$(I_n \mid X),$$

pero la matriz de carga $X \in C$ es restringida a ser un codeword del código con métrica del rango, en lugar de datos sin codificar.

Nuestras razones para la elección de C como un código matricial con la métrica del rango se hace evidente a partir del siguiente lema.

3.3.2 Lema. Sea $C \subseteq \mathrm{Mat}(n \times m, \mathbb{F}_q)$ y $X, X' \in C$. Entonces

(a) $d(I(X), I(X')) = 2d_R(X, X')$

(b) $d(I(C)) = 2D_R(C) = 2\bar{d}$.

Demostración.

(a) Puesto que $\dim I(X) = \dim I(X') = n$, tenemos que:

$$\begin{aligned} d(I(X), I(X')) &= \dim(I(X) + I(X')) - \dim(I(X) \cap I(X')) \\ &= 2\ \dim(I(X) + I(X')) - \dim I(X) - \dim I(X') \\ &= 2\ \dim(I(X) + I(X'))) - 2n \\ &= 2\ \mathrm{Rang} \begin{pmatrix} I & X \\ I & X' \end{pmatrix} - 2n \\ &= 2\ \mathrm{Rang} \begin{pmatrix} I & X \\ 0 & X' - X \end{pmatrix} - 2n \\ &= 2\ (n + \mathrm{Rang}\ (X - X')) - 2n \\ &= 2\ \mathrm{Rang}\ (X - X') \\ &= 2\ d_R(X, X'). \end{aligned}$$

(b) Se sigue de (a). □

La anterior proposición muestra que un código de subespacio construido por elevación hereda las propiedades de distancia de su código matricial con la métrica del rango subyacente. La pregunta de si estos códigos levantados son "buenos" en comparación con la clase de todos los códigos de dimensión constante se aborda en el siguiente lema, pero antes demos lugar a la siguiente definición de suboptimalidad para códigos de dimensión constante.

3.3.3 Definición. Sea $C \subseteq \mathbb{P}(W, \ell)$ un código de dimensión constante con $\dim W = N$ y $d(C) = 2d$ definimos la ***suboptimalidad*** de C de la siguiente manera

$$\alpha(C) := \frac{\log_q A_q[N, 2d, \ell] - \log_q |C|}{\log_q A_q[N, 2d, \ell]}.$$

3.3.4 Lema. Sea $C \subseteq \mathrm{Mat}(n \times m, \mathbb{F}_q)$ un MRD-código. Para cualquier código $C' \subseteq \mathbb{P}(\mathbb{F}_q^{n+m}, n)$, con $D_R(C') \geq d(I(C))$ se verifica que

$$|C'| < 4|I(C)| = 4|C|.$$

Además, para los parámetros del código, la suboptimalidad de $I(C)$ en $\mathbb{P}(\mathbb{F}_q^{n+m}, n)$ satisface lo siguiente

$$\alpha(I(C)) < \frac{4}{(n+m)\log_2 q}$$

Demostración. Haciendo uso del teorema 2.5.7 y del hecho de que C es un MRD-código, es decir que C alcanza la cota de singleton para códigos matriciales tenemos que

$$\begin{aligned} \log_q \frac{|C'|}{4} &< \max\{n,m\}(\min\{n,m\} - \tfrac{d(I(C))}{2} + 1) \\ &< \max\{n,m\}(\min\{n,m\} - \bar{d} + 1) \\ &= \log_q |C|, \end{aligned}$$

por lo tanto

$$|C'| < 4|C|,$$

donde $\bar{d} = D_R(C) = \frac{1}{2}\ d(\langle I(C)\rangle)$, también se obtiene que

$$\begin{aligned} \alpha(\langle I(C)\rangle) &= \frac{\log_q A_q[M, 2\bar{d}, n] - \log_q |\langle I(C)\rangle|}{\log_q A_q[M, 2\bar{d}, n]} \\ &< \frac{\log_q 4 + \log_q |C| - \log_q |C|}{\log_q 4|C|} \\ &< \frac{\log_q 4}{\log_q 4 + \log_q |C|} \\ &< \frac{\log_q 4}{\log_q |C|} \\ &< \frac{\log_q 4}{\max\{n,m\}(\min\{n,m\} - d + 1)} \\ &\leq \frac{\log_q 4}{\max\{n,m\}} \end{aligned}$$

$$\begin{aligned} &\leq \frac{\log_q 4}{(n+m)/2} \\ &= \frac{\log_2 4^2}{(n+m)\log_2 q} \\ &= \frac{4}{(n+m)\log_2 q}\ .\square \end{aligned}$$

Este lema muestra que, no hay perdida de optimalidad en restringir la atención a códigos con métrica del rango obtenidos por levantamiento.

Decodificación

Ahora nos concentraremos en el problema de decodificación (3.2), para el caso específico de códigos de subespacios obtenidos a partir de levantamientos de códigos matriciales con la métrica del rango. Veremos que es posible reformularlo como un problema que se asemeje a la decodificación convencional.

Sea $x \in C$ y $X = \begin{pmatrix} I_n & x \end{pmatrix}$ la matriz de transmisión, donde

$$C \subseteq \mathrm{Mat}(n \times m, \mathbb{F}_q)$$

es un código con métrica del rango. Escribimos la matriz recibida en la forma

$$Y = \begin{pmatrix} \hat{A} & y \end{pmatrix}$$

donde $\hat{A} \in \mathrm{Mat}(N \times n, \mathbb{F}_q)$ y $y \in \mathrm{Mat}(N \times m, \mathbb{F}_q)$ y $\mathrm{Rang}\ Y = N$, puesto que cualquier paquete recibido linealmente dependiente no afecta el problema de decodificación y podrá ser desechado por el nodo de destino. Se define

$$\mu := n - \mathrm{Rang}\ \hat{A} \text{ y } \delta := N - \mathrm{Rang}\ \hat{A},$$

donde μ mide la deficiencia del rango de $\hat{A}$ con respecto a las columnas, mientras que δ mide la deficiencia de rango de $\hat{A}$ con respecto a las filas.

Antes de examinar el problema general, se estudia el caso especial simple que surge cuando $\mu = \delta = 0$.

3.3.5 Lema. Si $\mu = \delta = 0$, entonces

$$d(\langle X \rangle, \langle Y \rangle) = 2d_R(x, r)$$

donde $r = \hat{A}^{-1}y$.

Demostración. Si $\mu = \delta = 0$, entonces Rang $\hat{A} = n = N$ y se tiene que $\hat{A}$ es invertible. Entonces,

$$\bar{Y} = \begin{pmatrix} I_n & \hat{A}^{-1}y \end{pmatrix}$$

es equivalente por filas a Y. Es decir,

$$\langle \bar{Y} \rangle = \langle Y \rangle.$$

Aplicando el lema 3.3.2, se obtiene el resultado deseado. □

El lema anterior muestra que, cuando $\hat{A}$ es invertible, una solución de (3.2) puede encontrarse resolviendo el tradicional problema de decodificación del rango, lo ilustraremos mediante el siguiente ejemplo.

3.3.6 Ejemplo. Sea $n = 4$ y $q = 5$. Sean x_1, x_2, x_3, x_4 denotan las filas de un codeword para $x \in C$.

Supongamos que

$$A = \begin{pmatrix} 2 & 4 & 2 & 4 \\ 0 & 0 & 3 & 3 \\ 1 & 0 & 4 & 3 \\ 0 & 4 & 1 & 4 \end{pmatrix},$$

$B = \begin{pmatrix} 4 & 0 & 1 & 0 \end{pmatrix}^T$ y $z = \begin{pmatrix} 1 & 2 & 3 & 4 & z \end{pmatrix}$. Sea

$$x = \begin{pmatrix} 1 & 0 & 0 & 0 & x_1 \\ 0 & 1 & 0 & 0 & x_2 \\ 0 & 0 & 1 & 0 & x_3 \\ 0 & 0 & 0 & 1 & x_4 \end{pmatrix}.$$

Luego

$$Ax = \begin{pmatrix} 2 & 4 & 2 & 4 & 2x_1 + 4x_2 + 2x_3 + 4x_4 \\ 0 & 0 & 3 & 3 & 3x_3 + 3x_4 \\ 1 & 0 & 4 & 3 & x_1 + 4x_3 + 3x_4 \\ 0 & 4 & 1 & 4 & 4x_2 + x_3 + 4x_4 \end{pmatrix}$$

$$Bz = \begin{pmatrix} 4 & 3 & 2 & 1 & 2x_1 + 4z \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 & z \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Por lo tanto

$$Y = Ax + Bz = \begin{pmatrix} 1 & 2 & 4 & 0 & 2x_1 + 4x_2 + 2x_3 + 4x_4 + 4z \\ 0 & 0 & 3 & 3 & 3x_3 + 3x_4 \\ 2 & 2 & 2 & 2 & x_1 + 4x_3 + 3x_4 + z \\ 0 & 4 & 1 & 4 & 4x_2 + x_3 + 4x_4 \end{pmatrix}.$$

Procediendo por eliminación Gaussiana tenemos

$$\bar{Y} = \begin{pmatrix} I & r \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 4 & 0 & * \\ 0 & 0 & 3 & 3 & * \\ 2 & 2 & 2 & 2 & * \\ 0 & 4 & 1 & 4 & * \end{pmatrix} \; {\scriptstyle 4F_4 \to F_4}$$

$$\begin{pmatrix} 1 & 2 & 4 & 0 & 2x_1 + 4x_2 + 2x_3 + 4x_4 + 4z \\ 0 & 0 & 3 & 3 & 3x_3 + 3x_4 \\ 2 & 2 & 2 & 2 & x_1 + 4x_3 + 3x_4 + z \\ 0 & 1 & 4 & 1 & 4x_2 + 4x_3 + x_4 \end{pmatrix} \; {\scriptstyle \substack{F_2 \leftrightarrow F_4 \\ 3F_1 + F_3 \to F_3}}$$

$$\begin{pmatrix} 1 & 2 & 4 & 0 & 2x_1 + 4x_2 + 2x_3 + 4x_4 + 4z \\ 0 & 1 & 4 & 1 & x_2 + 4x_3 + x_4 \\ 0 & 3 & 4 & 2 & 2x_1 + 2x_2 + 3z \\ 0 & 0 & 3 & 3 & 3x_3 + 3x_4 \end{pmatrix} \; {\scriptstyle 2F_2 + F_3 \to F_3}$$

$$\begin{pmatrix} 1 & 2 & 4 & 0 & 2x_1 + 4x_2 + 2x_3 + 4x_4 + 4z \\ 0 & 1 & 4 & 1 & x_2 + 4x_3 + x_4 \\ 0 & 0 & 2 & 4 & 2x_1 + 4x_2 + 3x_3 + 2x_4 + 3z \\ 0 & 0 & 3 & 3 & 3x_3 + 3x_4 \end{pmatrix} \; {\scriptstyle F_3 + F_4 \to F_4}$$

$$\begin{pmatrix} 1 & 2 & 4 & 0 & 2x_1 + 4x_2 + 2x_3 + 4x_4 + 4z \\ 0 & 1 & 4 & 1 & x_2 + 4x_3 + x_4 \\ 0 & 0 & 2 & 4 & 2x_1 + 4x_2 + 3x_3 + 2x_4 + 3z \\ 0 & 0 & 0 & 2 & 2x_1 + 4x_2 + x_3 + 3z \end{pmatrix} \; {\scriptstyle 3F_4 + F_3 \to F_3}$$

$$\begin{pmatrix} 1 & 2 & 4 & 0 & 2x_1 + 4x_2 + 2x_3 + 4x_4 + 4z \\ 0 & 1 & 4 & 1 & x_2 + 4x_3 + x_4 \\ 0 & 0 & 2 & 0 & 3x_1 + x_2 + x_3 + 2x_4 + 2z \\ 0 & 0 & 0 & 2 & 2x_1 + 4x_2 + x_3 + 3z \end{pmatrix} \; {\scriptstyle \substack{3F_3 + F_1 \to F_1 \\ 3F_3 + F_2 \to F_2}}$$

$$\begin{pmatrix} 1 & 2 & 0 & 0 & x_1 + 2x_2 \\ 0 & 1 & 0 & 1 & 4x_1 + 4x_2 + 2x_4 + 2x_3 + z \\ 0 & 0 & 2 & 0 & 3x_1 + x_2 + x_3 + 2x_4 + 2z \\ 0 & 0 & 0 & 2 & 2x_1 + 4x_2 + x_3 + 3z \end{pmatrix} \; {\scriptstyle 2F_4 + F_2 \to F_2}$$

$$\left(\begin{array}{cccc|c} 1 & 2 & 0 & 0 & x_1+2x_2 \\ 0 & 1 & 0 & 0 & 3x_1+2x_2+4x_3+2x_4+2z \\ 0 & 0 & 2 & 0 & 3x_1+x_2+x_3+2x_4+2z \\ 0 & 0 & 0 & 2 & 2x_1+4x_2+x_3+3z \end{array}\right) \; {\scriptstyle 3F_2+F_1\to F_1}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 3x_2+2x_3+x_4+z \\ 0 & 1 & 0 & 0 & 3x_1+2x_2+4x_3+2z \\ 0 & 0 & 2 & 0 & 3x_1+x_2+x_3+2x_4+2z \\ 0 & 0 & 0 & 2 & 2x_1+4x_2+x_3+3z \end{array}\right) \; \begin{array}{l} {\scriptstyle 3F_3\leftrightarrow F_3} \\ {\scriptstyle 3F_4\leftrightarrow F_4} \end{array}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 3x_2+2x_3+x_4+z \\ 0 & 1 & 0 & 0 & 3x_1+2x_2+4x_3+2z \\ 0 & 0 & 1 & 0 & 4x_1+3x_2+3x_3+x_4+z \\ 0 & 0 & 0 & 1 & x_1+2x_2+3x_3+4z \end{array}\right)$$

Debemos tener en cuenta que, siempre que no se produzcan errores, se espera encontrar $r = x$.

$$r = \begin{pmatrix} 3x_2+2x_3+x_4+z \\ 3x_1+2x_2+4x_3+2x_4+2z \\ 4x_1+3x_2+3x_3+x_4+z \\ x_1+2x_2+3x_3+4z \end{pmatrix}.$$

Ahora, podemos escribir a r como se muestra a continuación.

$$r = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 1 \\ 4 \end{pmatrix} \begin{pmatrix} 4x_1+3x_2+2x_3+x_4+z \end{pmatrix}.$$

Por lo tanto,

$$\text{Rang } (r-x) = 1.$$

Podemos pensar en esto como un codeword de error $e = r - x$ de rango 1 aplicado a X. Este error puede ser corregido, si

$$d_R(C) \geq 3.$$

Si se piensa en el caso general, donde $\hat{A}$ no necesariamente es invertible, en primer lugar se tendrá un enfoque relativamente sencillo, que no obstante conduce a un problema de decodificación poco atractivo, sin embargo es posible demostrar que

$$d(\langle X\rangle, \langle Y\rangle) = 2 \text{ Rang } (y - \hat{A}x) + \mu - \delta$$

Demostración.

$$d(\langle X\rangle, \langle Y\rangle) = 2 \text{ Rang } (y - \hat{A}x) + \mu - \delta$$

$$\begin{aligned} d(\langle X\rangle, \langle Y\rangle) &= \dim(\langle x\rangle + \langle y\rangle) - \dim(\langle x\rangle \cap \langle y\rangle) \\ &= 2\dim(\langle x\rangle + \langle y\rangle) - \dim\langle x\rangle - \dim\langle y\rangle \\ &= 2\text{Rang} \begin{pmatrix} x \\ y \end{pmatrix} - \text{Rang } x - \text{Rang } y \\ &= 2[n + \text{Rang } (y - \hat{A}x)] - \text{Rang } x - \text{Rang } y \\ &= 2\text{Rang } (y - \hat{A}x) + n - N \\ &= 2\text{Rang } (y - \hat{A}x) + n - \text{Rang } \hat{A}x) - N + \text{Rang } \hat{A}x) \\ &= 2\text{Rang } (y - \hat{A}x) + \mu - \delta, \end{aligned}$$

donde $\mu = n - \hat{A}x$ y $\delta = N - \hat{A}x$. □

lo anterior conduce el siguiente problema de decodificación:

$$\hat{x} = \underset{x \in C}{\text{argmin}} \text{ Rang } (y - \hat{A}x). \tag{3.6}$$

Si se define un nuevo código

$$C' = \hat{A}C = \{Ax \mid x \in C\},$$

entonces una solución para (3.6) puede ser encontrada resolviendo primero

$$\hat{x}' = \underset{x' \in C'}{\text{argmin}} \text{ Rang } (y - x')$$

usando un decodificador tradicional del rango para C' y eligiendo cualquier

$$\hat{x} \in \{x \mid \hat{A}x = \hat{x}'\}$$

como una solución.

Una desventaja obvia de este enfoque es que se requiere un nuevo código C' para ser utilizado en cada instancia o nodo de la decodificación. Esto probablemente aumente la complejidad de la decodificación, ya que la existencia de un algoritmo eficiente para C no implica la existencia de un algoritmo eficiente para $C' = \hat{A}C$ para todo $\hat{A}$. Además, incluso si existen algoritmos eficientes para cada C', correr un algoritmo diferente por cada matriz recibida puede ser poco practico y hasta indeseable desde el punto de vista de la implementación.

A continuación se busca una expresión para $d(\langle X\rangle, \langle Y\rangle)$ donde la estructura de C puede ser explotada con el fin de motivar nuestro enfoque consideramos los siguiente ejemplos, los cuales generalizan el primer ejemplo.

3.3.7 Ejemplo. Consideremos nuevamente el ejemplo anterior, pero ahora supongamos

$$A = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 1 & 3 & 0 & 3 \\ 1 & 4 & 0 & 3 \\ 2 & 0 & 4 & 0 \\ 1 & 1 & 2 & 4 \end{pmatrix},$$

$B = \begin{pmatrix} 4 & 0 & 1 & 0 & 0 \end{pmatrix}^T$ y $z = \begin{pmatrix} 1 & 2 & 3 & 4 & z \end{pmatrix}$. Entonces

$$Ax = \begin{pmatrix} 1 & 0 & 2 & 3 & x_1 + 2x_3 + 3x_4 \\ 1 & 3 & 0 & 3 & x_1 + 3x_2 + 3x_4 \\ 1 & 4 & 0 & 3 & x_1 + 4x_2 + 3x_4 \\ 2 & 0 & 4 & 0 & 2x_1 + 4x_3 \\ 1 & 1 & 2 & 4 & x_1 + x_2 + 2x_3 + 4x_4 \end{pmatrix}$$

$$Bz = \begin{pmatrix} 4 & 3 & 2 & 1 & 4z \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 & z \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Por lo tanto

$$Y = \begin{pmatrix} 0 & 3 & 4 & 4 & x_1 + 2x_3 + 3x_4 + 4z \\ 1 & 3 & 0 & 3 & x_1 + 3x_2 + 3x_4 \\ 2 & 1 & 3 & 2 & x_1 + 4x_2 + 3x_4 + z \\ 2 & 0 & 4 & 0 & 2x_1 + 4x_3 \\ 1 & 1 & 2 & 4 & x_1 + x_2 + 2x_3 + 4x_4 \end{pmatrix} = \begin{pmatrix} \hat{A} & y \end{pmatrix}.$$

Aunque $\hat{A}$ no es invertible, sin embargo, podemos realizar eliminación gaussiana en Y para obtener

$$\hat{Y} = \begin{pmatrix} I & r \\ 0 & \hat{E} \end{pmatrix} \tag{3.7}$$

donde

$$r = \begin{pmatrix} 2x_1 + 2x_2 + 3x_3 + 4x_4 + 4z \\ 4x_1 + 4x_2 + 2x_3 + x_4 + z \\ 2x_1 + 4x_2 + 2x_3 + 3x_4 + 3z \\ 3x_1 + x_2 + 4x_3 + 3x_4 + 2z \end{pmatrix}$$

y

$$\hat{E} = 2x_1 + 4x_2 + x_3 + 3x_4 + 3z.$$

Observe que

$$e = r - x = \begin{pmatrix} x_1 + 2x_2 + 3x_3 + 4x_4 + 4z \\ 4x_1 + 3x_2 + 2x_3 + x_4 + z \\ 2x_1 + 4x_2 + x_3 + 3x_4 + 3z \\ 3x_1 + x_2 + 4x_3 + 2x_4 + 2z \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 4 \end{pmatrix} \hat{E}.$$

Note que si efectuamos sobre e las siguientes operaciones:

$$e \underset{F_1+F_2\to F_2}{\longrightarrow} \begin{pmatrix} f_1 \\ 0 \\ f_3 \\ f_4 \end{pmatrix} \underset{F_3+F_4\to F_4}{\longrightarrow} \begin{pmatrix} f_1 \\ 0 \\ f_3 \\ 0 \end{pmatrix} \underset{2F_3+F_1\to F_1}{\longrightarrow} \begin{pmatrix} 0 \\ 0 \\ f_3 \\ 0 \end{pmatrix}$$

no solo se obtiene que Rang $e = 1$, sino que también se recupera parte de su descomposición como producto externo, concretamente, el vector $\hat{E}$.

3.3.8 Ejemplo. Sea $n = 4$ y $q = 5$ y sea

$$A = \begin{pmatrix} 3 & 2 & 1 & 1 \\ 0 & 4 & 3 & 2 \\ 2 & 1 & 0 & 4 \end{pmatrix}$$

y si supone que no hay errores, entonces

$$Y = \begin{pmatrix} 3 & 2 & 1 & 1 & 3x_1 + 2x_2 + x_3 + x_4 \\ 0 & 4 & 3 & 2 & 4x_2 + 3x_3 + 2x_4 \\ 2 & 1 & 0 & 4 & 2x_1 + x_2 + 4x_4 \end{pmatrix} = \begin{pmatrix} \hat{A} & y \end{pmatrix}$$

Una vez más, no es posible invertir $\hat{A}$, sin embargo, después de realizar eliminación Gaussiana sobre Y e insertar una fila de ceros en la tercera posición, se obtiene

$$\begin{aligned} \hat{Y} &= \begin{pmatrix} 1 & 0 & 4 & 0 & x_1 + 4x_3 \\ 0 & 1 & 2 & 0 & x_2 + 2x_3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_4 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 4 & 0 & x_1 + 4x_3 \\ 0 & 1 & 2 & 0 & x_2 + 2x_3 \\ 0 & 0 & 1-1 & 0 & x_3 - x_3 \\ 0 & 0 & 0 & 1 & x_4 \end{pmatrix} \\ &= \begin{pmatrix} I + \hat{L}I_3^T & x + \hat{L}x_3 \end{pmatrix} \\ &= \begin{pmatrix} I + \hat{L}I_3^T & r \end{pmatrix} \end{aligned}$$

donde

$$\hat{L} = \begin{pmatrix} 4 \\ 2 \\ -1 \\ 0 \end{pmatrix} \text{ y } I_3^T = \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix}.$$

luego

$$\hat{L}I_3^T = \begin{pmatrix} 4 \\ 2 \\ -1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

entonces

$$I + \hat{L}I_3^T = \begin{pmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\hat{L}I_3^T = \begin{pmatrix} 4x_3 \\ 2x_3 \\ -x_3 \\ 0 \end{pmatrix}$$

$$X + \hat{L}I_3^T = \begin{pmatrix} x_1 + 4x_3 \\ x_2 + 2x_3 \\ 0 \\ x_4 \end{pmatrix}$$

Nuevamente puede observar que el codeword de error tiene rango 1, y que se ha recuperado parte de su descomposición como un producto externo. Es decir, se tiene que

$$e = r - x = \hat{L}x_3$$

donde $\hat{L}$ es conocido.□

3.3.9 Observación. Rang $e = 1$ en efecto,

$$e = r - x = x + \hat{L}x_3 - x = \hat{L}x_3 = \begin{pmatrix} 4x_3 \\ 2x_3 \\ -x_3 \\ 0 \end{pmatrix}.$$

Efectuando operaciones elementales adecuadas se llega a

$$\begin{pmatrix} f_1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Después de haber visto a partir de estos dos ejemplos cómo la información lateral (conocimiento parcial de la matriz de error) surge en la salida del canal, se aborda el caso general en el siguiente lema, pero antes se hará referencia a algunas propiedades de las matrices de I_U y I_{U^C}, donde $I = I_n, U \subseteq \{1, \ldots, n\}$ y $U^C = \{1, \ldots, n\} \backslash U$.

3.3.10 Definición. Sea $I_U \in \mathbb{F}_q^{n \times |U|}$, I_U se define de la siguiente manera

$$I_U = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix},$$

donde f_j son las filas de la matriz I_U para $j = 1, \ldots, n$.

Las f_j filas de I_U coincide con las j-ésima filas de I_n, y se cumple que para todo $j \in U$

$$f_j = 0.$$

Note que para cualquier $A \in \mathrm{Mat}(n \times k, \mathbb{F}_q)$, la matriz $I_U^T A$ extrae las filas de A que son indexadas por U. Mientras que para cualquier $A \in \mathrm{Mat}(k \times n, \mathbb{F}_q)$ la matriz AI_U extrae las columnas de A que son indexadas por U.

Recíprocamente, para algún $B \in \mathrm{Mat}(|U| \times k, \mathbb{F}_q)$ la matriz $I_U B$ redistribuye las filas de B a las posiciones indexadas por U. Mientras que para cualquier $B \in \mathrm{Mat}(k \times |U|, \mathbb{F}_q)$ la matriz BI_U^T redistribuye las columnas de B a las posiciones indexadas por U, en cada caso las filas o columnas nulas se insertan en las posiciones indexadas por U^C. Es importante tener en cuenta que I_{U^C} e I_U satisfacen las siguientes propiedades:

$$\begin{aligned} I &= I_U I_U^T + I_{U^C} I_{U^C}^T \\ I_U^T I_U &= I_{|U|} \\ I_U^T I_{U^C} &= 0. \end{aligned}$$

3.3.11 Lema. Sea Y, μ y δ, definidos como antes. Entonces existen matrices $r \in \mathrm{Mat}(n \times m, \mathbb{F}_q), \hat{L} \in \mathrm{Mat}(n \times \mu, \mathbb{F}_q), \hat{E} \in \mathrm{Mat}(\delta \times m, \mathbb{F}_q)$ y un conjunto $U \subseteq \{1, ..., N\}$ que satisfacen

$$|U| = \mu \tag{3.8}$$

$$I_U^T r = 0 \tag{3.9}$$

$$I_U^T \hat{L} = -I_\mu \tag{3.10}$$

$$\mathrm{Rang}\ \hat{E} = \delta \tag{3.11}$$

de tal manera que

$$\left\langle \begin{pmatrix} I + \hat{L} I_U^T & r \\ 0 & \hat{E} \end{pmatrix} \right\rangle = \langle Y \rangle. \tag{3.12}$$

Demostración. Denotaremos con $\mathrm{FER}(Y)$ la forma escalonada reducida por filas de la matriz Y.

Para $i \in \{1, \ldots, N\}$, sea p_i la posición de la columna de entrada principal de la i-ésima fila de $\mathrm{FER}(Y)$.

Sea $U^c = \{p_1, \ldots, p_{n-\mu}\}$ y $U = \{1, \ldots, n\} \backslash U^c$. Note que $|U| = \mu$. De las propiedades de la forma escalonada reducida por filas, podemos escribir

$$\mathrm{FER}(Y) = \begin{pmatrix} W & \bar{r} \\ 0 & \hat{E} \end{pmatrix}$$

donde $\bar{r} \in \mathrm{Mat}((n-\mu) \times m, \mathbb{F}_q)$, $\hat{E} \in \mathrm{Mat}(\delta \times m, \mathbb{F}_q)$ tiene rango δ, y $W \in \mathrm{Mat}((n-\mu) \times m, \mathbb{F}_q)$, la cual satisface que

$$W I_{U^C} = I_{(n-\mu)}.$$

Sea, ahora

$$\bar{Y} = \begin{pmatrix} I_{U^C} & 0 \\ 0 & I_{\delta \times \delta} \end{pmatrix} \mathrm{FER}(Y) = \begin{pmatrix} I_{U^C} W & r \\ 0 & \hat{E} \end{pmatrix}$$

donde $r = I_{U^C} \bar{r}$. Dado que $I = I_{U^C} I_{U^C}^T + I_U I_U^T$, tenemos

$$\begin{aligned} I_{U^C}W &= I_{U^C}W(I_{U^C}I_{U^C}^T + I_U I_U^T) \\ &= I_{U^C}WI_{U^C}I_{U^C}^T + I_{U^C}WI_U I_U^T \\ &= I_{U^C}I_{U^C}^T + I_{U^C}WI_U I_U^T \\ &= I - I_U I_{U^T} + I_{U^C}WI_U I_U^T \\ &= I + (-I_U + I_{U^C}WI_U)I_U^T \\ &= I + \hat{L}I_U^T. \end{aligned}$$

donde $\hat{L} = -I_U + I_{U^C}WI_U$. Por otro lado, dado que $I_U^T I_U = I_\mu$ y $I_U^T I_{U^C} = 0$, se tiene que $I_U^T\hat{L} = -I_\mu$ y $I_U^T r = 0$. Por lo tanto,

$$\bar{Y} = \begin{pmatrix} I + \hat{L}I_U^T & r \\ 0 & \hat{E} \end{pmatrix}$$

es una matriz con el mismo espacio de filas que Y, lo cual completa la demostración. □

La anterior proposición muestra que cada matriz Y es equivalente por filas a una matriz

$$\bar{Y} = \begin{pmatrix} I + \hat{L}I_U^T & r \\ 0 & \hat{E} \end{pmatrix}$$

que es esencialmente la matriz Y en forma escalonada reducida, la cual se asemeja al levantamiento de algún $r \in \mathrm{Mat}(n \times m, \mathbb{F}_q)$.

Podemos pensar en $r, \hat{L}, \hat{E}$ y el conjunto U como un suministro de una forma compacta para describir el subespacio $\langle Y \rangle$. El conjunto U es de hecho redundante y U se puede omitir de la descripción, como se muestra en el siguiente lema.

3.3.12 Lema. Sea $r \in \mathrm{Mat}(n \times m, \mathbb{F}_q)$, $\hat{L} \in \mathrm{Mat}(n \times \mu, \mathbb{F}_q)$, $\hat{E} \in \mathrm{Mat}(\delta \times m, \mathbb{F}_q)$ y $U \subseteq \{1, \ldots, n\}$ la tupla de $(r, \hat{L}, \hat{E})$ satisface (3.8)- (3.11).

Si $S \subseteq \{1, \ldots, n\}$, $T \in \mathrm{Mat}(\mu, \mathbb{F}_q)$ and $R \in \mathrm{Mat}(\delta, \mathbb{F}_q)$ tal que $(r, \hat{L}T, R\hat{E}, S)$ satisface (3.8)- (3.11), entonces

$$\left\langle \begin{pmatrix} I + \hat{L}TI_S^T & r \\ 0 & R\hat{E} \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} I + \hat{L}I_U^T & r \\ 0 & \hat{E} \end{pmatrix} \right\rangle.$$

Demostración.

De (A.5) y del hecho de que R es no singular (dado que Rang $(R\hat{E}) = \delta$), esto equivale a demostrar que

$$\left\langle \begin{pmatrix} I + \hat{L}TI_S^T & r \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} I + \hat{L}I_U^T & r \end{pmatrix} \right\rangle.$$

Sea $W_1 = I + \hat{L}I_U^T$ y $W_2 = I + \hat{L}TI_S^T$. Note que

$$W_1 I_{U^C} = I_{U^C}, I_U^T \begin{pmatrix} W_1 & r \end{pmatrix} = 0.$$

Por lo tanto $I_{U^C}^T W_1$ tiene rango máximo. Similarmente $I_S^T \begin{pmatrix} W_2 & r \end{pmatrix} = 0$ y $I_{S^C}^T W_2$ también tiene rango máximo. Entonces es suficiente demostrar que

$$M \begin{pmatrix} W_2 & r \end{pmatrix} = \begin{pmatrix} W_1 & r \end{pmatrix} \tag{3.13}$$

para algún $M \in \mathrm{Mat}(n \times n, \mathbb{F}_q)$.

Sea $A = U \cup S$ y $B = U \cap S$. Observe que M puede ser particionada en tres submatrices,

$$MI_{A^C}, MI_S \text{ y } MI_{U \setminus B}$$

escogemos $MI_{A^C} = I_{A^C}$ y MI_S arbitrariamente y a $MI_{U \setminus B}$ de modo que (3.13) se satisfaga. En primer lugar, se debe tener en cuenta que

$$Mr = M(I_{A^C} I_{A^C}^T + I_A I_A^T)r = I_{A^C} I_{A^C}^T r = r$$

De lo anterior $I_{A^C}^T r = 0$. Por lo tanto, sólo tenemos que considerar $MW_2 = W_1$ en (3.13). Por otra parte, observemos que

$$\begin{aligned} MW_2 &= M(I_{A^C} I_{A^C}^T + I_S I_S^T + I_{U \setminus B} I_{U \setminus B}^T) W_2 \\ &= I_{A^C} I_{A^C}^T W_2 + (M)(I_{U \setminus B})(I_{U \setminus B}^T W_2). \end{aligned}$$

Ahora, consideremos el sistema $MW_2 = W_1$. A partir del álgebra lineal básica, podemos resolver $MI_{U \setminus B}$ si y sólo si

$$\mathrm{Rang} \begin{bmatrix} I_{U \setminus B}^T W_2 \\ W_1 - I_{A^C} I_{A^C}^T W_2 \end{bmatrix} \leq |U \setminus B|.$$

Dado que $I_{U \setminus B}^T W_1 = 0$ y $I_S^T W_2 = 0$, podemos reorganizar las filas para obtener

$$\begin{aligned} \mathrm{Rang} \begin{pmatrix} I_{U \setminus B}^T W_2 \\ W_1 - I_{A^C} I_{A^C}^T W_2 \end{pmatrix} &= \mathrm{Rang} \begin{pmatrix} I_{U \setminus B}^T (W_1 - W_2) \\ I_{(U \setminus B)^C} I_{(U \setminus B)^C}^T (W_1 - W_2) \end{pmatrix} \\ &= \mathrm{Rang}\ (W_1 - W_2). \end{aligned}$$

Para completar la prueba, demostramos que Rang $(W_1 - W_2) \leq |U \setminus B|$. Tenemos

$$\begin{aligned}
\text{Rang } (W_1 - W_2) &= \text{Rang } (\hat{L}I_U^T - \hat{L}TI_S^T) \\
&\leq \text{Rang } (I_U^T - TI_S^T) \text{ multiplicando por } (I_S^T\hat{L} = -T^{-1}) \\
&= \text{Rang } (I_S^T\hat{L}I_U^T + I_S^T) \\
&= \text{Rang } (I_S^T W_1) \\
&= \text{Rang } (I_S I_S^T W_1) \\
&= \text{Rang } (I_{S\setminus B}I_{S\setminus B}^T + I_B I_B^T)W_1 \\
&= \text{Rang } I_{S\setminus B}I_{S\setminus B}^T W_1 \\
&\leq |S \setminus B| = |U \setminus B|. \qquad \square
\end{aligned}$$

El anterior lema muestra que, dada una tupla $(r, \hat{L}, \hat{E})$ obtenida a partir del lema (3.3.11), el conjunto U se puede encontrar como cualquier conjunto que satisfaga (3.8)- (3.11).

Por otra parte, la matriz de $\hat{L}$ puede ser multiplicada a la derecha por cualquier matriz no singular (a condición de que satisfaga los resultantes de (3.8)- (3.11) para algún U), y la matriz $\hat{E}$ se puede multiplicar a la izquierda por una matriz no singular y ninguna de estas operaciones cambian el subespacio descrito por $(r, \hat{L}, \hat{E})$. La noción de una descripción concreta de un subespacio $\langle Y \rangle$ se refleja en la siguiente definición.

3.3.13 Definición. Sea $r \in \text{Mat}(n \times m, \mathbb{F}_q)$, $\hat{L} \in \text{Mat}(n \times \mu, \mathbb{F}_q)$ y $\hat{E} \in \times\text{Mat}(\delta \times m, \mathbb{F}_q)$, si la terna $(r, \hat{L}, \hat{E})$ satisface las propiedades (3.8)-(3.12) para algún $U \subseteq \{1, \ldots, n\}$, entonces es llamada una reducción de la matriz Y.

A continuación se muestra la demostración del teorema principal de esta sección.

3.3.14 Teorema. Sea $(r, \hat{L}, \hat{E})$ una reducción de Y. Entonces

$$d(\langle X \rangle, \langle Y \rangle) = 2\text{Rang } \begin{pmatrix} \hat{L} & r - x \\ 0 & \hat{E} \end{pmatrix} - (\mu + \delta).$$

Demostración. Tenemos lo siguiente que

$$\begin{aligned}
\text{Rang}\begin{pmatrix} X \\ Y \end{pmatrix} &= \text{Rang}\begin{pmatrix} I & x \\ I+\hat{L}I_U^T & r \\ 0 & \hat{E} \end{pmatrix} \\
&= \text{Rang}\begin{pmatrix} -\hat{L}I_U^T & x-r \\ I+\hat{L}I_U^T & x \\ 0 & \hat{E} \end{pmatrix} \\
&= \text{Rang}\begin{pmatrix} \hat{L}I_U^T & r-x \\ I_{U^C}(I+\hat{L}I_U^T)^T & I_{U^C}r \\ 0 & \hat{E} \end{pmatrix} \\
&= \text{Rang}\begin{pmatrix} \hat{L}I_U^T & r-x \\ I_{U^C}^T & I_{U^C}x \\ 0 & \hat{E} \end{pmatrix} \\
&= \text{Rang}\begin{pmatrix} \hat{L}I_U^T & r-x \\ 0 & \hat{E} \end{pmatrix} + \begin{pmatrix} I_{U^C}^T & I_{U^C}x \end{pmatrix} \\
&= \text{Rang}\begin{pmatrix} \hat{L} & r-x \\ 0 & \hat{E} \end{pmatrix} + n - \mu.
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
d(\langle X\rangle, \langle Y\rangle) &= 2\ \text{Rang}\begin{pmatrix} X \\ Y \end{pmatrix} - \text{Rang}\ X - \text{Rang}\ Y \\
&= 2\ \text{Rang}\begin{pmatrix} \hat{L} & r-x \\ 0 & \hat{E} \end{pmatrix} - (\mu+\delta). \qquad \square
\end{aligned}$$

Una consecuencia del teorema anterior, es que bajo la construcción del levantamiento, el problema de decodificación (3.2) para la codificación en red aleatoria se puede abstraer a un problema de decodificación generalizada para códigos matriciales con la métrica del rango. Más precisamente, si x es enviado a través de un canal, suponiendo que x es levantado para obtener a $X = (I_n \mid x)$ y se obtiene a Y que además es reducido a la terna

$$(r, \hat{L}, \hat{E}),$$

entonces el problema (3.2) puede escribirse como se indica a continuación:

Sea $C \subseteq \text{Mat}(n \times m, \mathbb{F}_q)$ un código con métrica del rango, $r \in \text{Mat}(n \times m, \mathbb{F}_q)$, $\hat{L} \in \text{Mat}(n \times \mu, \mathbb{F}_q)$, $\hat{E} \in \times\text{Mat}(\delta \times m, \mathbb{F}_q)$, y la terna recibida $(r, \hat{L}, \hat{E})$ con Rang $\hat{L} = \mu$ y Rang $\hat{E} = \delta$, se tiene que

$$\hat{x} = \underset{x \in\ C}{\text{argmin}}\ \text{Rang}\begin{pmatrix} \hat{L} & r-x \\ 0 & \hat{E} \end{pmatrix}. \tag{3.14}$$

El anterior problema se conoce como la generalización del problema de decodificación para códigos con métricas del rango, en la siguiente sección se estudiara dicho problema y su solución para **MRD** códigos.

3.4. Generalización del problema de decodificación para códigos con métrica del rango

En esta sección desarrollamos una perspectiva sobre la generalización del problema de decodificación para códigos con métrica del rango, el cual será de utilidad para la comprensión de la capacidad de corrección que poseen códigos matriciales con métrica del rango; así como la formulación de un algoritmo eficiente de decodificación.

Ubicación de errores y valores de errores.

3.4.1 Definición. Sea $C \subseteq \mathrm{Mat}(n \times m, \mathbb{F}_q)$ un código con la métrica del rango. Para un codeword transmitido x y un codeword recibido r, se define a

$$e := r - x$$

como la palabra error.

Debemos tener en cuenta que si la palabra error tiene Rang τ, entonces podemos escribir

$$e = LE$$

para alguna matriz $L \in \mathrm{Mat}(n \times \tau, \mathbb{F}_q)$ y $E \in \mathrm{Mat}(\tau \times m, \mathbb{F}_q)$, como en (A.3), donde L y E son matrices de rango completo.

Sean $L_1, \ldots, L_\tau \in \mathrm{Mat}(1 \times n, \mathbb{F}_q)$ y $E_1, \ldots, E_\tau \in \mathrm{Mat}(1 \times m, \mathbb{F}_q)$ vectores que denotan las columnas de L y las filas de E respectivamente. Entonces es posible extender a e como una sumatoria de productos externos.

$$e = LE = \sum_{j=1}^{\tau} L_j E_j. \tag{3.15}$$

Es posible tomar prestado algo de la terminología de la teoría clásica de códigos, así que recuerde que un vector de error $e \in \mathbb{F}_q^n$ de peso τ (con respecto a la métrica de Hamming) se puede ampliar unívocamente como una suma de productos externos.

$$e = \sum_{j=1}^{\tau} I_{i_j} e_j,$$

donde $1 \leq i_1 < \ldots < i_\tau \leq n$ y $e_1, \ldots, e_\tau \in \mathbb{F}_q$. El índice i_j (o la unidad de vector I_{i_j}) especifica la ubicación del j-ésimo error, mientras e_j especifica el valor del error j-ésimo.
Análogamente, en la extensión (3.15) nos referimos a $L_1, \ldots, L_\tau$ como las ubicaciones o localizaciones de errores y a $E_1, \ldots, E_\tau$ como los correspondientes valores de errores. La localización L_j (como valor columna) indica que, para $i = 1, \ldots, n$ el j-ésimo valor de error E_j (como vector fila) ocurrió en la i-ésima fila multiplicado por el coeficiente L_{ij}. Por supuesto si se tiene que $L_{ij} = 0$ esto significa que el j-ésimo valor de error se presenta en la i-ésima fila.

Es importante mencionar que en contraste con la teoría clásica de codificación, la expansión de (3.15) no es única, ya que

$$e = LE = LT^{-1}TE$$

para cualquier T no singular $T \in \mathrm{Mat}(\tau, \mathbb{F}_q)$. Por lo tanto, estrictamente hablando, $L_1, \ldots, L_\tau$ y $E_1 \ldots, E_\tau$ son sólo unos posibles conjuntos de ubicaciones y valores de error que describen a la palabra error e.

Borraduras y desviaciones.

Ahora se reformulara el problema generalizado de decodificación en una forma que se facilite su comprensión y solución.

En primer lugar, observe que el problema (3.14) es equivalente al problema de encontrar una palabra de error $\hat{e}$, propuesta por

$$\hat{e} = \underset{e \in r-C}{\operatorname{argmin}} \ \mathrm{Rang} \begin{pmatrix} \hat{L} & e \\ 0 & \hat{E} \end{pmatrix}, \tag{3.16}$$

en el cual la salida puede ser calculada por

$$\hat{x} = r - \hat{e}.$$

3.4.2 Lema. Sea $e \in \mathrm{Mat}(n \times m, \mathbb{F}_q)$, $\hat{L} \in \mathrm{Mat}(n \times \mu, \mathbb{F}_q)$ y $\hat{E} \in \mathrm{Mat}(\delta \times \mu, \mathbb{F}_q)$. Las siguientes afirmaciones son equivalentes.

(a) $\tau^* = \mathrm{Rang} \begin{pmatrix} \hat{L} & e \\ 0 & \hat{E} \end{pmatrix}$

(b) $\tau^* - \mu - \delta$ es el valor mínimo de

$$\text{Rang } (e - \hat{L}E^{(1)} - L^{(2)}\hat{E}))$$

para todo $E^{(1)} \in \text{Mat}(\mu \times m, \mathbb{F}_q)$ y $L^{(2)} \in \text{Mat}(n \times \delta, \mathbb{F}_q)$

(c) τ^* es el valor mínimo de τ para el cual existen $L_1, \ldots, L_\tau \in \mathbb{F}_q$ y $E_1, \ldots, E_\tau \in \text{Mat}(1 \times m, \mathbb{F}_q)$ satisface que:

$$\begin{aligned} e &= \sum_{j=1}^{\tau} L_j E_j \\ L_j &= \hat{L}_j,\ j = 1, \ldots, \mu \\ E_{\mu+j} &= \hat{E},\ j = 1, \ldots, \delta. \end{aligned}$$

Demostración. Sea

$$\epsilon' = \min_{E^{(1)},\ L^{(2)}} \text{Rang } (e - \hat{L}E^{(1)} - L^{(2)}\hat{E}).$$

Demostremos inicialmente la equivalencia entre (a), (b). Del lema A.2.3 se tiene que,

$$\min_{L^{(2)}} \text{Rang } (e - \hat{L}E^{(1)} - L^{(2)}\hat{E}) = \text{ Rang } \begin{pmatrix} e - \hat{L}E^{(1)} \\ \hat{E} \end{pmatrix} - \text{Rang } \hat{E}.$$

Similarmente del lema A.2.3 tenemos

$$\begin{aligned} \min_{E^{(1)}} \text{Rang } \begin{pmatrix} e - \hat{L}E^{(1)} \\ \hat{E} \end{pmatrix} - \text{Rang } &= \min_{E^{(1)}} \text{Rang } \left(\begin{pmatrix} e \\ \hat{E} \end{pmatrix} - \begin{pmatrix} \hat{L} \\ 0 \end{pmatrix} E^{(1)} \right) \\ &= \text{ Rang } \begin{pmatrix} \hat{L} & e \\ 0 & \hat{E} \end{pmatrix} - \text{Rang } \hat{L}. \end{aligned}$$

Entonces

$$\epsilon' = \text{Rang } \begin{pmatrix} \hat{L} & e \\ 0 & \hat{E} \end{pmatrix} - \mu - \delta.$$

y así la equivalencia es mostrada.

Ahora, observe que el inciso (c) es equivalente al hecho que $\tau^* - \mu - \delta$ es el valor mínimo de ϵ para el cual existen

$$E^{(1)} \in \text{Mat}(\mu \times m, \mathbb{F}_q)$$

$$L^{(2)} \in \text{Mat}(n \times \delta, \mathbb{F}_q)$$

$$L^{(3)} \in \mathrm{Mat}(n \times \epsilon, \mathbb{F}_q)$$

y por ultimo

$$E^{(3)} \in \mathrm{Mat}(\epsilon \times m, \mathbb{F}_q)$$

que satisfacen la siguiente igualdad

$$e = \hat{L}E^{(1)} + L^{(2)}\hat{E} + L^{(3)}E^{(3)}.$$

Para mostrar la equivalencia entre (b) y (c), podemos mostrar que $\epsilon' = \epsilon''$, donde

$$\epsilon'' = \mathrm{mín}\{\epsilon \mid e = \hat{L}E^{(1)} + L^{(2)}\hat{E} + L^{(3)}E^{(3)}\}.$$

Podemos escribir ϵ'' como

$$\begin{aligned} \epsilon'' &= \min_{E^{(1)},\ L^{(2)}} \mathrm{mín}\{\epsilon \mid e - \hat{L}E^{(1)} - L^{(2)}\hat{E} = L^{(3)}E^{(3)}\} \\ &= \min_{E^{(1)},\ L^{(2)}} \mathrm{Rang}\ (e - \hat{L}E^{(1)} - L^{(2)}\hat{E}) \\ &= \epsilon. \qquad \square \end{aligned}$$

Con la ayuda del lema 3.4.2, la influencia de $\hat{L}$ y $\hat{E}$ en el problema de decodificación puede ser interpretado como sigue. Suponga que $e \in r - C$ es la única solución al problema (3.14). Entonces, e puede ser interpretado como

$$e = \sum_{j=1}^{\tau} L_j E_j,$$

donde $L_1, \ldots, L_\mu$ y $E_{\mu+1}, \ldots, E_{\mu+\delta}$ son conocidos por el decodificador. En otras palabras, el problema de decodificación se facilita, ya que el decodificador tiene información lateral acerca de la expansión de e.

Introduciremos un nuevo término para manejar el caso en que se conoce el valor de un error, pero no su ubicación a diferencia de la teoría clásica de códigos. En la extensión de (3.15) de la palabra de error, cada término $E_j L_j$ se denominará.

- Una borradura si L_j es conocido;
- Una desviación si E_j es conocido; y
- Un error total o completo (o simplemente un error), si L_j ni E_j son conocidos.

Las borraduras, desviaciones y los errores se conocen comúnmente como "erratas". Decimos que un patrón de erratas es corregible, cuando (3.14) tiene una única solución igual al codeword transmitido inicialmente.

El siguiente teorema caracteriza la capacidad de corrección de erratas de un código con métrica del rango.

3.4.3 Teorema. Un código $C \subseteq \text{Mat}(n \times m, \mathbb{F}_q)$ con métrica del rango y distancia mínima d es capaz de corregir todos los patrones de ϵ errores , μ borraduras y δ desviación si y sólo si

$$2\epsilon + \mu + \delta \leq d - 1$$

.

Demostración. Sea $x \in C$ el codeword transmitido y sea $(r, \hat{E}, \hat{L}) \in \text{Mat}(n \times m, \mathbb{F}_q) \times \text{Mat}(n \times \mu, \mathbb{F}_q) \times \text{Mat}(\delta \times m, \mathbb{F}_q)$ una terna recibida tal que

$$\text{Rang} \begin{pmatrix} \hat{L} & r - x \\ 0 & \hat{E} \end{pmatrix} = \mu + \delta + \epsilon.$$

Supongamos $x' \in C$ es otro código tal que

$$\text{Rang} \begin{pmatrix} \hat{L} & r - x' \\ 0 & \hat{E} \end{pmatrix} = \mu + \delta + \epsilon',$$

donde $\epsilon' \leq \epsilon$. Por la anterior proposición es posible escribir

$$e = r - x = \hat{L}E^{(1)} + L^{(2)}\hat{E} + L^{(3)}E^{(3)}$$

$$e' = r - x' = \hat{L}E^{(4)} + L^{(5)}\hat{E} + L^{(6)}E^{(6)}$$

para algunos $E^{(1)}, L^{(2)}, \ldots, E^{(6)}$ con las dimensiones apropiadas tal que el rango Rang $L^{(3)}E^{(3)} = \epsilon$ y Rang $L^{(6)}E^{(6)} = \epsilon'$. Por lo tanto,

$$e - e' = \hat{L}(E^{(1)} - E^{(4)}) + (L^{(2)} - L^{(5)})\hat{E} + L^{(3)}E^{(3)} + L^{(6)}E^{(6)}$$

y

$$\text{Rang } (x' - x) = \text{Rang } (e - e') \leq \mu + \delta + \epsilon + \epsilon' \leq d - 1$$

contradiciendo la distancia mínima del código C.

Recíprocamente, sea $x, x' \in C$ tales que Rang $(x' - x) = d$. Para todo μ, δ y ϵ tal que $\mu + \delta + 2\epsilon \geq d$, podemos escribir

$$x' - x = L^{(1)}E^{(1)} + L^{(2)}E^{(2)} + L^{(3)}E^{(3)} + L^{(4)}E^{(4)}$$

donde los cuatro términos anteriores tienen igual dimensiones a μ, δ y ϵ y $\epsilon' = d - \mu - \delta - \epsilon - \epsilon$, respectivamente.

Sea

$$e = L^{(1)}E^{(1)} + L^{(2)}E^{(2)} + L^{(3)}E^{(3)}$$

$$e' = -L^{(4)}E^{(4)}$$

y observamos también que $x' - x = e - e'$. Sea $r = x + e = x' + e', \hat{L} = L^{(1)}$ y $\hat{E} = E^{(2)}$. Supongamos que x es transmitida y que $(r, \hat{L}, \hat{E})$ es recibido. Entonces

$$\text{Rang} \begin{pmatrix} \hat{L} & r - x \\ 0 & \hat{E} \end{pmatrix} = \begin{pmatrix} \hat{L} & e \\ 0 & \hat{E} \end{pmatrix} = \mu + \delta + \epsilon$$

$$\text{Rang} \begin{pmatrix} \hat{L} & r - x' \\ 0 & \hat{E} \end{pmatrix} = \begin{pmatrix} \hat{L} & e' \\ 0 & \hat{E} \end{pmatrix} = \mu + \delta + \epsilon'.$$

Si se tiene que $\epsilon' = d - \mu - \delta - \epsilon \leq \epsilon$, es posible deducir que x no puede ser la única solución a (3.14) y por lo tanto el patrón de errata no se puede corregir.□

Este teorema muestra que al tomar en cuenta la información acerca de las borraduras y desviaciones cuando estas ocurren se puede aumentar la capacidad de corrección de un error en un código con métrica del rango. En efecto, supongamos que un codeword de Rang $t = \mu + \delta + \epsilon$ se aplica a un codeword, donde μ, δ y ϵ son el número de borraduras, desviaciones y errores totales, respectivamente, en el patrón de erratas. Se deduce que un decodificador convencional de rango (hace caso omiso de la información acerca de borraduras y desviaciones) puede garantizar una decodificación exitosa si

$$2t \leq \bar{d} - 1.$$

Por otro lado, la generalización del decodificador del rango requiere sólo que $2\epsilon + \mu + \delta \leq \bar{d} - 1$, o $2t \leq \bar{d} - 1 + \mu + \delta$, con el fin de garantizar la decodificación correcta. En este caso, la capacidad de corrección de error se incrementa en

$$\frac{(\mu + \delta)}{2}$$

si el decodificador es usado en lugar del decodificador convencional.

Ahora si hay una borradura de fila significa que todas las entradas de la fila son reemplazamos por un símbolo de borradura, y de manera similar para una borradura de columna. El problema de decodificación en este entorno naturalmente se define como la búsqueda de un codeword de tal manera que, cuando las entradas de borraduras en la palabra recibida sean sustituidas por los del codeword, la diferencia entre esta nueva matriz y el codeword es la que posee el menor grado posible. Ahora demostramos que este problema es un caso especial de (3.14).

En primer lugar, si obligamos a la palabra recibida r a estar en $\mathrm{Mat}(n \times m, \mathbb{F}_q)$ reemplazando cada símbolo de borradura con un símbolo arbitrario en $\mathbb{F}_q$, por ejemplo 0. Supongamos que las filas $i_1, \ldots, i_\mu$ y las columnas $k_1, \ldots, K_\delta$ son borraduras. Sea $\hat{L} \in \mathrm{Mat}(n \times \mu, \mathbb{F}_q)$ donde $\hat{L}_{i_j,j} = 1$ y $\hat{L}_{i,j} = 0$, para todo $i \neq i_j$, para $j = 1, \ldots, \delta$ sea $\hat{E} \in \mathrm{Mat}(\delta \times m, \mathbb{F}_q)$ dado por $\hat{E}_{j,k_j} = 1$ y $\hat{E}_{j,k} = 0$, para todo $k \neq k_j$, $j = 1, \ldots, \delta$. Dado que

$$\begin{pmatrix} \hat{L} & r - x \\ 0 & \hat{E} \end{pmatrix} = \begin{pmatrix} \hat{L} & r \\ 0 & \hat{E} \end{pmatrix} - \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} \tag{3.17}$$

es fácil ver que podemos realizar operaciones sobre las columnas en (3.17) para reemplazar las filas borradas de r con la mismas entradas que x, y de igual forma nos es posible realizar operaciones de renglón en dicha ecuación para reemplazar columnas con borraduras de r con las mismas entradas que x.

El problema de decodificación (3.14) no se modifica por la realización de estas operaciones y se reduce exactamente a el problema de decodificación con borraduras de filas y columnas descrito en el párrafo anterior. Un ejemplo de ello se da a continuación.

3.4.4 Ejemplo. Sea $n = m = 3$. Supongamos que la tercera fila y la segunda columna han sufrido borraduras en la palabra recibida. Entonces

$$r = \begin{pmatrix} r_{11} & 0 & r_{13} \\ r_{21} & 0 & r_{23} \\ 0 & 0 & 0 \end{pmatrix}, \quad \hat{L} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \hat{E} = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$$

dado que

$$\begin{pmatrix} 0 & r_{11} & 0 & r_{13} \\ 0 & r_{21} & 0 & r_{23} \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ y } \begin{pmatrix} 0 & r_{11} & x_{12} & r_{13} \\ 0 & r_{21} & x_{22} & r_{23} \\ 1 & x_{31} & x_{32} & x_{33} \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

son equivalentes por filas, entonces obtenemos que

$$\begin{aligned} \text{Rang} \begin{pmatrix} \hat{L} & r-x \\ 0 & \hat{E} \end{pmatrix} &= \text{Rang} \begin{pmatrix} 0 & r_{11}-x_{11} & 0 & r_{13}-x_{13} \\ 0 & r_{21}-x_{21} & 0 & r_{23}-x_{23} \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ &= 2 + \text{Rang} \begin{pmatrix} r_{11}-x_{11} & 0 & r_{13}-x_{13} \\ r_{21}-x_{21} & 0 & r_{23}-x_{23} \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

la cual es esencialmente la misma función objetivo que fue descrita anteriormente.

Mientras que las borraduras existentes en las filas o columnas son casos especiales, también es cierto que este último puede ser siempre transformado en el primero. Esto se puede realizar multiplicando todos los codewords con métrica del rango a la izquierda y a la derecha por matrices no singulares de tal manera, que la correspondiente $\hat{L}_j$ y $\hat{E}_j$ son convertidos en vectores unitarios. El inconveniente de esta aproximación, como se ha señalado es que la estructura del código cambia en cada instancia de decodificación, lo que puede aumentar la complejidad y los problemas de aplicación.

Por lo tanto, es más practico fijar la estructura del código y construir un decodificador que se puede manejar con las nociones generalizadas de borraduras y desviaciones. Enfoque que adoptaremos a continuación.

3.5. La métrica del rango y los códigos de Gabidulin

En esta sección, dirigimos nuestra atención hacia el diseño de un decodificador eficiente para códigos con métrica del rango, el cual pueda corregir cualquier patrón de ϵ errores, μ borraduras y δ desviaciones, los cuales satisfacen

$$2\epsilon + \mu + \delta \leq \bar{d} - 1,$$

donde $\bar{d}$ es la distancia mínima del código. El decodificador será aplicado a los códigos de Gabidulin, una clase de códigos MRD, por tal razón otro importante objetivo se está sección es la creación de un algoritmo que genere

este tipo de códigos así como su implementación.

Los códigos con la métrica del rango en $\mathrm{Mat}(n \times m, \mathbb{F}_q)$ son típicamente construidos como códigos de bloque de longitud n sobre la extensión campo $\mathbb{F}_{q^m}$.

Más precisamente, mediante la fijación de una base para $\mathbb{F}_{q^m}$ como un espacio de vectorial m-dimensional sobre $\mathbb{F}_q$, se puede considerar cualquier elemento de $\mathbb{F}_{q^m}$ como un vector fila de longitud m sobre $\mathbb{F}_q$ (y viceversa). Del mismo modo, podemos considerar cualquier vector columna de longitud n sobre $\mathbb{F}_{q^m}$ como una matriz $n \times m$ sobre $\mathbb{F}_q$ (y viceversa).

Todos los conceptos que fueron definidos previamente para matrices en $\mathrm{Mat}(n \times m, \mathbb{F}_q)$ pueden ser naturalmente aplicados a los vectores en $\mathbb{F}_{q^m}^n$; en particular, al rango de un vector $x \in \mathbb{F}_{q^m}^n$ como un elemento de $\mathrm{Mat}(n \times m, \mathbb{F}_q)$.

1. ***Polinomios linealizados:*** Una clase de polinomios que juegan un papel importante en el estudio de los códigos con la métrica del rango son los polinomios linealizados, un polinomio linealizado (o q-polinomio)

   sobre $\mathbb{F}_{q^m}$ es un polinomio de la forma

   $$f(x) = \sum_{i=o}^{t} f_i x^{[i]},$$

   donde $f_i \in \mathbb{F}_{q^m}$ si $f_t \neq 0$ llamamos a t el grado de $f(x)$. Los polinomios linealizados reciben su nombre debido a la siguiente propiedad:

   Para todo $a_1, a_2 \in \mathbb{F}_q$ y $\beta_1, \beta_2 \in \mathbb{F}_{q^m}$,

   $$f(a_1\beta_1 + a_2\beta_2) = a_1 f(\beta_1) + a_2 f(\beta_2).$$

   Es decir, la evaluación de un polinomio linealizado es una función de $\mathbb{F}_q$ a $\mathbb{F}_{q^m}$ la cual es lineal sobre $\mathbb{F}_q$. En particular, el conjunto de todas las raíces en $\mathbb{F}_{q^m}$ de un polinomio linealizado es un subespacio de $\mathbb{F}_{q^m}$.

   Sea $A(x)$ y $B(x)$ polinomios linealizados con q-grados t_A y t_B, respectivamente.

   El producto de $A(x)$ y $B(x)$ se define como el polinomio

   $$A(x) \otimes B(x) := A(B(x)).$$

Es fácil verificar que

$$P(x) = A(x) \otimes B(x)$$

es un polinomio linealizado de q-grado $t = t_A + t_B$ cuyos coeficientes pueden ser calculados de la siguiente manera:

$$P_\ell = \sum_{i=\text{máx}\{0,\ \ell - t_B\}}^{\text{mín}\{\ell,\ t_A\}} A_i B_{\ell-i}^{[i]} = \sum_{j=\text{máx}\{0,\ \ell - t_A\}}^{\text{mín}\{\ell,\ t_B\}} A_{\ell-j} B_j^{[\ell-j]}$$

para $\ell = 0, \ldots, t$. En particular, si $t_A \leq t_B$, luego

$$P_\ell = \sum_{i=0}^{t_A} A_i B_{\ell-j}^{[i]}, \quad t_A \leq t \leq t_B, \tag{3.18}$$

mientras que si $t_B \leq t_A$, entonces,

$$P_\ell = \sum_{j=0}^{t_B} A_{\ell-j} B_j^{[\ell-j]}\ , \quad t_B \leq t \leq t_A. \tag{3.19}$$

Se sabe que el conjunto de polinomios linealizados sobre $\mathbb{F}_{q^m}$ junto con las operaciones de suma de polinomios y multiplicación simbólica forman un anillo no conmutativo con identidad.

Aunque este anillo no es conmutativo, tiene muchas de las propiedades de un dominio entero euclidiano, incluyendo por ejemplo, la no existencia de divisores de cero.

Se Define ahora el q-opuesto de un polinomio linealizado

$$f(x) = \sum_{i=o}^{t} f_i x^{[i]}$$

como el polinomio

$$\bar{f}(x) = \sum_{i=o}^{t} \bar{f}_i x^{[i]}.$$

Donde $\bar{f}_i = f_{t-i}^{[i-t]}$ para $i = 0, \ldots, t$ (cuando t no sea especificado asumimos que t es el q-grado de $f(x)$).

Para un conjunto $S \subseteq \mathbb{F}_{q^m}$, se define el mínimo polinomio linealizado de S con respecto a $\mathbb{F}_{q^m}$ y se denota por $M_S(x)$ o mín poly$\{S\}(x)$, como el polinomio linealizado mónico sobre $\mathbb{F}_{q^m}$ de menor grado cuyo espacio de raíces contiene al conjunto S.

Además $M_S(x)$ también se puede definirse como

$$M_S(x) := \prod_{\beta \in \langle S \rangle} (X - \beta)$$

por lo que el $q-$grado de $M_S(x)$ es igual a la dim$\langle S \rangle$. Por otra parte, si $f(x)$ es un polinomio linealizado cuyo espacio raíces contiene a S, entonces

$$f(x) = Q(x) \otimes M_S(x)$$

para algún polinomio linealizado $Q(x)$. Esto implica que

$$M_{S \cup \{\alpha\}}(x) = M_{M_S(\alpha)}(x) \otimes M_S(x)$$

para cualquier α. Entonces $M_S(x)$ se puede calcular en $O(t^2)$ operaciones de $\mathbb{F}_{q^m}$, formando una base $\{\alpha_1, \ldots, \alpha_t\}$ para $\langle S \rangle$ y calculando $M_{\{\alpha_1,\ldots,\alpha_t\}}(x)$ de forma recursiva para $i = 1, \ldots, t$.

El diseño de un algoritmo para esta clase de polinomios y su implementación puede ser visto en el apéndice en la página

2. ***Códigos de Gabidulin:*** Iniciaremos dando a conocer una definición formal para este tipo de códigos.

3.5.1 Definición. Sea $g = (g_1, g_2, \ldots, g_n)$ un conjunto ordenado con $n \leq m$ elementos de $\mathbb{F}_{q^m}$, los cuales son linealmente independientes sobre $\mathbb{F}_q$.

El código de Gabidulin de soporte g y dimensión k es el código generado por la matriz.

$$G_{g,k} = \begin{pmatrix} g_1^{[0]} & g_2^{[0]} & \cdots & g_n^{[0]} \\ g_1^{[1]} & g_2^{[1]} & \cdots & g_n^{[1]} \\ \vdots & \vdots & \ddots & \vdots \\ g_1^{[k-1]} & g_2^{[k-1]} & \cdots & g_n^{[k-1]} \end{pmatrix},$$

donde $[i] = q^i$. La distancia mínima de un código de Gabidulin es $\bar{d} = n-k+1$, satisfaciendo la cota de Singleton en la métrica del rango.

Note que esta definición es semejante a la de códigos de Reed-Solomon, en donde un conjunto de diferentes elementos son reemplazados por un conjunto de elementos linealmente independientes, y la potencia g_i^j es reemplazada por la "potencia de Frobenius" $q_i^{[j]}$, esta clase de códigos conocidos como códigos Reed-Solomon pueden ser igualmente vistos, como la evaluación de polinomios de grado menor que k sobre un conjunto de n elementos.

De manera equivalente se puede obtener una interpretación para códigos de Gabidulin, en este caso estos códigos se generaran a través de la evaluación de polinomios linealizados sobre un conjunto de n elementos linealmente independientes.

Sustituiremos x^{q^i} por $x^{[i]}$ para mayor claridad, teniendo a q fijo. Por lo tanto bajo esta notación, un polinomio linealizado sobre $\mathbb{F}_{q^m}$ puede ser escrito como

$$f(x) = \sum_{i=0}^{t} f_i x^{[i]}.$$

Claramente el código de Gabidulin $G_{g,k}$ es el conjunto de codewords

$$x = f(g) = (f(g_1), f(g_2), \ldots, f(g_n)),$$

donde $f(x)$ es algún polinomio linealizado de grado menor que t.

Diseño de un algoritmo para generar códigos de Gabidulin sobre un cuerpo binario.

Para crear este tipo de códigos se hizo uso del siguiente algoritmo, y se consideraros inicialmente estas asignaciones que $t := \bar{d}$ y $a_i := f_i$.

Procedimiento: Determinación de n elementos linealmente independientes sobre $\mathbb{F}_{q^m}$ y creación de polinomios linealizados de grado a lo mas d,

$$g = (g_1, g_2, \ldots, g_n)$$

$$f(x) = \sum_{i=0}^{d} a_i x^{q^i},$$

donde $a_i \in \mathbb{F}_{q^m}$, donde $q = 2$.

entrada: n y k.

salida: Un código de Gabidulin con parámetros $[n, k, n-k+1]$.

Inicio

$U :=$ El conjunto de elementos de $\mathbb{F}_{2^m}$ pasados a vectores.
$M :=$ El conjunto de combinaciones $d+1$ de U.
$N :=$ El conjunto de combinaciones $d-1$ de elementos de $\mathbb{F}_{2^c}$.

si $d < 0$ o $n < 0$ **entonces**:
imprima Error el valor 2^d es el grado de un polinomio, por lo tanto d debe ser mayor que 1.

si no:

para i desde 0 hasta $|M|$:

s=0
defina el espacio vectorial sobre el cual se desea trabajar así como su base.
crear una lista Q para guardar los elementos linealmente independientes.

si $|base| = n$:
guarde el elemento que es base.

si no:
guarde la base.

imprima algún elemento guardado en Q.

fin para

fin si no

fin si

para $ii < |N|$:
Crear los polinomios linealizados de grado menor que d

$s := \sum_{i=0}^{d} a_{[i][j]} x^{2^i}$
si $s \neq 0$:

imprima s

fin si

fin para

Evaluar los elementos de la base de la lista Q en todos los polinomios linealizados s.

fin

En la siguiente sección se mostrara los cálculos efectuados con SAGE, a medida que se explique la implementación del algoritmo en este programa, se generaran un código de Gabidulin con parámetros $[2, 2]$.

Cálculos realizados con sage

Suponga que desea generar un código lineal de Gabidulin con parámetros $[2, 2]$, entonces la distancia mínima para este código será $\bar{d} = n-k+1 = 2$, es decir que estaremos bajo un código de Gabidulin con parámetros $[2, 2, 2]$.

En este caso en particular se estaría trabajando con $\mathbb{F}_4 = \{0, 1, a+1, a\}$, los elementos de $\mathbb{F}_4$ deben ser pasados a representación vectorial; para

esto se usa la instrucción `vector`, se hace esto con el fin de escoger conjuntos de vectores linealmente independientes, entonces los vectores serán visto de la siguiente manera:

$$U = \{(0,0);(1,0);(1,1);(0,1)\}.$$

Para lo anterior se crea la lista $U = [z.\texttt{vector}() \text{ for } z \quad in \quad Z]$, donde Z son los elemento de $\mathbb{F}_4$, los distintos conjuntos de vectores linealmente independientes de U que se obtendrán y guardaran en una lista Q serán:

$$Q = [\{(0,1),(1,1)\};\{(1,0),(1,1)\};\{(1,0),(0,1)\}]$$

lo anterior se logra indicándole al programa que haga n combinaciones de los elementos de U usando la función `Combinations` y con la función `randint` se elegirá alguno de esos conjuntos aleatoriamente con el fin de evaluarlo en todos los polinomios linealizados de grado menor que 2, en tal caso suponga que el conjunto elegido aleatoriamente es

$$NL = [\{(0,1),(1,1)\}]$$

tal conjunto debe ser pasado a su representación polinómica nuevamente y llamados de otra manera, antes de ser evaluados en los polinomios linealizados, es decir,

$$pp = NL = [a, a+1],$$

lo cual se logra realizando un bucle anidado con la instrucción `for` cuyo objetivo será que la variable del polinomio, en este caso a recorra todos los elementos vectoriales del conjunto NL.

El paso a seguir ahora es crear los polinomios linealizados de grado menor que 2 con coeficientes en en el cuerpo finito $\mathbb{F}_4$, en tal caso $m = 2$ y $d = 2$ los cuales son los valores de entrada. Para cálculos sobre

$$\mathbb{F}_q^m,$$

se necesita fijar una extensión algebraica de grado m sobre $\mathbb{F}_q$. Para hacer esto en SAGE se usa la instrucción $\texttt{GF}(q,a)$, entonces la instrucción antes mencionada debe escribirse $\texttt{GF}(2^2,a)$, escrito de esta manera se generan los elementos del cuerpo finito $\mathbb{F}_4$ el cual escribe sus elementos usando la variable a, este conjunto lo guardamos como una lista digitando

$$\texttt{L} = \texttt{G.list()}.$$

Dado que $\bar{d} = 2$, las combinaciones que se deben crear serán $d + 1$, es decir de tres elementos, con tal fin se usa la instrucción `Tuple`, dicha instrucción crea combinaciones en una lista, la cual nos permite obtener los coeficientes para cada polinomio, tal combinación se guardara en una lista de la siguiente manera

$$\texttt{N} = \texttt{Tuples(L,m+1).list()}.$$

El siguiente paso es recorrer la lista anterior, multiplicando cada coeficiente encontrado por x^{p^i}, donde $p = 2$ para $i = 0, 1, 2$, ese recorrido se realizara usando la instrucción `while` como se indica a continuación.

```
ii=0
while ii<len(N):
      s=0
      k=0
      while k<(m+1):
            s+=N[ii][k]*(x^2^(k))
            k=k+1
      ii=ii+1
      print '   ',s
      C+=[s]
print 'Number of linearized polynomial', len(C)
```

Ahora se debe evaluar cada elemento de *pp* en los polinomios anteriormente creados, para ello se usa nuevamente la instrucción `while` cuidando que el codeword $[0, 0]$ de longitud $n = 2$ sea incluido, ya que SAGE no entiende la evaluación del vector

$$NL = [\{(0, 1), (1, 1)\}]$$

en el polinomio cero, eso se logra con una instrucción `if` dentro de un `while` en donde se le indica al programa que debe hacer una repetición del elemento cero cada vez que se desee evaluar algún elemento en el polinomio cero, así:

```
l=0
u=0
while u<len(C):
      CL=[]
      y=0
      if u==0:
          while y<len(pp):
                 CC=0
                 y=y+1
                 CL+=[CC]
          GAB+=[CL]
      else:
          while y<len(pp):
                CC=C[u](pp[y])
                GP=CC.full_simplify()
                y=y+1
                CL+=[R(GP)]
          GAB+=[CL]
      u=u+1
```

Note que la evaluación se realiza con `CC=C[u](pp[y])`, pero si solo se deja así; el programa arrojara los polinomios en factores por tal motivo en la linea siguiente se usa la función `full_simplify` aplicada al codeword `CC` para que lo exprese de forma simplificada.

Posteriormente se le indicamos al programa que muestre lo que deseamos visualizar usando la instrucción `print` si desea incluir algún mensaje, este debe escribirse dentro de comillas y después una coma con la función a imprimir.

Para finalizar se debe poner a correr el programa y así el programa arrojara un polinomio de Gabidulin con parámetros $[2, 2, 2]$, para ello lo que se hace es llamar al programa en una nueva celda de trabajo en la misma `Notebook` el nombre de nuestro programa el cual se definió como `gab`$(2, 2, 2)$, indicándole así que $n = 2$ $k = 2$ y por lo tanto $\bar{d} = 2$, lo que SAGE nos mostrara al dar clic en la opción `evaluate` es lo siguiente:

```
Linearly independent vectors [(0, 1), (1, 1)]
Linearly independent vectors  in polynomial representation [a, a + 1]
Linearized polynomial of linear degree less than 2 :
    0
    a*x
    (a + 1)*x
    x
    a*x^2
    a*x^2 + a*x
    a*x^2 + (a + 1)*x
    a*x^2 + x
    (a + 1)*x^2
    (a + 1)*x^2 + a*x
    (a + 1)*x^2 + (a + 1)*x
    (a + 1)*x^2 + x
    x^2
    x^2 + a*x
    x^2 + (a + 1)*x
    x^2 + x
Number of linearized polynomial 16
The Gabidulin code with parameters [2, 2, 1] is:
CODE_GAB= [[0, 0], [a^2, a^2 + a], [a^2 + a, a^2 + 1], [a, a + 1], [a^3,
a^3 + a], [a^3 + a^2, a^3 + a^2], [a^3 + a^2 + a, a^3 + a^2 + a + 1],
[a^3 + a, a^3 + 1], [a^3 + a^2, a^3 + a^2 + a + 1], [a^3, a^3 + 1], [a^3
+ a, a^3 + a], [a^3 + a^2 + a, a^3 + a^2], [a^2, a^2 + 1], [0, a + 1],
[a, 0], [a^2 + a, a^2 + a]]
```

Figura 3.1: Código de Gabidulin con parámetros $[2, 2, 2]$

Note que el programa nos permite visualizar los vectores L.I (linealmente independiente) los cuales son elegidos aleatoriamente, es decir, que cada vez que se pulse la instrucción `evaluate`, estos vectores posiblemente cambiaran, además el programa también permite visualizar estos mismos vectores pero en representación polinómica, así como los polinomios linealizados de grado menor que d y la evaluación de los vectores en los polinomios, los cuales serán los codewords del código de Gabidulin, el cual le indicamos al programa que mostrara como `CODE_GAB`.

En la siguiente página se muestra la programación completa para códigos de Gabidulin con parámetros deseados, recuerde que cuando implemente el programa el n que se debe ingresar tiene que ser menor o igual que m en tal caso será un código de Gabidulin los cuales también son MRD-códigos, es decir un código de Gabidulin es de longitud a lo mas m.

Programación completa de un algoritmo que genera códigos de Gabidulin en SAGE.

```
def gab_yz(m,K):
    n=m
    var('w,a')
    qq=2^(n)
    GG=GF(qq,'a')
    Z=GG.list()
    nset=[y for y in Z if y!=0]
    N=Combinations(nset,n).list()
    U=[z.vector() for z in Z]
    mset = [w for w in U if w!=0]
    M=Combinations(mset,n).list()
    Q=[]
    i=0
    while i<len(M):
          V=VectorSpace(GF(2), n)
          V= V.subspace(M[i])
          base=V.basis()
          if len(base)==n:
              Q+=[M[i]]
          else:
              Q+=[base]
          i=i+1
    j=randint(0,len(M)-1)
    NL=Q[j]
    print'Linearly independent vectors g=', NL
    pp=[]
    for t in range(len(NL)):
        ss=0
        for g in range(n):
            ss+=NL[t][g]*(a^(g))
            #if ss!=0:
        pp+=[ss]
    print 'Linearly independent vectors  in polynomial representation',  pp
    mj=K
    r=m
    d=mj-1
    q=2^(r)
    G= GF(q,'a')
    var('x','a')
    L=list(G)
    N=Tuples(L,d+1).list()
    C=[]
    print'Linearized polynomial of linear degree less than',mj,':'
```

```
ii=0
while ii<len(N):
      s=0
      k=0
      while k<(d+1):
            s+=N[ii][k]*(x^2^(k))
            k=k+1
      ii=ii+1
      print '    ',s
      C+=[s]
#print 'Number of linearized polynomial', len(C)

GAB=[]
CC=[]
SS=[]
R=PolynomialRing(Zmod(2),'a')
l=0
u=0
while u<len(C):
      CL=[]
      y=0
      if u==0:
          while y<len(pp):
                 CC=0
                 y=y+1
                 CL+=[CC]
          GAB+=[CL]
      else:
          while y<len(pp):
                CC=C[u](pp[y])
                GP=CC.full_simplify()
                y=y+1
                CL+=[R(GP)]
          GAB+=[CL]
      u=u+1
print'The Gabidulin code with parameters',[n,mj,n-mj+1],'is:'
print 'CODE_GAB=',GAB
print 'Number of codeword is:',len(GAB)
```

3. ***Decodificación de códigos de Gabidulin:*** Recordemos, que en el problema de decodificación convencional del rango con τ errores, donde $2\tau \leq d-1$, y donde la palabra recibida $r \in \mathbb{F}_{q^m}^n$ esta dada, lo que se deseaba era encontrar una palabra de error única

$$e \in r - C$$

tal que el Rang $e = \tau$.

A continuación se revisa el procedimiento usual de decodificación, que consiste en encontrar los valores de errores $E_1, \ldots, E_\tau \in \mathbb{F}_{q^m}$ y las ubicaciones de los errores $L_1, \ldots, L_\tau \in \mathbb{F}_q^n$ tales que

$$e = \sum_{j=1}^{\tau} L_j E_j.$$

Dado que $e \in r - C$, podemos formar los *síndromes*

$$(S_0, \ldots, S_{d-2})^T := H_r = H_e$$

los cuales pueden ser relacionados con los valores de error y las ubicaciones de error según

$$\begin{aligned} S_\ell = \sum_{i=1}^{n} h_i^{[\ell\ ]} e_i &= \sum_{j=1}^{\tau} L_{ij} E_j \\ &= \sum_{j=1}^{\tau} X_j^{[\ell\ ]} E_j,\ \ell = 0, \ldots, d-2 \end{aligned} \tag{3.20}$$

donde

$$X_j = \sum_{i=1}^{n} L_{ij} h_i,\ j = 1, \ldots, \tau \tag{3.21}$$

son llamados los localizadores de error asociados con $L_1, \ldots, L_\tau$.

Supongamos, por ahora, que los valores de errores $E_1, \ldots, E_\tau$ (donde los τ elementos son linealmente independientes y además satisfacen $\langle e \rangle = \langle E_1, \ldots, E_\tau \rangle$) ya han sido identificados. Entonces, el localizador de errores puede determinarse resolviendo (3.20) o, de forma equivalente, mediante la solución de

$$\bar{S}_\ell = S_{d-2-\ell}^{[\ell-d+2]} = \sum_{j=1}^{\tau} E_j^{[\ell-d+2]} X_j,\ \ell = 0, \ldots, d-2 \tag{3.22}$$

que es un sistema de ecuaciones de la forma

$$B_\ell = S_{d-2-\ell}^{[\ell-d+2]} = \sum_{j=1}^{\tau} A_j^{[\ell]} X_j,\ \ell = 0, \ldots, d-2 \tag{3.23}$$

que consiste en $d-1$ ecuaciones lineales (sobre $\mathbb{F}_{q^m}$) de τ variables $X_1, \ldots, X_\tau$. Tal sistema tiene una única solución (cuando existe), siempre que $\tau \leq d-1$ y $A_1, \ldots, A_\tau$ sean linealmente independientes (Ver [15], [16]).

Después de que el error ha sido encontrado, $L_1, \ldots, L_\tau$ pueden ser recuperados fácilmente por la solución de (3.21). Formalmente, sea $h \in \mathrm{Mat}(n \times m, \mathbb{F}_q)$ la matriz cuyas filas son $h_1, \ldots, h_n$, y sea $Q \in \mathrm{Mat}(n \times m, \mathbb{F}_q)$ la inversa por la derecha de h, es decir, $hQ = I_{n\times n}$, entonces

$$L_{ij} = \sum_{k=1}^{m} X_{jk} Q_{ki},\ i = 1, \ldots, n,\ j = 1, \ldots, \tau.$$

El cálculo de los valores de error pueden hacerse indirectamente a través de un polinomio de error $\sigma(x)$.

Sea $\sigma(x)$ un polinomio linealizado con $q-$grado τ que tiene como raíces todas las combinaciones lineales de $E_1, \ldots, E_\tau$. Entonces, $\sigma(x)$ puede estar relacionada con el síndrome polinomial

$$S(x) = \sum_{j=0}^{d-2} S_j x^{[j]} \tag{3.24}$$

a través de la siguiente ecuación:

$$\sigma(x) \otimes S(x) \equiv \rho(x) \bmod x^{[d-1]} \tag{3.25}$$

donde $\rho(x)$ es un polinomio linealizado con q-grado menor o igual que $\tau - 1$. Una forma equivalente de expresar (3.25) es

$$\sum_{i=0}^{\tau} \sigma_i S_{\ell-i}^{[i]} = 0,\ \ell = \tau, \ldots, d-2. \tag{3.26}$$

Después de que el polinomio de errores es calculado, los valores de error se pueden conseguir mediante el cálculo de una base $E_1, \ldots, E_\tau$ para el espacio de la raíz de $\sigma(x)$. Este puede hacerse usando bien sea el algoritmo propuesto en [18] o por el método propuesto en [19].

Capítulo A

Rango de una matriz

A.1. Preliminares

Sean V y W espacios vectoriales sobre K con dimensiones finitas y sea $A \in \hom_K(V, W)$.

Problema. ¿Cómo pueden elegirse bases para V y W de tal forma que la representación matricial de A con respecto a estas bases tenga forma lo mas sencilla posible?

Daremos respuesta a esta pregunta en esta sección. Para ello introducimos el concepto de rango de una matriz, el cual jugará un papel importante a la hora de construir la respuesta.

A.1.1 Definición. Sea R un anillo con elemento identidad. Una matriz de tipo (m, n) sobre R es un esquema rectangular de la forma

$$(a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

con $a_{ij} \in R$, que consta de m filas

$$f_j = (a_{j1} \cdots a_{jn}),\ j = 1, \ldots, m$$

y n columnas

$$c_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix}, \ k = 1, \ldots, n.$$

El conjunto de todas las matrices de tipo $n \times m$ sobre R lo notaremos con $\mathrm{Mat}(n \times m, K)$. Las matrices del tipo $n \times n$ se llamaran cuadradas y el conjunto de estas se notará con $\mathrm{Mat}(n, K)$. Si $a_{ij} = 0$, para todo $i > j$, entonces A se llama **triangular inferior**, si $a_{ij} = 0$, para todo $i < j$, entonces A se llama **triangular superior** y si $a_{ij} = 0$, para todo $i \neq j$, entonces A se llama **triangular diagonal**.

A.1.2 Definición. Sean K un cuerpo, $A, B \in \mathrm{Mat}(n \times m, K)$, con $A = (a_{ij})$ y $B = (b_{ij})$. Definimos las siguientes operaciones:

$$(a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij}), \text{ para todo } i, j.$$

$$k(a_{ij}) := (ka_{ij}), \text{ para todo } k \in K.$$

Se verifica que $\mathrm{Mat}(n \times m, K)$ es un espacio vectorial sobre K de dimensión mn y por lo tanto isomorfo a K^{mn}.

A.1.3 Definición. Sea $A = (a_{ij}) \in \mathrm{Mat}(n \times m, K)$. Consideremos los vectores filas

$$f_j = (a_{j1}, , a_{jn}), \ j = 1, \ldots, m$$

y los vectores columnas

$$c_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix}, \ k = 1, \ldots, n$$

de A como elementos de K^n y K^m respectivamente y definamos

$$r_f(A) := \langle f_1, \ldots, f_m \rangle \tag{A.1}$$

$$r_c(A) := \langle c_1, \ldots, c_n \rangle. \tag{A.2}$$

Llamaremos a $r_f(A)$ el **rango (por filas)** de A y a $r_c(A)$ el **rango (por columna)** de A. Es claro que $r_f(A), r_c(A) \leq \text{mín}\{m, n\}$.

A.1.4 Ejemplo. Cálculo de rangos.

(a) Sea $A \in \mathrm{Mat}(3 \times 4, K)$ dada por

$$A = \begin{pmatrix} 1 & -1 & 2 & 1 \\ 2 & 0 & 10 & 5 \\ 3 & -3 & 6 & 3 \end{pmatrix}.$$

Note que

$$\begin{pmatrix} 1 & -1 & 2 & 1 \\ 2 & 0 & 10 & 5 \\ 3 & -3 & 6 & 3 \end{pmatrix} \overset{-3F_1+F_3 \to F_3}{=} \begin{pmatrix} 1 & -1 & 2 & 1 \\ 2 & 0 & 10 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Dado que los vectores filas primero y segundo son linealmente independientes se sigue que $r_f(A) = 2$.

(b) Sea $A \in \mathrm{Mat}(4, K)$ dada por

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

Dada la independencia lineal de las filas y de las columnas, se sigue inmediatamente que $r_f(A) = r_c(A) = 4$.

(c) Sean $a, b \in K$, $n \geq 2$ y $A \in \mathrm{Mat}(n \times m, \mathbb{F}_q)$ definida por

$$A = \begin{pmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & & & & \vdots \\ b & b & b & \cdots & a \end{pmatrix}.$$

Si $a = b = 0$, entonces $r_f(A) = r_c(A) = 0$ y si $a = b \neq 0$, entonces $r_f(A) = r_c(A) = 1$. Supongamos que $a \neq b$. Entonces efectuando para cada $j \in \{2, \ldots, n\}$ las operaciones $-F_1 + F_j \to F_j$, se obtiene la matriz

$$\begin{pmatrix} a & b & b & \cdots & b & b \\ b-a & a-b & 0 & \cdots & 0 & 0 \\ b-a & 0 & a-b & \cdots & 0 & 0 \\ \vdots & & & & & \vdots \\ b-a & 0 & 0 & \cdots & 0 & a-b \end{pmatrix}.$$

Efectuando la operación $C_1+\cdots+C_n \to C_1$, (se suman todas las columnas a la columna 1) se obtiene la matriz

$$\begin{pmatrix} a+(n-1)b & b & b & \cdots & b \\ 0 & a-b & 0 & \cdots & 0 \\ 0 & 0 & a-b & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & a-b \end{pmatrix}.$$

Efectuando ahora la operación $F_1 - \frac{b}{a-b}(F_2+\cdots+F_n) \to F_1$, se obtiene la matriz

$$\begin{pmatrix} a+(n-1)b & 0 & 0 & \cdots & 0 \\ 0 & a-b & 0 & \cdots & 0 \\ 0 & 0 & a-b & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & a-b \end{pmatrix}.$$

Conclusión:

$$r_c(A) = \begin{cases} n & \text{si } (a+(n-1)b)(a-b)^{n-1} \neq 0 \\ n-1 & \text{si } (a+(n-1)b) = 0 \neq a-b. \end{cases}$$

A.1.5 Lema.

1. Sean V, W espacios vectoriales sobre K y $B_1 = \{v_1, \ldots, v_n\}$ y $B_2 = \{w_1, \ldots, w_m\}$ bases respectivas para V y W. Si $A \in \mathrm{End}_K(V,W)$ con $A(v_j) = \sum_{i=1}^{n} a_{ij} w_i,\ j = 1, \ldots, n$, entonces

$$r(A) = r_c((a_{ij})).$$

2. Si $A \in \mathrm{Mat}(n \times m, K)$, $B \in \mathrm{Mat}(m, K)$ y $C \in \mathrm{Mat}(n, K)$ con B y C invertibles, entonces

$$r_c(CAB) = r_c(A).$$

3. Si $A \in \mathrm{Mat}(n \times m, K)$ y $B \in \mathrm{Mat}(n \times s, K)$, entonces

$$r_c(AB) \leq \text{mín}\{r_c(A), r_c(B)\}.$$

Demostración.

1. Definamos la función $B : W \longrightarrow K^m$ de la siguiente manera:

$$B\left(\sum_{i=1}^{m} k_i w_i\right) := \begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix},\ k_i \in K.$$

Es claro que B es un isomorfismo. De 3.3.14 (3) (ver [20] Capitulo 3) se tiene que $r(A) = r(BA) = \dim_K(\mathrm{Im}(BA))$. Note que $\mathrm{Im}(BA)$ es el generado de los vectores

$$BA(v_j) = B\left(\sum_{i=1}^{m} a_{ij} w_i\right) := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = c_j,\ j = 1, \ldots, n.$$

Por lo tanto,

$$r(A) = \dim_K \langle c_1, \ldots, c_n \rangle = r_c((a_{ij})).$$

2. Es una traducción de 3.3.14 (3) (ver [20]) en términos de matrices.

3. Corresponde a 3.3.14 (2) (ver [20]). □

A.1.6 Teorema.

1. Sean V, W espacios vectoriales sobre K y con dimensiones finitas y

$$A \in \mathrm{End}_K(V, W).$$

Entonces existen bases B_1 y B_2 de V y W respectivamente, tales que

$${}_{B_2}A_{B_1} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

donde $r = r(A)$.

2. Si $A \in \mathrm{Mat}(n \times m, \mathbb{F}_q)$, entonces existen matrices invertibles con $B \in \mathrm{Mat}(m, K)$ y $C \in \mathrm{Mat}(n, K)$ tales que

$$BAC = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

donde $r = r_c(A)$.

Demostración.

1. Sea $B = \{v_1, \ldots, v_n\}$ una base para V, de tal forma que $(v_{r+1}, \ldots, v_n)$ sea una base para el $\ker(A)$, con r adecuado. El teorema de isomorfía asegura que

$$\begin{aligned} \dim_K \operatorname{Im}(A) &= \dim_K V/\ker(A) \\ &= \dim_K V - \dim_K \ker(A) \\ &= n - \dim_K \ker(A) \\ &= r. \end{aligned}$$

   Probemos ahora que $(A(v_1, \ldots, A(v_r)))$ es una base para $\operatorname{Im}(A)$.
   Sea $y \in \operatorname{Im}(A)$, entonces $y = A(x)$ para algún $x \in V$, entonces

$$\begin{aligned} x = \textstyle\sum_{j=1}^{n} \alpha_j v_j \quad \Rightarrow \quad y &= \textstyle\sum_{j=1}^{n} \alpha_j A(v_j) \\ &= \textstyle\sum_{j=1}^{r} \alpha_j A(v_j), \end{aligned}$$

   entonces $\operatorname{Im}(A) = \langle \beta' \rangle$, veamos que son L.I
   Sea $\sum_{i=1}^{r} \alpha_i A(v_i) = 0$, entonces $A\left(\sum_{i=1}^{r} \alpha_i A v_i\right) = 0$, por lo tanto $\sum_{i=1}^{r} \alpha_i v_i \in \ker(A)$ así que $\sum_{i=1}^{r} \alpha_i v_i = \sum_{j=r+1}^{n} \beta_j v_j$, luego tenemos que $\sum_{j=r+1}^{n} \delta_i v_i = 0$, entonces $\alpha_i = 0$, para todo $i = 1, \ldots, r$ y $\delta_i = 0$, para todo $i = 1, \ldots, n$. Con lo anterior mostramos que $(A(v_1, \ldots, A(v_r)))$ es una base para $\operatorname{Im}(A)$

   Entonces si elegimos una base $(w_1, \ldots, w_m)$ de W, de tal manera que $w_j = A(v_j)$ para $j = 1, \ldots, r$. Entonces

$$A(v_j) = w_j, \text{ para } j = 1, \ldots, r$$

$$A(v_j) = 0, \text{ para } j = r+1, \ldots, n.$$

   Esto demuestra que

$${}_{B_2}A_{B_1} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

   con $r = r(A)$.

2. Se sigue de lo demostrado anteriormente y del teorema de cambio de base (ver [20]). □

A.1.7 Definición. Sea $A = (a_{ij}) \in \operatorname{Mat}(n \times m, K)$. Se define la transpuesta de A, notada por A^T como la matriz que se obtiene inter cambiando las filas por columnas en A. Es decir, $A^T = (a_{ji})$.

A.1.8 Teorema. Si $A \in \text{Mat}(n \times m, K)$, entonces

$$r_f(A) = r_c(A) \leq \text{mín}\{m, n\}.$$

Este teorema nos permite hablar del rango de una matriz y escribiremos para ello $r(A)$.

Demostración. Usando el teorema A.1.6 (2) se tiene que existen matrices invertibles $B \in \text{Mat}(m, K)$ y $C \in \text{Mat}(n, K)$ tales que

$$BAC = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

donde $r = r_c(A)$. Entonces haciendo uso del lema 4.1.21 (ver [20])

$$C^t A^t B^t = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Usando nuevamente el teorema A.1.6 (2) obtenemos que

$$r_c(A) = r = r_c(C^t A^t B^t) = r_f(A). \ \square$$

A.2. Propiedades del rango de una matriz.

Sea $X \in \text{Mat}(n \times m, \mathbb{F}_q)$. Por definición, el Rang $X = \dim\langle X \rangle$, sin embargo, hay muchas formas equivalentes y útiles de caracterizar el rango de X. Por ejemplo, que el Rang X es el natural más pequeño r para el que existen matrices $A \in \text{Mat}(n \times r, \mathbb{F}_q)$ y $B \in \text{Mat}(n \times m, \mathbb{F}_q)$ tal que $X = AB$, es decir,

$$\text{Rang } X = \text{mín}\{r \mid r, A \in \text{Mat}(n \times r, \mathbb{F}_q), B \in \text{Mat}(r \times m, \mathbb{F}_q), X = AB.\} \tag{A.3}$$

sabemos que, para cualquier $X, Y \in \text{Mat}(n \times m, \mathbb{F}_q)$, se tiene que el

$$\text{Rang } (X + Y) \leq \text{Rang } X + \text{Rang } Y$$

y que, para $X \in \text{Mat}(n \times m, \mathbb{F}_q)$ y $A \in \text{Mat}(N \times m, \mathbb{F}_q)$, tenemos

$$\text{Rang } (AX) \geq \text{Rang } A + \text{Rang } X - n. \tag{A.4}$$

Recordemos que si U y V son subespacios de un espacio vector fijo, entonces la suma

$$U + V = \{u + v : u \in U, v \in V\}$$

es el más pequeño subespacio que contiene tanto U y V. Recordemos también que

$$\dim(U + V) = \dim U + \dim V - \dim(U \cap V).$$

Vamos a hacer un uso extensivo del hecho de que

$$\left\langle \begin{pmatrix} X \\ Y \end{pmatrix} \right\rangle = \langle X \rangle + \langle Y \rangle \tag{A.5}$$

y por lo tanto

$$\begin{aligned} \text{Rang } \begin{pmatrix} X \\ Y \end{pmatrix} &= \dim(\langle X \rangle + \langle Y \rangle) \\ &= \text{Rang } X + \text{Rang } Y - \dim(\langle X \rangle \cap \langle Y \rangle). \end{aligned}$$

A.2.1 Lema. Sean $X, Y \in \text{Mat}(n \times m, \mathbb{F}_q)$. Entonces

$$\text{Rang } \begin{pmatrix} X \\ Y \end{pmatrix} = \text{Rang } (Y - X) + \text{mín}\{\text{Rang } X + \text{Rang } Y\}.$$

Demostración.

$$\text{Rang} \begin{pmatrix} X \\ Y \end{pmatrix} = \text{Rang } \begin{pmatrix} X \\ Y - X \end{pmatrix} \leq \text{Rang}(Y - X) + \text{Rang } X.$$

De igual forma se puede tomar

$$\text{Rang } \begin{pmatrix} X \\ Y \end{pmatrix} = \text{Rang } \begin{pmatrix} Y - X \\ Y \end{pmatrix} \leq \text{Rang } (Y - X) + \text{Rang } Y.$$

entonces

$$\text{Rang } \begin{pmatrix} X \\ Y \end{pmatrix} \leq \text{Rang } (Y - X) + \text{mín}\{\text{Rang } X, \text{Rang } Y\}.$$

A.2.2 Corolario Sean $X, Z \in \text{Mat}(n \times m, \mathbb{F}_q)$ y $Y = X + Z$. Entonces

$$d(\langle X \rangle, \langle Y \rangle) \leq 2 \text{ Rang } Z - |\text{Rang } X - \text{Rang } Y|.$$

Demostración. Recordemos que $\dim(\langle X\rangle + \langle Y\rangle) = \text{Rang } \begin{pmatrix} X \\ Y \end{pmatrix}$ y que

$$\dim(\langle X\rangle \cap \langle Y\rangle) = \text{Rang } X + \text{Rang } Y - \text{Rang } \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Ahora bien teniendo en cuenta la proposición A.2.1 obtenemos;

$$\begin{aligned} d(\langle X\rangle, \langle Y\rangle) &= \dim(\langle X\rangle + \langle Y\rangle) - \dim(\langle X\rangle \cap \langle Y\rangle) \\ &= 2 \text{ Rang } \begin{pmatrix} X \\ Y \end{pmatrix} - \text{Rang } X - \text{Rang } Y \\ &\leq 2 \text{ Rang} Z + 2\,\text{mín}\{\text{Rang } X, \text{Rang } Y\} - \text{Rang } X - \text{Rang } Y \\ &= 2 \text{ Rang } Z - |\text{Rang } X - \text{Rang } Y|. \quad \square \end{aligned}$$

A.2.3 Lema. Sea $X \in \text{Mat}(n \times m, \mathbb{F}_q)$, $Y \in \text{Mat}(N \times m, \mathbb{F}_q)$ tenemos

$$\min_{A \in \text{Mat}(N\times m, \mathbb{F}_q)} \text{Rang } (Y - AX) = \text{Rang } \begin{pmatrix} X \\ Y \end{pmatrix} - \text{Rang } X.$$

y para $X \in \text{Mat}(n \times m, \mathbb{F}_q)$ y $Y \in \text{Mat}(N \times m, \mathbb{F}_q$. Tenemos que

$$\min_{B \in \text{Mat}(m\times M, \mathbb{F}_q)} \text{Rang } (Y - XB) = \text{Rang } (XY) - \text{Rang } X.$$

Demostración. Para cualquier $A \in \text{Mat}(N \times m, \mathbb{F}_q)$ se tiene que

$$\begin{aligned} \text{Rang } \begin{pmatrix} X \\ Y \end{pmatrix} &= \text{Rang } \begin{pmatrix} X \\ Y - AX \end{pmatrix} \\ &\leq \text{Rang } X + \text{Rang } (Y - AX). \end{aligned}$$

Con lo cual se tiene la cota inferior para Rang $(Y - AX)$

$$\text{Rang } \begin{pmatrix} X \\ Y \end{pmatrix} - \text{Rang } X \leq \text{Rang } (Y - AX).$$

Demostramos ahora que esta cota inferior es alcanzable.
Sea $Z \in \text{Mat}(t \times m, \mathbb{F}_q)$ tal que

$$\langle Y\rangle = \langle X\rangle \cap \langle Y\rangle \oplus \langle Z\rangle$$

donde $t = \text{Rang } Y - \omega$ y $\omega = \dim\langle X\rangle \cap \langle Y\rangle$.
Sea $B \in \text{Mat}(\omega \times n, \mathbb{F}_q)$ tal que

$$\langle BX\rangle = \langle X\rangle \cap \langle Y\rangle$$

podemos escribir

$$Y = T \begin{pmatrix} BX \\ Z \end{pmatrix}$$

para alguna $T \in \mathrm{Mat}(N \times \omega + t, \mathbb{F}_q)$ con rango completo.

Sea ahora

$$A = T \begin{pmatrix} BX \\ U \end{pmatrix} \in \mathrm{Mat}(N \times m, \mathbb{F}_q).$$

Entonces

$$\begin{aligned} \text{Rang } (Y - AX) &= \text{Rang } \left(\begin{pmatrix} BX \\ Z \end{pmatrix} - T \begin{pmatrix} BX \\ 0 \end{pmatrix} \right) \\ &= T \begin{pmatrix} 0 \\ Z \end{pmatrix} \\ &= \text{Rang } Z \\ &= \text{Rang } Y + \dim(\langle X \rangle \cap \langle Y \rangle) \\ &= \text{Rang } \begin{pmatrix} X \\ Y \end{pmatrix} - \text{Rang } X. \end{aligned}$$

Esto demuestra la primera afirmación. La segunda afirmación es solo la versión transpuesta de la primera. □

El objetivo de la siguiente sección es diseñar un algoritmo que genere todos polinomios linealizados de un mismo grado lineal.

Capítulo B

Polinomios linealizados sobre cuerpos binarios.

Los polinomios linealizados fueron previamente definidos, por tal razón en lo que se centrara este capitulo es en el diseño de un algoritmo para generarlos y en la aplicación de dicho algoritmo haciendo uso del programa SAGE.

Diseño de un algoritmo que genere polinomios linealizados sobre cuerpos binarios.

Para la creación de este tipo de polinomios con coeficientes sobre cuerpos binarios haremos uso de la definición

$$f(x) = \sum_{i=0}^{t} f_i x^{q^i}.$$

Procedimiento: Creación de polinomios linealizados $f(x) = \sum_{i=0}^{t} f_i x^{q^i}$, con $f_i \in \mathbb{F}_{q^m}$ y $q = 2$.

entrada: m, y t.

salida: Todos los polinomios linealizados $f(x)$ de grade q^t con coeficientes en $\mathbb{F}_{2^m}$.

Inicio

Haga las siguientes asignaciones $m := n$, $d := t$ y $f_i := a_i$
$L :=$ El conjunto de elementos de $\mathbb{F}_{2^n}$.
$N :=$ El conjunto de combinaciones $d+1$ de L.

si $d < 0$ **entonces**:
imprima Error el valor 2^d es el grado de un polinomio, por lo tanto debe ser positivo.

si no:
s=0
para i desde 0 hasta $|N|$
elementos nulos:=verdadero

para j desde 0 hasta d:

si $a_{[i][j]} \neq 0$, $a_{[i][j]} \in N$ **entonces**:
elementos cero:=falso

si elementos nulos:=verdadero o $a_{[i][j]} \neq 0$, donde $a_{[i][j]} \in N$:
para k desde 0 hasta $d+1$:

$s := \sum_{i=0}^{d} a_{[i][j]} x^{2^i}$
si $s \neq 0$:
imprima s
fin si
fin para
fin si
fin para
fin para
fin

Suponga entonces que usted desea generar polinomios de grado 1, tendiendo en cuenta que los polinomios linealizados se definen como:

$$f(x) = \sum_{i=0}^{d} a_i x^{q^i},$$

con $a_i \in \mathbb{F}_{q^n}$, $0 \leq i \leq d$.

Entonces los polinomios a generar de grado uno serán:

$$x$$

$$\alpha x$$

$$(\alpha + 1)x.$$

En la siguiente sección se mostrara los cálculos efectuados con SAGE, a medida que se explique la implementación del algoritmo en este programa se generaran los polinomios de grado cuatro con coeficientes en $\mathbb{F}_4$ con el fin de ilustrar la aplicación del algoritmo con mas detalle.

Cálculos realizados con sage

El programa SAGE es un sistema algebraico computacional (en inglés CAS) escrito en Python y es una versión modificada de Pyrex (llamada inicialmente SageX y posteriormente Cython).

Suponga que se desea obtener todos los polinomios linealizados de grado 4 con coeficientes en en el cuerpo finito $\mathbb{F}_4$, en tal caso $n = 2$ y $d = 2$ los cuales son los valores de entrada. Recuerde que los elementos de $\mathbb{F}_4$ son:

$$\mathbb{F}_4 := \{0, 1, \alpha, \alpha + 1\}$$

Para cálculos sobre

$$\mathbb{F}_q^n,$$

se necesita fijar una extensión algebraica de grado n sobre $\mathbb{F}_q$. Para hacer esto en SAGE se usa la instrucción `GF`(q, a), entonces la instrucción antes mencionada debe escribirse `GF`$(2^2, a)$, escrito de esta manera se generan los elementos del cuerpo finito $\mathbb{F}_4$ la cual escribe sus elementos usando la variable a, este conjunto lo guardamos como una lista digitando `L = G.list()` y si se usa la instrucción `print L` es posible visualizar los elementos de $\mathbb{F}_4$ en SAGE, los cuales se mostraran así:

```
[0, a, a + 1, 1]
```

Dado que $d = 2$ las combinaciones a crear serán $d + 1$, es decir de tres elementos, con tal fin se usa la instrucción `Tuple`, dicha instrucción crea combinaciones en una lista, la cual nos permite permite obtener lo siguiente,

```
[[0, 0, 0], [a, 0, 0], [a + 1, 0, 0], [1, 0, 0], [0, a, 0], [a, a, 0],
[a + 1, a, 0], [1, a, 0], [0, a + 1, 0], [a, a + 1, 0], [a + 1, a + 1,
0], [1, a + 1, 0], [0, 1, 0], [a, 1, 0], [a + 1, 1, 0], [1, 1, 0], [0,
0, a], [a, 0, a], [a + 1, 0, a], [1, 0, a], [0, a, a], [a, a, a], [a +
1, a, a], [1, a, a], [0, a + 1, a], [a, a + 1, a], [a + 1, a + 1, a],
[1, a + 1, a], [0, 1, a], [a, 1, a], [a + 1, 1, a], [1, 1, a], [0, 0, a
+ 1], [a, 0, a + 1], [a + 1, 0, a + 1], [1, 0, a + 1], [0, a, a + 1],
[a, a, a + 1], [a + 1, a, a + 1], [1, a, a + 1], [0, a + 1, a + 1], [a,
a + 1, a + 1], [a + 1, a + 1, a + 1], [1, a + 1, a + 1], [0, 1, a + 1],
[a, 1, a + 1], [a + 1, 1, a + 1], [1, 1, a + 1], [0, 0, 1], [a, 0, 1],
[a + 1, 0, 1], [1, 0, 1], [0, a, 1], [a, a, 1], [a + 1, a, 1], [1, a,
1], [0, a + 1, 1], [a, a + 1, 1], [a + 1, a + 1, 1], [1, a + 1, 1], [0,
1, 1], [a, 1, 1], [a + 1, 1, 1], [1, 1, 1]]
```

nuevamente la anterior lista se hará visible únicamente si se usa la instrucción `print N`, donde `N = Tuples(L,m+1).list()`.

El siguiente paso a realizar es recorrer la lista anterior, multiplicando cada coeficiente encontrado por x^{p^i}, donde $p = 2$ para $i = 0, 1, 2$. Estos elementos serán recorridos entonces por las partes variables del polinomio de grado 4 (ejemplo en consideración).

$$x, x^2, x^4$$

esto teniendo presente la definición de polinomios linealizados, donde $d = 2$ y $q = 2$.

Para ser mas exactos el procedimiento exige que cada elemento de $a_{[i][j]}$ de la lista N sea multiplicado por x, x^2 y x^4, este recorrido se lograra usando una recurrencia con las alternativas lógicas `if` y `else`, además para efectos del funcionamiento del programa es indispensable formar bucles anidados con la instrucción `for` con el fin de que cada elemento de $a_{[i][j]}$ en la lista `N` sea tomado en consideración, con lo que se logra el resultado deseado.

A continuación se muestra toda la implementación del algoritmo usando SAGE.

```
# n es la potencia de q=2 del cuerpo finito F_q
# d es el grado del polinomio que usted desea generar (x)^(q^d)

def pol_linealizados(n,d):
    m=d
    var('x')
    q=2^n
    G= GF(q,'a')
    L=G.list()
    #print L
    N=Tuples(L,m+1).list()
    #print N
    for i in range(0,len(N)):
         s=0
         elementos_nulos=true
         for j in range(m):
             if N[i][j]!=0:
                elementos_nulos=false

         if N[i][m]!=0 or elementos_nulos==true:
               for k in range(m+1):
                s+=N[i][k]*(x^(2^k))
               if s!=0:
                print s
```

Note que en esta codificación `print L` aparece precedido con el símbolo numeral escrito de está manera se le esta indicando a SAGE que no nos muestre esta lista o que no lea esta instrucción ya que lo único que deseamos que nos muestre son los polinomios.

Para finalizar se debe correr el programa y de esta manera será posible visualizar todos los polinomios de grado cuatro que se han creado con el programa ya mostrado, con tal fin lo que se debe hacer es dígitar el nombre que se le ha dado al programa en una nueva celda de trabajo en la misma `Notebook`. El nombre que se le asigno al programa fue `pol_ linealizados`$(2, 2)$, al digitarlo se debe indicar que $n = 2$ y $d = 2$, entonces SAGE mostrara los polinomios al dar clic en la opción `evaluate` arrojando el siguiente resultado:

```
a*x^4 + a*x
a*x^4 + (a + 1)*x
a*x^4 + x
a*x^4 + a*x^2
a*x^4 + a*x^2 + a*x
a*x^4 + a*x^2 + (a + 1)*x
a*x^4 + a*x^2 + x
a*x^4 + (a + 1)*x^2
a*x^4 + (a + 1)*x^2 + a*x
a*x^4 + (a + 1)*x^2 + (a + 1)*x
a*x^4 + (a + 1)*x^2 + x
a*x^4 + x^2
a*x^4 + x^2 + a*x
a*x^4 + x^2 + (a + 1)*x
a*x^4 + x^2 + x
(a + 1)*x^4
(a + 1)*x^4 + a*x
(a + 1)*x^4 + (a + 1)*x
(a + 1)*x^4 + x
(a + 1)*x^4 + a*x^2
(a + 1)*x^4 + a*x^2 + a*x
(a + 1)*x^4 + a*x^2 + (a + 1)*x
(a + 1)*x^4 + a*x^2 + x
(a + 1)*x^4 + (a + 1)*x^2
(a + 1)*x^4 + (a + 1)*x^2 + a*x
(a + 1)*x^4 + (a + 1)*x^2 + (a + 1)*x
(a + 1)*x^4 + (a + 1)*x^2 + x
(a + 1)*x^4 + x^2
(a + 1)*x^4 + x^2 + a*x
(a + 1)*x^4 + x^2 + (a + 1)*x
(a + 1)*x^4 + x^2 + x
x^4
x^4 + a*x
x^4 + (a + 1)*x
x^4 + x
x^4 + a*x^2
x^4 + a*x^2 + a*x
x^4 + a*x^2 + (a + 1)*x
x^4 + a*x^2 + x
x^4 + (a + 1)*x^2
x^4 + (a + 1)*x^2 + a*x
x^4 + (a + 1)*x^2 + (a + 1)*x
x^4 + (a + 1)*x^2 + x
x^4 + x^2
x^4 + x^2 + a*x
x^4 + x^2 + (a + 1)*x
x^4 + x^2 + x
```

Figura B.1: Todos los polinomios de grado cuatro con coeficientes en $\mathbb{F}_4$

Lista de símbolos

$G, H, \ldots$	conjuntos, grupos.
$\|G\|$	cardinalidad, número de elementos del grupo G.
$C^{\perp}$	el código dual de C.
C^{T}	el código transpuesto del código C.
$B_t(x)$	esfera con centro en x y radio t.
$K[x]$	conjunto de todos los polinomios con coeficientes en K.
$\hat{C}$	extensión del código C.
$\breve{C}(i)$	reducción del código C en la i-ésima coordenada.
$\dot{C}(i)$	perforación del código C en la i-ésima coordenada.
$d(u, v)$	distancia, número de coordenadas en que u y v difieren.
$d(C)$	distancia mínima del código C.
$\omega t(x)$	peso, número de coordenadas no nulas de x.
$\mathbb{P}(W)$	espacio proyectivo, conjunto de los subespacios de W.
$d_R(X, Y)$	distancia del rango entre las matrices X y Y.
$d_R(X, Y)$	distancia del rango de las matrices X y Y.
$I(X)$	levantamientos de la matriz X.
$I(C)$	levantamientos de un código matricial C.
$\alpha(C)$	suboptimalidad de un código C.
$D_R(C)$	distancia mínima de un código C con métrica del rango.

Bibliografía

[1] Rudolf Ahlswede, Ning Cai, Shuo-Yen Robert Li and Raymond W. Yeung *Network Information Flow*, IEEE Transactions on Information Theory, Vol. 46, No. 4, (2000).

[2] J.Adámek. *theory and applications of error-correcting codes with and introduction to cryptography adn information theory,*A wiley intersciense publication,137-147, (1991).

[3] R. Koetter and F. R. Kschischang. *Coding for Errors and Erasures in Random Network Coding,* IEEE Transactions on Information Theory, Volumen 54, (2008).

[4] R. Koetter, D. Silva and F. R. Kschischang. A *Rank-Metric Approach to Error Control in Random Network Coding,* IEEE Transactions on Information Theory, Volumen 54, (2008).

[5] J.H.VanLint and R.M.Wilson. *A Course in combinatorics, 2nd ed. Cambridge, U.K,*Cambridge Univ. Press, 2001.

[6] I. Gutiérrez. *Notas de clase,* 15-66, 71-79, (2011).

[7] Y. Raymond. *Information theory and Network coding,* SPIN spinger's internal poject number, 411-418, (2008).

[8] E.R, Berlekamp. *The tecnology of error-correcting codes,* IEEE, Volumen 68,564-593, (Mayo 1980).

[9] A.E. Brouwer, A.M. Cohen, and A. Neumaier. *Distance-Regular Graphs,* New York: Springer-verlag,(1989).

[10] E.M. Gabidulin. *A fast matrix decoding algorithm for rank-error-correcting codes,* Lecture Notes in Computer Science, Volumen 573, 126-133, (1992).

[11] E.M. Gabidulin. *Theory of codes with maximum rank distance, Probl. Inform. Transm,* volumen 21, no 1,1-12, (1985).

[12] D. Silva and F. R. Kschischang. *On metrics for error correction in network coding,* (2008). P.

[13] A. Chou, Y. Wu, and K. Jain. *Practical network coding, in Proc. Allerton Conf. on Comm., Control, and Computing,* Monticello, IL, Oct. (2003).

[14] T. Ho, M. M´edard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong. *A random linear network coding approach to multicast,* IEEE Trans. Inf. Theory, vol. 52, no. 10, 4413-4430,Oct. (2006).

[15] R. Lidl and H. Niederreiter. *Finite Fields.* Reading, MA: Addison-Wesley, (1983).

[16] F. MacWilliams and N. Sloane. *The Theory of Error-Correcting Codes,* Amsterdam: North-Holland, (1977).

[17] G. Richter and S. Plass. *Error and erasure decoding of rank-codes with a modified Berlekamp-Massey algorithm,* Germany, (2004).

[18] V. Skachek and R. M. Roth. *Probabilistic algorithm for finding roots of linearized polynomials, Designs, Codes and Cryptography* vol. 46, no. 1,17-23, (2008) tambien en *Comput. Sci. Dept, Technion, Tech. Rep. CS-2004-08,* Jun. (2004).

[19] E. R. Berlekamp. *Algebraic Coding Theory,* New York, McGraw-Hill, (1968).

[20] I. Gutiérrez y J. Robinson. *Algebra lineal, Ediciones uninorte,* (2012).