Índice general

1 Criptosistemas simétricos o de clave privada **1**
1.1 Criptografía 1
1.2 Criptoanálisis 6
1.3 Entropía 8

Capítulo 1

Criptosistemas simétricos o de clave privada

1.1. Criptografía

La *criptografía* es la ciencia de representar información de forma opaca para que sólo los agentes autorizados (personas o dispositivos diversos) sean capaces de desvelar el mensaje oculto. El proceso de ocultar la información se llama *cifrado*, pero a menudo también se llama *encriptado* por influencia del inglés. El proceso de desvelarla se llama *descifrado* o *desencriptado*. El concepto de criptosistema modela los procesos de cifrado y descifrado.

Un *criptosistema simétrico*, también llamado *de clave privada*, está formado por un conjunto K, cuyos elementos llamamos *claves* o *llaves*, y una regla que asocia dos aplicaciones a cada clave $k \in K$:

$$c_k : M_k \longrightarrow C_k, \quad d_k : C_k \longrightarrow M_k$$

de forma que

$$d_k(c_k(x)) = x, \text{ para todo } x \in M_k$$

En la práctica el protocolo criptográfico también incluye un algoritmo generador de claves, es decir, uno que tiene como salida un elemento de K, pero nosotros no vamos a tener en cuenta esta parte del criptosistema.

Extendemos las aplicaciones c_k a $M_k^\infty = \cup_{n\geq 1} M_k^n$ y d_k a $C_k^\infty = \cup_{n\geq 1} C_k^n$ poniendo

$$c_k(x_1 \dots x_n) = c_k(x_1) \cdots c_k(x_n), \quad x_1, \dots, x_n \in M_k$$
$$d_k(y_1 \dots y_n) = d_k(y_1) \cdots d_k(y_n), \quad y_1, \dots, y_n \in C_k$$

Obsérvese que representamos los elementos de M^n como concatenación de elementos de M.

Utilizaremos la siguiente terminlogía para una clave $k \in K$:

- Elementos de M_k^∞: *Mensajes en claro.*
- Elementos de C_k^∞: *Mensajes en cifrados o encriptados.*
- Elementos de c_k: *Función de cifrado o función de encriptado.*
- Elementos de d_k: *Función de descifrado o función de desencriptado.*

Un mensaje en claro o cifrado diremos que es *básico* si tiene longitud 1. Sin embargo, en muchas situaciones diremos mensajes para referirnos a mensajes básicos, bien en claro o cifrados.

Frecuentemente M_k es el mismo conjunto para todas las claves y lo mismo ocurre con los C_k. En este caso ponemos $M = M_k$ y $C = C_k$.

Ejemplo 1.1. *Criptosistemas.*
Sea A un conjunto finito y denotemos por S_A el conjunto de permutaciones de los elementos de A.

(1) *Sustitución.* Tomamos como conjunto de claves $K = S_A$, como conjuntos de mensajes básicos $M = C = A$ y como funciones de cifrado y descifrado:

$$c_\sigma(x) = \sigma(x) \quad \text{y} \quad d_\sigma(y) = \sigma^{-1}(y)$$

(2) *Reordenamiento.* Ponemos $K = \cup_{n\geq 2} S_n$, donde $S_n = S_{\{1,\dots,n\}}$. Si la clave σ está en S_n entonces ponemos $M_\sigma = C_\sigma = A^n$ y ciframos reordenando las posiciones de los símbolos de los mensajes en claro. Más precisamente:

$$c_\sigma(x_1 \cdots x_n) = x_{\sigma(1)} \cdots x_{\sigma(n)} \quad \text{y} \quad d_\sigma(y_1 \cdots y_n) = x_{\sigma^{-1}(1)} \cdots x_{\sigma^{-1}(n)}$$

En la práctica casi todos los protocolos criptográficos son combinaciones de los del ejemplo 1.1. Sin embargo, en estos criptosistemas no se dice nada sobre las permutaciones elegidas. La naturaleza de estas permutaciones es lo que hace un criptosistema bueno o malo. La bondad de un criptosistema depende de que satisfaga las siguientes condiciones:

- *Rapidez de los cálculos.* Es importante disponer de un algoritmo eficiente (polinomial, con exponente pequeño) para calcular $c_k(x)$ y $d_k(y)$.
- *Seguridad.* Debe ser difícil descubrir un valor concreto de x a partir del valor de $c_k(x)$ sin conocer k.

Las nociones de "algoritmo eficiente", "difícil de calcular" y "tiempo razonable" son ambiguas. Más adelante daremos conceptos más precisos. De momento, veamos ejemplos concretos.

Criptosistema de César

Vamos a identificar las 27 letras mayúsculas del alfabeto español asignando a cada una de ellas un número del 0 al 26, de forma que estamos identificando los elementos de A con los de $\mathbb{Z}_{27}$ usando la siguiente tabla

A	B	C	D	E	F	G	H	I	J	K	L	M	N
0	1	2	3	4	5	6	7	8	0	10	11	12	13
$\tilde{N}$	O	P	Q	R	S	T	U	V	W	X	Y	Z	
14	15	16	17	18	19	20	21	22	23	24	25	26	

(1.1)

Ponemos $K = M = C = \mathbb{Z}_{27}$ y

$$c_k(x) = (x + k \mod 27)$$
$$d_k(x) = (x - k \mod 27)$$

Por ejemplo, supongamos que la clave es la letra W, que corresponde con el número 23 y que queremos cifrar el mensaje "ésta es la primera vez que vamos a cifrar un mensaje". En primer lugar obsérvese que nuestro alfabeto sólo contiene letras mayúsculas sin acentuar, con lo que el mensaje a cifrar es:

ESTAESLAPRIMERAVEZQUEVAMOSACIFRARUNMENSAJE

Ahora tenemos que convertir el mensaje en una lista de números de acuerdo a la tabla anterior:

$$4, 19, 20, 0, 4, 19, 11, 0, 16, 18, 8, 12, 4, 18, 0, 22, 4, 26, 17, 21,$$
$$4, 22, 0, 12, 15, 19, 0, 2, 8, 5, 18, 0, 18, 21, 13, 12, 4, 13, 19, 0, 9, 4$$

Después tenemos que sumar a cada uno de estos números la clave, es decir, sumamos 23 módulo 27, o lo que es lo mismo restamos 4 módulo 27.

$$0, 15, 16, 23, 0, 15, 7, 23, 12, 14, 4, 8, 0, 14, 23, 18, 0, 22, 13, 17,$$
$$0, 18, 23, 8, 11, 15, 23, 25, 4, 1, 14, 23, 14, 17, 9, 8, 0, 9, 15, 23, 5, 0$$

Finalmente tenemos que sustituir los números por las letras correspondientes y obtendremos el mensaje cifrado.

AOPWAOHWMÑEIAÑWRAVNQARWILOWYEBÑWÑQJIAJOWFA

Criptosistema de Vigenère

Es similar al de César con el alfabeto español identificado con $\mathbb{Z}_{27}$, pero $K = \mathbb{Z}_{27}^{\infty}$ y si $k \in \mathbb{Z}_{27}^{d}$ entonces $M_k = C_k = \mathbb{Z}_{27}^{d}$ y

$$c_{k_1 \cdots k_d}(x_1 \cdots x_d) = [x_1 + k_1]_{27} \cdots [x_d + x_k]_{27}$$
$$d_{k_1 \cdots k_d}(y_1 \cdots y_d) = [y_1 - k_1]_{27} \cdots [y_d - y_k]_{27}$$

Es decir se divide el mensaje en bloques de longitud d y se trabaja como en el Criptosistema de César usando la clave k_i para los símbolos en posiciones con $j \equiv i \mod d$.

En realidad el criptosistema de César y Vigenére se pueden considerar como casos particulares de otros más generales. Simplemente hemos restringido la presentación a 27 símbolos para hacer más simple la explicación. Sin embargo el conjunto de símbolos puede ser cualquiera (letras mayúsculas y minúsculas, comas, espacios, etc), y podemos siempre identificar el conjunto de símbolos con $\mathbb{Z}_n$ donde n es el cardinal del conjunto de símbolos. De esta forma podemos obtener versiones del Criptosistema de César y del de Vigenère con un conjunto mayor de símbolos. Por otro lado, se puede cambiar $\mathbb{Z}_n$ por cualquier otro grupo sin que cambie de forma esencial la naturaleza de estos dos criptosistemas. Nos referiremos a este grupo como plataforma. Por ejemplo, en lugar de identificar los símbolos del alfabeto con los elementos de $\mathbb{Z}_n$ los podemos identificar con los elementos de un grupo finito y utilizar la operación del grupo para hacer las operaciones. No es esencial que la aplicación que asocia a un símbolo un elmenento del grupo plataforma sea sobreyectiva, aunque si que es necesario que sea inyectiva.

Por otro lado, los símbolos pueden no ser elementos de un alfabeto natural sino combinaciones de varios símbolos. Es decir, podemos elegir una cierta longitud, pongamos n, e identificar listas de n símbolos en nuestro alfabeto, con el grupo aditivo $\mathbb{Z}_{b^n}$, donde b es el cardinal del alfabeto. Por ejemplo, supongamos que utilizamos el mismo alfabeto de antes, es decir, las letras mayúsculas del alfabeto en español y que las queremos agrupar en bloques de longitud cuatro. Entonces podemos identificar estas sucesiones con los elementos de $\mathbb{Z}_{27^4}$ mediante la siguiente regla:

$$AAAA = 0, AAAB = 1, \ldots, AAAZ = 26, AABA = 27, AABB = 28, \ldots, ZZZZ = 27^4 - 1$$

Obsérvese que el bloque $X_1X_2X_3X_4$ se identificará con

$$x_1 \cdot 27^3 + x_2 \cdot 27^2 + x_3 \cdot 27 + x_4 \cdot 27^0$$

donde x_i es el número identificado con la letra X_i. Por ejemplo, "MESA" se identificaría con

$$12 \cdot 27^3 + 4 \cdot 27^2 + 19 \cdot 27 + 0 \cdot 27^0 = 239625$$

De esta forma $M = C = \mathbb{Z}_{27^4}$ y podríamos utilizar cualquiera de los criptosistemas anteriores de forma similar.

En resumen, podemos establecer diferentes formas de asociar los mensajes básicos con los elementos de un grupo G. Todo esto no tiene ninguna naturaleza criptográfica, es simplemente la codificación de la información, de forma que podemos identificar el grupo plataforma G con los tres conjuntos fundamentales del protocolo criptográfico: $K \equiv M \equiv C \equiv G$. Lo que tiene la naturaleza criptográfica son las funciones de cifrado y descifrado, que en este caso toma la forma,

$$c_k(x) = kx, \quad \text{y} \quad d_k(y) = k^{-1}y$$

Otra cosa importante, es que la naturaleza del grupo plataforma no es indiferente, pues para que podamos considerar el criptosistema como tal, necesitamos disponer de algoritmos eficientes para calcular las funciones de cifrado y descifrado y queremos que, sin el conocimiento de la clave sea difícil descifrar. Así por ejemplo $(\mathbb{Z}_n, +)$, $(\mathbb{Z}_n^*, \cdot)$ y los grupos aditivo y multiplicativo de un cuerpo finito o el grupo aditivo de un espacio vectorial sobre un cuerpo finito cumplen la primera condición pues la aritmética en estos grupos es sencilla.

Veamos un último ejemplo de criptosistema sencillo.

Criptosistemas afines

Sea n un entero positivo y pongamos $M = C = \mathbb{Z}_n$ y $K = \mathbb{Z}_n^* \times \mathbb{Z}_n$, donde $\mathbb{Z}_n^*$ representa el conjunto de elementos invertibles de $\mathbb{Z}_n$, es decir, los números naturales menores que n y coprimos con n. Las funciones de cifrado y descifrado son

$$c_{(a,b)}(x) = (ax + b \mod n)$$
$$d_{(a,b)}(x) = (a^{-1}(x - b) \mod n)$$

donde a^{-1} representa el inverso de a en $\mathbb{Z}_n$.

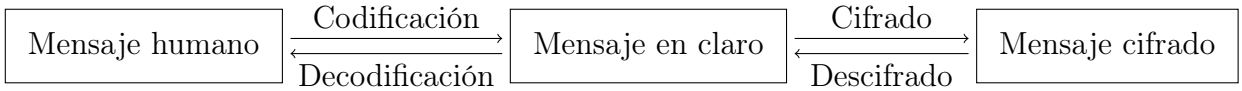
Más generalmente, si R es un anillo y R^* denota el grupo de las unidades de R, podemos poner $K = R^* \times R$, $M = C = R$ y utilizar las mismas funciones de cifrado y descifrado: $c_{(a,b)}(x) = ax + b$ y $d_{(a,b)}(x) = a^{-1}(x - b)$, con operaciones en el anillo plataforma R.

Otra alternativa es tomar $K = \mathrm{GL}_d(R) \times R^d$, donde $\mathrm{GL}_d(R)$ denota el conjunto de las matrices invertibles $d \times d$ con entradas en R. Entonces, usando multiplicación matricial el criptosistema afín viene dado por

$$c_{(A,b)} = Ax + b$$
$$d_{(A,b)} = A^{-1}(x - b)$$

Más generalmente, se puede fijar como conjunto de mensajes básicos, un $R-$módulo $M = C$ y elegir como conjunto de claves $K = \mathrm{Aut}(M) \times M$, donde $\mathrm{Aut}(M)$ es el grupo de los automorfismos de M, y como funciones de cifrado y descifrado $c_{(f,b)}(x) = f(x) + b$ y $d_{(f,b)}(y) = f^{-1}(y - b)$. En este caso la plataforma está formada por un anillo y un módulo y, de nuevo, las naturalezas algebraicas de las plataformas son importantes.

Cerramos esta sección con un esquema sencillo del proceso criptográfico:



En el proceso de codificación, transformamos un texto en símbolos que puedan ser procesados por algoritmos. Por ejemplo, en el Criptosistema de César los mensajes humanos son las letras de un alfabeto que las codificamos convirtiéndolas en números.

	Codificación		Clave = 3	
A	$\longmapsto$	0	$\longmapsto$	3
B	$\longmapsto$	1	$\longmapsto$	4
$\vdots$	$\longmapsto$	$\vdots$	$\longmapsto$	$\vdots$
W	$\longmapsto$	23	$\longmapsto$	26
X	$\longmapsto$	24	$\longmapsto$	0
Y	$\longmapsto$	25	$\longmapsto$	1
Z	$\longmapsto$	26	$\longmapsto$	2

(1.2)

También podríamos hacer bloques de una longitud determinada, por ejemplo 4, y convertirlas en números considerando los símbolos del alfabeto original como los guarismos de un sistema de numeración, con lo que las palabras de una determinada longitud se interpretan como números de esa longitud en base b, donde b es el número de símbolos del alfabeto.

	Codificación		Clave = 300	
$AAAA$	$\longmapsto$	0	$\longmapsto$	300
$AAAB$	$\longmapsto$	1	$\longmapsto$	301
$\vdots$	$\longmapsto$	$\vdots$	$\longmapsto$	$\vdots$
$AAAZ$	$\longmapsto$	26	$\longmapsto$	326
$AABA$	$\longmapsto$	27	$\longmapsto$	327
$\vdots$	$\longmapsto$	$\vdots$	$\longmapsto$	$\vdots$
$BEYC$	$\longmapsto$	23276	$\longmapsto$	23576
$\vdots$	$\longmapsto$	$\vdots$	$\longmapsto$	$\vdots$
$ZZZZ$	$\longmapsto$	$27^4 - 1 = 531440$	$\longmapsto$	299

Con este ejemplo el proceso de codificación de BEYC tiene dos pasos: primero BEYC se convierte en la lista de números $(1, 4, 25, 2)$ y después se convierte en este número

$$1 \cdot 27^3 + 4 \cdot 27^2 + 25 \cdot 27^1 + 2 \cdot 27^0 = 23276$$

Después ciframos sumando la clave. En el proceso de descifrado deberíamos aplicar el proceso inverso primero descifrando y después convirtiendo el mensaje en claro en un mensaje humano. Por ejemplo, supongamos que el mensaje cifrado es 25936. Desciframos restando 300, con lo que obtendríamos 25636. Para decodificarlo tenemos que escribir este número en base 27 dividiendo sucesivamente por 27 y quedándonos con los restos:

$$\begin{aligned} 25636 &= 949 \cdot 27 + 13 \\ 949 &= 35 \cdot 27 + 4 \\ 35 &= 1 \cdot 27 + 8 \end{aligned}$$

Luego 25636 se corresponde con la lista $(1, 8, 4, 13)$ que, de acuerdo con la tabla 1.1 corresponde con la palabra BIEN.

A nosotors nos importa poco el proceso de codificación y nos interesaremos más en los procesos de cifrado y descifrado. Sin embargo, es importante tener en cuenta que en el proceso de codificación se debe elegir una buena plataforma para hacer los cálculos de encriptado. En los ejemplos

más sencillos la plataforma era $\mathbb{Z}_n$ para un número n; en los más complicados era otra estructura algebraica más abstracta. Pero lo importante es tener métodos de cálculo eficientes. Por ejemplo en $\mathbb{Z}_n$ las operaciones se reducen a operaciones aritméticas sencillas: sumas, multiplicaciones y divisiones. No hay ningún problema en trabajar con cuerpos finitos ya que veremos métodos de cálculo rápidos con dichos cuerpos. Tampoco habría grandes problemas en utilizar como plataforma anillos de matrices de cuerpos finitos o anillos de polinomios de dichos cuerpos, ya que las operaciones con cuerpos finitos son fácilmente transferibles a estos otros anillos. Se podrían también utilizar anillos de naturaleza discreta como el de los números enteros o el cuerpo de los números racionales, así como extensiones finitas de este. Sin embargo, no se pueden utilizar plataformas en las que no se pueda describir los elementos de una forma finita. Por ejemplo, el cuerpo de los números reales no es apropiado como plataforma.

1.2. Criptoanálisis

Criptoanálisis es la ciencia (¡o arte!) de desvelar un mensaje oculto sin necesidad de la clave con la que fue encriptado. La combinación de criptografía y criptoanálisis se llama *criptología*.

La posibilidad de éxito del criptoanalista depende de la cantidad de información que posea. Por ejemplo, imaginemos las siguientes situaciones:

(1) El criptoanalista cuenta con un mensaje cifrado de longitud suficentemente grande.

(2) El criptoanalista posee uno o varios mensajes sin cifrar con su correspondiente mensaje cifrado.

(3) El criptoanalista puede generar tanta parejas (m, c) de mensajes en claro como desee junto con su correspondiente mensaje cifrado.

Se supone que el criptoanalista conoce las funciones c y d de cifrado y descifrado y, o bien desea encontrar la clave utilizada para poder descifrar los mensajes, o bien simplemente quiere descifrar un mensaje concreto. El problema de descifrar los mensajes enviados con una clave se llama *Problema de Ruptura de un Criptosistema*.

Veremos como realizar un criptoanálisis del Criptosistema de César. Supongamos que estamos en la situación de menor información posible, es decir, tenemos un mensaje cifrado con el Criptosistema de César (en la versión que los mensajes básicos son las 27 letras del alfabeto español):

ÑWEWUEPKBNÑUKVKWMQKNÑMEIYWYVLBÑWYAERÑBYKMYBNKBVÑ

Suponemos que sabemos que el mensaje en claro es una frase en español. Como las letras más frecuentes son las vocales, las letras más frecuentes en el mensaje cifrado deben tener una gran intersección con las correspondientes a las vocales. Las dos letras más frecuentes en el mensaje anterior son la K y la $\tilde{\text{N}}$, cada una de las cuales aparece 6 veces. Probablemente una de éstas es el mensaje básico cifrado de la A. Probamos a descifrar con estas letras y descubrimos que el mensaje obtenido descifrando con la K es

ENUNLUGARDELAMANCHADECUYONOMBRENOQUIEROACORDARME

Claramente es difícil que la clave no sea K.

Por tanto, aunque el Criptosistema de César satisface la primera condición para ser un buen criptosistema (rapidez de cifrado y descifrado) no cumple la condición de seguridad. La razón básica de por qué el Criptosistema de César no es seguro es que el número de claves es muy pequeño. Hemos hecho un análisis de frecuencia de las letras para evitar probar con las 27 claves pero con

la ayuda de ordenadores esto sería inmediato con lo que se rompería muy fácilmente. Un ataque consistente en probar todas las claves es lo que se llama *ataque por fuerza bruta* mientras que un análisis de frecuencias de los símbolos utilizados se suele llamar *ataque por análisis de frecuencias*.

El Criptosistema de Vigenère es sólo un poco más seguro. Mejor dicho su seguridad dependerá de la longitud de la palabra clave. Aunque no se conozca dicha longitud si no es muy grande, un análisis de frecuencia de los caracteres que aparecen en el mensaje a intervalos regulares rompería el criptosistema de Vigenère. Esto lo veremos más adelante.

Analicemos ahora el criptosistema lineal. En el caso en el que el conjunto de mensajes se identifica con $\mathbb{Z}_n$ y las claves son parejas $(a, b) \in \mathbb{Z}_n^* \times \mathbb{Z}_n$, para romper el criptosistema es suficiente encontrar dos parejas de mensajes básicos (m_1, c_1) y (m_2, c_2), donde c_i es el mensaje cifrado de m_i, que nos sirvan para resolver el sistema de ecuaciones:

$$\begin{aligned} am_1 + b &= c_1 \\ am_2 + b &= c_2 \end{aligned}$$

en $\mathbb{Z}_n$. Obsérvese que si los vectores (m_1, c_1) y (m_2, c_2) son linealmente independientes, lo cual sucede con alta probabilidad, entonces el sistema tiene solución única. ¿Esto es cierto? no estamos en espacios vectoriales...

En el caso en el que $M = C = \mathbb{Z}_n^d$ y $K = \mathrm{GL}_d(\mathbb{Z}_n) \times \mathbb{Z}_n^d$, el análisis es similar, aunque cuanto mayor sea d, más seguro será el criptosistema. En cualquier caso, si se cuenta con una cantidad suficientemente grande de parejas de mensajes básicos en claro con sus corresondientes cifrados, el sistema se puede romper de forma bastante fácil. Más concretamente, supongamos que disponemos de una lista de parejas de mensajes en claro y cifrados $(m_1, c_1), \ldots, (m_k, c_k)$, con $m_i, c_i \in \mathbb{Z}_n^d$, $i = 1, \ldots, k$ y deseamos calcular la clave (X, y) que se ha utilizado para cifrar con $X \in \mathrm{GL}_d(\mathbb{Z}_n)$ e $y \in \mathbb{Z}_n^d$. Planteamos el siguiente sistema de ecuaciones:

$$\left\{ \begin{array}{rcl} Xm_1 + y & = & c_1 \\ & \vdots & \\ Xm_k + y & = & c_k \end{array} \right.$$

Desarrollando el sistema se obtiene un sistema de ecuaciones lineales en el que las incógnitas son las entradas de la matriz X y el vector y. Si k es suficientemente grande probablemente el sistema es determinado y podremos encontrar la solución. Más concretamente, escribimos cada m_i como $(m_{i1}, \ldots, m_{id})$ y cada c_i como $(c_{i1}, \ldots, c_{id})$. Escribimos también las incógnitas X e y, que corresponden a la clave buscada, en términos de sus coordenadas:

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{d1} & x_{d2} & \cdots & x_{dd} \end{pmatrix}, \quad y = (y_1, \ldots, y_d)$$

Entonces la ecuación $Xm_i + y = c_i$ se convierte en un sistema de ecuaciones lineales:

$$\left. \begin{array}{rcl} x_{11}m_{i1} + x_{12}m_{i2} + \cdots + x_{12}m_{id} + y_1 & = & c_{i1} \\ x_{21}m_{i1} + x_{22}m_{i2} + \cdots + x_{22}m_{id} + y_2 & = & c_{i2} \\ & \vdots & \\ x_{d1}m_{i1} + x_{d2}m_{i2} + \cdots + x_{d2}m_{id} + y_d & = & c_{id} \end{array} \right\} \quad i = 1, 2, \ldots, k$$

Si para cada i elegimos la ecuación j-ésima obtenemos:

$$\left.\begin{array}{rcl} m_{11}x_{j1} + m_{12}x_{j2} + \cdots + m_{1d}x_{jd} + y_j &=& c_{1j} \\ m_{21}x_{j1} + m_{22}x_{j2} + \cdots + m_{2d}x_{jd} + y_j &=& c_{2j} \\ & \vdots & \\ m_{k1}x_{j1} + m_{k2}x_{j2} + \cdots + m_{kd}x_{jd} + y_j &=& c_{kj} \end{array}\right\} \quad j = 1, 2, \ldots, d$$

Este sistema es compatible pues los c_{ij} se han calculado a partir de los m_{ij} utilizando estas expresiones para ciertos valores de los x_{ij} e y_i. Obsérvese que todos los sistemas de ecuaciones lineales anteriores tienen la misma matriz de coeficientes:

$$A = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1d} & 1 \\ m_{21} & m_{22} & \cdots & m_{2d} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ m_{21} & m_{22} & \cdots & m_{2d} & 1 \end{pmatrix}$$

Si la matriz A tiene rango $d+1$, el sistema tiene una única solución que podemos obtener con un poquito de álgebra lineal. Pero esto solo se puede hacer en espacios vectoriales y estamos en $\mathbb{Z}_n$.

La solución obtenida proporciona la fila j-ésima de la matriz X y la coordenada j-ésima de y.

1.3. Entropía

Teorema 1.1. *Una función continua definida sobre el conjunto de funciones de distribución de longitud n que cumpla las condiciones:*

(1) $H\left(\frac{1}{n}, \ldots, \frac{1}{n}\right) < H\left(\frac{1}{n+1}, \ldots, \frac{1}{n+1}\right)$

(2) $H\left(\frac{1}{n}, \ldots, \frac{1}{n}\right) = H\left(\frac{k_1}{n}, \ldots, \frac{k_m}{n}\right) + \sum_{i=1, k_i \neq 0}^{m} \frac{k_i}{n} H\left(\frac{1}{k_i}, \ldots, \frac{1}{k_i}\right)$ *siempre que* $\sum_{i=1}^{m} k_i = n$

es de la forma

$$H(p_1, \ldots, p_n) = \sum_{i=1, p_i \neq 0}^{n} p_i \log_b\left(\frac{1}{p_i}\right) = -\sum_{i=1, p_i \neq 0}^{n} p_i \log_b p_i$$

para algún $b > 1$.

Demostración. Si $m|n$, entonces

$$H\left(\frac{1}{n}, \ldots, \frac{1}{n}\right) = H\left(\frac{m}{n}, \ldots, \frac{m}{n}\right) + \sum_{i=1}^{n/m} \frac{m}{n} H\left(\frac{1}{m}, \ldots, \frac{1}{m}\right) =$$
$$= H\left(\frac{m}{n}, \ldots, \frac{m}{n}\right) + H\left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$$

En particular, si $n = m^s$, entonces

$$H\left(\frac{1}{m^s}, \ldots, \frac{1}{m^s}\right) = H\left(\frac{1}{m^{s-1}}, \ldots, \frac{1}{m^{s-1}}\right) + H\left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$$

Sea $g(n) = H\left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$, entonces

$$g(m^s) = g(m^{s-1}) + g(m)$$

y por inducción sobre s, se obtiene que

$$g(m^s) = sg(m)$$

La condición (1), implica que g es estrictamente creciente y, por tanto, para todo $m > 1$ tenemos $g(m^s) < g(m^{s+1})$, es decir, $sg(m) < (s+1)g(m)$. Por tanto $g(m)$ es positivo.

Sean n, k y m enteros mayores a 1 y sea s

$$s = \text{máx}\left\{j \in \mathbb{Z} : j \geq 0, m^j \leq n^k\right\}$$

entonces $m^s \leq n^k < m^{s+1}$. Como g es estrictamente creceinte, $g(m^s) \leq g(n^k) \leq g(m^{s+1})$, o equivalentemente

$$sg(m) \leq kg(n) \leq (s+1)g(m)$$

Como log también es una función creciente también tenemos

$$s\log(m) \leq k\log(n) \leq (s+1)\log(m)$$

Por tanto,

$$\frac{s}{k} \leq \frac{g(n)}{g(m)} \leq \frac{s+1}{k} \quad \text{y} \quad \frac{s}{k} \leq \frac{\log(n)}{\log(m)} \leq \frac{s+1}{k}$$

luego

$$\left|\frac{g(n)}{g(m)} - \frac{\log(n)}{\log(m)}\right| \leq \frac{1}{k}$$

Como k es arbitrario,

$$\frac{g(n)}{g(m)} = \frac{\log(n)}{\log(m)}$$

es decir,

$$\frac{g(n)}{\log(n)} = \frac{g(m)}{\log(m)} = C$$

Luego $g(n) = C\log(n)$ para algún número postivo C. Por tanto, si elegimos una base b adecuada, tendremos que $g(n) = \log_b n$.

Supongamos ahora que $(p_1, \ldots, p_k)$ es una distribución de probabilidad formada por números racionales. Poniéndolos con común denominador podemos suponer que $p_i = \dfrac{b_i}{n}$ y, de la propiedad (2) tenemos,

$$H(p_1, \ldots, p_k) = H\left(\frac{b_1}{n}, \ldots, \frac{b_k}{n}\right) = g(n) - \sum_{i=1, b_i \neq 0}^{k} \frac{b_i}{n} g(b_i) = \log_b n - \sum_{i=1, b_i \neq 0}^{k} \frac{b_i}{n} \log_b b_i =$$

$$= \sum_{i=1, b_i \neq 0}^{k} \frac{b_i}{n} \log_b \frac{n}{b_i} = \sum_{i=1, p_i \neq 0}^{k} p_i \log_b \frac{1}{p_i}$$

Como H es continua, entonces

$$H(p_1, \ldots, p_k) = \sum_{i=1, p_i \neq 0}^{k} p_i \log_b \frac{1}{p_i}$$

para toda $k-$upla $(p_1, \ldots, p_k)$ de números reales en el dominio de H. □

Definición 1.1. Sea b un número real mayor que 1. Se llama *entropía* en base b de una distribución de probabilidad $P = (p_1, \ldots, p_k)$ a

$$H_b(p_1, \ldots, p_k) = \sum_{i=1}^{k} p_i \log_b \frac{1}{p_i}$$

La entropía de una variable aleatoria discreta es la entropía de su distribución de probabilidad.

La base b en la que se calcule la función de entropía sólo implica un cambio de escala debido a la igualdad $\log_b x = \log_{b'} x \cdot \log_b b'$ que implica

$$H_b(X) = H_{b'}(X) \cdot \log_b b'$$

Proposición 1.1. *Sea $(p_1, \ldots, p_n, q_1, \ldots, q_m)$ una distribución de probabilidad. Si $a = \sum_{i=1}^{n} p_i$, con $0 < a < 1$ entonces*

$$H(p_1, \ldots, p_n, q_1, \ldots, q_m) = H(a, 1-a) + aH\left(\frac{p_1}{a}, \ldots, \frac{p_n}{a}\right) + (1-a)H\left(\frac{q_1}{1-a}, \ldots, \frac{q_m}{1-a}\right)$$

Demostración.

$$\begin{aligned}
H(p_1, \ldots, p_n, q_1, \ldots, q_m) &= \sum_{i=1}^{n} p_i \log \frac{1}{p_i} + \sum_{i=1}^{m} q_i \log \frac{1}{q_i} = \sum_{i=1}^{n} p_i \log \frac{a}{ap_i} + \sum_{i=1}^{m} q_i \log \frac{1-a}{(1-a)q_i} = \\
&= \sum_{i=1}^{n} p_i \left(\log \frac{a}{p_i} + \log \frac{1}{a}\right) + \sum_{i=1}^{m} q_i \left(\log \frac{1-a}{q_i} + \log \frac{1}{1-a}\right) = \\
&= \sum_{i=1}^{n} p_i \log \frac{a}{p_i} + \sum_{i=1}^{n} p_i \log \frac{1}{a} + \sum_{i=1}^{m} q_i \log \frac{1-a}{q_i} + \sum_{i=1}^{m} q_i \log \frac{1}{1-a} = \\
&= a \log \frac{1}{a} + (1-a) \log \frac{1}{1-a} + \sum_{i=1}^{n} p_i \log \frac{a}{p_i} + \sum_{i=1}^{m} q_i \log \frac{1-a}{q_i} = \\
&= H(a, 1-a) + a \sum_{i=1}^{n} \frac{p_i}{a} \log \frac{a}{p_i} + (1-a) \sum_{i=1}^{m} \frac{q_i}{1-a} \log \frac{1-a}{q_i} = \\
&= H(a, 1-a) + aH\left(\frac{p_1}{a}, \ldots, \frac{p_n}{a}\right) + (1-a)H\left(\frac{q_1}{1-a}, \ldots, \frac{q_m}{1-a}\right)
\end{aligned}$$

□

Vamos ahora a ver cual es el rango de la función de entropía. Más concretamente vamos a demostrar el siguiente.

Teorema 1.2. *Sea X una variable aleatoria discreta con n sucesos posibles. Entonces*

$$0 \leq H_b(X) \leq \log_b n$$

Además $H_b(X) = 0$ precisamente si $P(X = x) = 1$ para algún suceso x y $H_b(X) = \log_b n$ si y sólo si la distribución de probabilidad de X es uniforme.

Para demostrar el Teorema 1.2 necesitaremos dos lemas. El primero es bien conocido:

Lema 1.1. *Para todo número real positivo x se verifica $\log x \leq x - 1$ y la igualdad se verifica precisamente si $x = 1$.*

El segundo es un poco más complicado:

Lema 1.2. *Sea* $P = (p_1, \ldots, p_n)$ *una distribución de probabilidad y* $Q = (q_1, \ldots, q_n) \in \mathbb{R}^n$ *con* $0 \leq q_i \leq 1$ *y* $\sum_{i=1}^n q_i \leq 1$. *Entonces*

$$\sum_{i=1, p_i \neq 0}^{n} p_i \log \frac{1}{p_i} \leq \sum_{i=1, q_i \neq 0} p_i \log \frac{1}{q_i}$$

Además la igualdad se verifica precisamente si $p_i = q_i$ *para todo* i.

Demostración. Del lema 1.1 se tiene que si $p \neq 0$ y $q \neq 0$ entonces

$$\log \frac{q}{p} \leq \frac{q}{p} - 1$$

y, por tanto,

$$p \log \frac{1}{p} \leq p \log \frac{1}{q} + q - p$$

Puesto que $\sum_{i=1}^n q_i \leq 1 = \sum_{i=1}^n p_i$, se tiene

$$\sum_{i=1, p_i \neq 0}^{n} p_i \log \frac{1}{p_i} \leq \sum_{i=1, p_i \neq 0, q_i \neq 0}^{n} \left(p_i \log \frac{1}{q_i} + q_i - p_i \right) \leq \sum_{i=1, q_i \neq 0}^{n} p_i \log \frac{1}{q_i}$$

Supongamos que se da la igualdad, esto es,

$$\sum_{i=1, p_i \neq 0}^{n} p_i \log \frac{1}{p_i} = \sum_{i=1, q_i \neq 0} p_i \log \frac{1}{q_i}$$

Entonces,

$$p_i \log \frac{1}{p_i} = p_i \log \frac{1}{q_i} + q_i - p_i$$

para todo i con $p_i \neq 0$ y $q_i \neq 0$, o equivalentemente,

$$\log \frac{q_i}{p_i} = \frac{q_i}{p_i} - 1$$

Pero del lema 1.1, esto equivale a que $p_i = q_i$ para todo i.

□