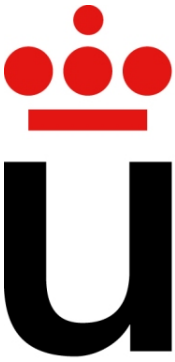
Ciencias Experimentales

Universidad Rey Juan Carlos
Servicio de Publicaciones

Victoria Ruiz Parrado
Joaquín Arias Herrero
Marina Cuesta Santa Teresa

Matemáticas Discreta y Álgebra

ISBN: 978-84-09-44732-9



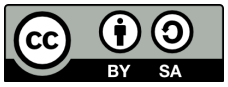
Universidad Rey Juan Carlos

ESCUELA TÉCNICA SUPERIOR
DE INGENIERÍA INFORMÁTICA

Matemáticas Discreta y Álgebra

TEORÍA Y PRÁCTICA POR Y PARA
LA COMPUTACIÓN Y LA CIBERSEGURIDAD

Autores: **Victoria Ruiz-Parrado**
Joaquín Arias
Marina Cuesta



Copyright ©2022 Victoria Ruiz-Parrado, Joaquín Arias, Marina Cuesta. Este obra está bajo la licencia CC BY-SA 4.0, Creative Commons Atribuciónn-CompartirIgual 4.0 Internacional. Como citar esta obra: Ruiz-Parrado, Victoria, Arias, Joaquín, y Cuesta, Marina (2022). Matemáticas Discreta y Álgebra: Teoría y Práctica por y para la Computación y la Ciberseguridad. Madrid: Servicio de Publicaciones de la Universidad Rey Juan Carlos. ISBN: 978-84-09-44732-9____

Resumen

La Matemática Discreta surge como una disciplina que unifica diversas áreas tradicionales de las Matemáticas (combinatoria, probabilidad, geometría de polígonos, aritmética, grafos, entre otros), como consecuencia, de ahí su interés en la informática, las telecomunicaciones, y en particular, la ciberseguridad. La información se manipula y almacena en los ordenadores en forma discreta (palabras formadas por ceros y unos), se necesita contar objetos (unidades de memorias, unidades de tiempo), se precisa estudiar relaciones entre conjuntos finitos (búsquedas en bases de datos), y es necesario analizar procesos que incluyan un número finito de pasos (algoritmos). La matemática discreta proporciona, por otro lado, algunas bases matemáticas para otros aspectos de la informática, como las estructuras de datos, algorítmica, bases de datos, teoría de autómatas, sistemas operativos y la investigación operativa. A su vez ayuda al desarrollo de ciertas capacidades fundamentales como la capacidad de formalizar, de razonar rigurosamente y/o de representar adecuadamente algunos conceptos.

El Álgebra Lineal es, seguramente, una de las herramientas fundamentales en las Ciencias de la Computación. Originariamente dedicada a la resolución de sistemas de ecuaciones, su abstracción y formalismo la hacen a veces un poco árida de entender. Sin embargo la inmensidad de sus aplicaciones bien vale el esfuerzo: Teoría de la Información, Teoría de Códigos, Ecuaciones Diferenciales, Optimización, etc.

Este manual combina Matemática Discreta y Álgebra Lineal, y es esencial para formar la base adecuada para comprender los modelos matemáticos que se presentan durante el desarrollo profesional en el campo de la informática y la ciberseguridad. Los objetivos que se buscan con este manual son aprender y utilizar técnicas y métodos propios de la Matemática Discreta y del Álgebra Lineal y su aplicación en el campo de la informática y la ciberseguridad. En concreto:

- Aprender métodos y conceptos básicos de algoritmos, aritmética modular, combinatoria y teoría de grafos.
- Discutir y resolver sistemas de ecuaciones lineales mediante el método de Gauss. Matrices y determinantes.
- Conocer la estructura de espacio vectorial, manejar la noción de aplicación lineal y su aplicación en diversos campos de la computación.
- Reconocer cuándo una matriz es diagonalizable y, si es así, saber diagonalizarla.

Agradecimientos

Este manual es el resultado de un trabajo de creación y mejora durante la docencia de la asignatura de "Matemáticas Discreta y Algebra" en el **Grado de Ingeniería de la Ciberseguridad** en los cursos 2019-2020, 2020-2021 y 2021-2022 (impartidos en el campus de Móstoles de la Escuela Técnica Superior de Ingeniería Informática de la Universidad Rey Juan Carlos). Con este trabajo queremos: (i) agradecer a los profesores que nos ayudaron a **amar las matemáticas**: Mercedes (Colegio Bérriz en las Rozas), profesores del DMATIC (Universidad Politécnica de Madrid), los profesores de la carrera de Matemáticas y Estadística de la Universidad Complutense de Madrid, y (ii) sembrar en alumnos *pasados* y futuros el deseo de conocer, aprender y **comprender** las matemáticas.

Para elaborar este texto nos hemos inspirado en nuestros apuntes de clase (de cuando éramos estudiantes) en presentaciones existentes, y hemos adaptado definiciones y ejemplos provenientes de diversos libros, incluyendo los siguientes textos:

Bratley, Paul y Brassard, Gilles (1996). **Fundamentos De Algoritmia**. Prentice Hall: London, England, pág. 600. ISBN: 978-8489660007.

Cirre-Torres, Francisco Javier (2004). **Matemática discreta**. Anaya Educación, pág. 184. ISBN: 978-8466730679.

Cobos-Gavala, Javier (2000). **Introducción a la Matemática Discreta**. Apuntes de clase. Departamento de Matemática Aplicada I: Universidad de Sevilla, pág. 124.

Gamboa-Mutuberría, José Manuel y M. Belén, Rodríguez-Rodríguez (2004). **Álgebra matricial**. Anaya Educación, pág. 144. ISBN: 978-8466726061.

Lay, David C. (2013). **Álgebra lineal y sus aplicaciones**. 4.ª ed. Pearson: Upper Saddle River, NJ. ISBN: 978-6073237451.

Merino-Gonzalez, Luis Miguel y Santo-Alaez, Evangelina (2006). **Álgebra Lineal Con Métodos Elementales**. Ediciones Paraninfo, S.A., pág. 408. ISBN: 978-8497324816.

Montenegro, José A. (2013). **Tema 6. Códigos lineales**. Transparencias. Departamento de Lenguajes y Ciencias de la Computación: Universidad de Málaga, pág. 76.

Rosen, Kenneth H. (2004). **Matemática discreta y Sus aplicaciones**. 5.ª ed. McGraw-Hill Interamericana de España S.L., pág. 888. ISBN: 978-8448140731.

Smart, Nigel P. (2002). **Cryptography**. McGraw Hill Education: Maidenhead, England, pág. 456. ISBN: 978-0077099879.

Índice general

Parte I

Matemáticas Discreta

Universidad Rey Juan Carlos

Tema 1

Algoritmos

1.1. Algoritmos

En matemáticas discreta, aparecen muchas clases de problemas genéricos. Por ejemplo, dada una sucesión de números enteros, encontrar el mayor; dado un conjunto, enumerar todos sus subconjuntos; dado un conjunto de enteros, ponerlos en orden creciente; dada una red de ordenadores, encontrar el camino más corto entre dos nodos. Cuando se presentan tales problemas, lo primero que debemos hacer es construir un modelo que traduzca el problema a un contexto matemático.

Fijar el modelo matemático apropiado es solamente una parte de la solución. Para completarla se necesita un método que resuelva el problema general utilizando el modelo. Idealmente, lo que se requiere es un procedimiento que siga una secuencia de pasos que conduzca a la repuesta deseada. Tal secuencia de pasos se denomina **algoritmo**. Formalmente,

Definición 1.1 (Algoritmo) *Conjunto finito de instrucciones precisas que sirve para realizar un cálculo o resolver un problema.*

Ejemplo 1.1 *Describe un algoritmo para encontrar el máximo valor de una sucesión finita de enteros.*

Un método consiste simplemente en utilizar el lenguaje natural para describir la sucesión de pasos utilizada. Ésta es la solución que proporciona este método:

1. *Fijamos provisionalmente el máximo igual al primer elemento de la sucesión.*
2. *Comparamos el siguiente entero de la sucesión con el máximo provisional, y si es más grande que el máximo provisional, fijamos el máximo provisional como este entero.*
3. *Se repite el paso anterior si hay más enteros en la sucesión.*
4. *Paramos cuando no queden más enteros en la sucesión por comparar. Llegado este punto, el máximo provisional es el mayor entero de la sucesión.*

Un algoritmo se puede describir empleando un lenguaje informático. Dado que solo se puede utilizar las instrucciones de dicho lenguaje, a veces la descripción del algoritmo es difícil de entender. Como comúnmente se utilizan varios lenguajes de programación distinto, para describir un algoritmo se suele utilizar el pseudocódigo. El pseudocódigo proporciona un paso intermedio entre una descripción de este algoritmo en lenguaje natural y en lenguaje informático. Es decir, los pasos del algoritmo se especifican usando instrucciones que recuerdan a las utilizadas en programación, pero se pueden incluir cualquier operación o sentencia bien definida.

Ejemplo 1.2 *Descripción en pseudocódigo del algoritmo para encontrar el máximo elemento de una sucesión finita.*

Algorithm 1: Búsqueda del elemento máximo en una sucesión finita

Input: $a_1, a_2, \cdots, a_n$; enteros
Output: Elemento mayor

```
1 max := a_1
2 for i:=2 to n do
3 |   if max <a_i then
4 |   |   max := a_i
5 |   end
6 end
7 return max
```

Este algoritmo primero asigna el término inicial de la sucesión a_1 a la variable *max*. El bucle *for* se usa para examinar sucesivamente los términos de la sucesión. Si un término a_i es mayor que el valor de *max* en ese momento, se asigna éste como nuevo valor de *max*.

Hay varias propiedades que generalmente comparten los algoritmos:

- *Entrada*: Un algoritmo tiene valores de entrada que son elementos de un conjunto especificado.
- *Salida:* Para cada conjunto de valores de entrada, un algoritmo produce valores de salida de un conjunto especificado. Los valores de salida son la solución del problema.
- *Definición:* Los pasos de un algoritmo deben definirse con precisión.
- *Corrección:* Un algoritmo debe producir valores de salida correctos para cada conjunto de valores de entrada.
- *Duración finita:* Un algoritmo debe producir la salida deseada tras un número finito (aunque quizás grande) de pasos para cualquier conjunto de valores de entrada.
- *Efectividad:* Debe ser posible realizar cada paso del algoritmo con exactitud y en un intervalo finito de tiempo.
- *Generalidad:* El procedimiento debería ser aplicable a todos los problemas de la forma deseada, no sólo para un conjunto particular de datos de entrada.

1.1.1. Algoritmos de Búsqueda

El problema de localizar un elemento de una lista ordenada se puede encontrar en muchos contextos. Por ejemplo, un programa que revisa que las palabras de un texto estén correctamente escritas busca estas palabras en un diccionario, que no es más que una lista de palabras ordenadas. Los problemas de este tipo se llama **problemas de búsqueda**.

Un problema de búsqueda general se puede describir como: localizar un elemento x en una lista de elementos distintos $a_1, a_2, \cdots, a_n$ o determinar que no está en la lista. La solución a este problema de búsqueda es la localización del término en la lista que es igual que x (esto es la solución es i si $x = a_i$) y es 0 si x no está en la lista.

Algunos tipos de búsqueda son:

Búsqueda Lineal El algoritmo de búsqueda lineal o búsqueda secuencial comienza por comparar x y a_1. Cuando $x = a_1$, la solución resulta ser la localización de a_1, i.e, 1. Cuando $x \neq a_1$ se compara con a_2. Este proceso se continúa hasta que se encuentra una coincidencia. La solución es la localización de ese término, a no ser que esta coin-

Algorithm 2: Algoritmo de Búsqueda Lineal

Input: $x, a_1, a_2, \cdots, a_n$
Output: Localización de x

```
1  i := 1
2  while i ≤ n y x ≠ a_i do
3  |   i := i + 1
4  end
5  if i ≤ n then
6  |   localizacion:=i
7  end
8  else
9  |   localizacion:=0
10 end
11 return localizacion
```

cidencia no ocurra. Si se ha recorrido la lista completa sin localizar x, la solución es 0. El pseudocódigo para el algoritmo de búsqueda lineal se encuentra en el Algoritmo 2.

Ejemplo 1.3 *Encuentra $x = 7$ en la siguiente lista:*

1	8	7	9	3

Para ello:

1. *Primero comprobamos que $7 \neq a_1 = 1$*
2. *Como no eran iguales comprobamos si se cumple con el siguiente elemento de la lista: $7 \neq a_2 = 8$*
3. *Probamos con el tercer elemento: $7 = a_3 = 7$.*
4. *Como sí ha habido coincidencia se devuelve $locacalizacion = 3$*

Búsqueda Binaria El algoritmo de búsqueda binaria se puede usar cuando la lista tiene los términos en orden creciente de tamaño. Se desarrolla comparando el elemento que se quiere localizar con el elemento central de la lista. La lista entonces se parte en dos sublistas más pequeñas. La búsqueda continúa restringiéndose a la lista apropiada, basándose en la comparación del elemento que se desea localizar con el término central. El pseudocódigo se puede ver en el Algoritmo 3.

Ejemplo 1.4 *Buscar el número 19 en la lista:*

1	*2*	*3*	*5*	*6*	*7*	*8*	*10*	*12*	*13*	*15*	*16*	*18*	*19*	*20*	*22*

Algorithm 3: Algoritmo de Búsqueda Binaria

Input: x, entero; $a_1, a_2, \cdots, a_n$, enteros ordenados de manera creciente
Output: Localización de x

```
1  i := 1                          // i es el extremo izquierdo del intervalo de búsqueda
2  j := n                          // j es el extremo derecho del intervalo de búsqueda
3  while i < j do
4  |  begin
5  |  |  m := ⌊(i+j)/2⌋
6  |  |  if x > a_m then
7  |  |  |  i := m+1
8  |  |  end
9  |  |  else
10 |  |  |  j := m
11 |  |  end
12 |  end
13 end
14 if x = a_i then
15 |  localizacion := i
16 end
17 else
18 |  localizacion := 0
19 end
20 return localizacion
```

1. *Hacemos i=1; j=16*
2. *Como* $i = 1 < j = 16$*, calculamos* $m = \lfloor (i+j)/2 \rfloor = \lfloor (1+16)/2 \rfloor = 8$*, y partimos la lista en dos sublistas:una de* a_1 *a* a_8*, y la otra de* a_9 *a* a_{16}*:*

1	*2*	*3*	*5*	*6*	*7*	*8*	*10*

12	*13*	*15*	*16*	*18*	*19*	*20*	*22*

3. *Se compara el elemento buscado (19) con el mayor término de las listas. Como 19>*$a_8 = 10$*, hacemos i=8+1=9.*
4. *Como* $i = 9 < j = 16$ *Volvemos a calcular* $m = \lfloor (i+j)/2 \rfloor = \lfloor (9+16)/2 \rfloor = 12$ *y partimos la sublista seleccionada en otras dos sublistas: una del elemento* a_9 *al* a_12*, y la otra del* a_{13} *al* a_{16}

12	*13*	*15*	*16*

18	*19*	*20*	*22*

5. *Como* $x = 19 > a_9 = 12$*, hacemos i=12+1=13.*
6. *Como i=13<j=16, calculamos* $m = \lfloor (13+16)/2 \rfloor = 14$*, partimos la lista en:*

18	*19*

20	*22*

7. *Como* $x = 19 \not> a_{14}$*,hacemos* $j = m = 14$*.*
8. *Como* $i = 13 < j = 14$*, calculamos* $m = \lfloor (13+14)/2 \rfloor = 13$*, dividimos la lista*

en

18		*19*

9. *Como* $x = 19 > a_{13} = 18$*, hacemos i=13+1=14.*
10. *Como* $i < j$ *hemos terminado. Como* $x = 19 = a_{14}$*, devolvemos localizacion* $=$ 14.

1.1.2. Algoritmos de Ordenación

Ordenar los elementos de una lista es un problema que se presenta en muchos contextos. Por ejemplo, para generar un directorio telefónico es necesario ordenar alfabéticamente los nombres de los usuarios.

Supongamos que tenemos una lista de elementos de un conjunto. Además, supongamos que conocemos una forma de ordenar estos conjuntos. Una **ordenación** es colocar estos elementos en una lista en la cual los elementos se disponen en orden creciente. Por ejemplo, ordenar la lista 7,2,1,4,5,6 produce la lista 1,2,4,5,7 y 9.

Este tipo de algoritmos han sido muy estudiados debido al alto coste computacional que tienen. Algunos de los métodos más simples, aunque no eficientes, son:

Método de la Burbuja Este método coloca los elementos de una lista creciente mediante la comparación sucesiva de elementos adyacentes, intercambiándolos si están en el orden equivocado. Para desarrollar el método de la burbuja, llevamos a cabo la operación básica, intercambiar un elemento mayor por otro menor que le siga, comenzando por el principio de la lista hasta llegar al final. Iteramos este proceso hasta que la operación se completa. Un pseudocódigo de este algoritmo es el Algoritmo 4

Ejemplo 1.5 *La ordenación de 3,2,4,1,5 mediante el método de la burbuja será:*

1. *Como* $a_1 > a_2$ *se intercambian estos valores:* *3,2,4,1,5* $\Rightarrow$ *2,3,4,1,5*
2. *Como* $a_1 < a_3$ *no hago nada. Como* $a_1 > a_4$ *se intercambian estos valores:* *2,3,4,1,5* $\Rightarrow$ *1,3,4,2,5*
3. *Como* $a_1 < a_5$*, no hago nada y paso a iterar con* a_2*. Como* $a_2 < a_3$ *no hago nada. Como* $a_2 > a_4$ *intercambios estos elementos:* *1,3,4,2,5* $\Rightarrow$ *1,2,4,3,5*
4. *Como* $a_2 < a_5$*, no hago nada y paso a iterar con* a_3*. Como* $a_3 > a_4$ *intercambios estos elementos:* *1,2,4,3,5* $\Rightarrow$ *1,2,3,4,5*
5. *Seguiríamos iterando hasta acabar todas las comprobaciones, pero en ninguna más habría que intercambiar valores, puesto que ya está ordenada.*

Algorithm 4: Algoritmo de Ordenación de la Burbuja

Input: $lista = [a_1, a_2, \cdots, a_n]$

Output: Lista ordenada

```
1  for i:=2 to n do
2      for j:=1 to n-i+1 do
3          if a_j > a_{j+1} then
4              aux=lista[a_j]
5              lista[a_j]=lista[a_{j+1}]
6              lista[a_{j+1}]=aux
7          end
8      end
9  end
10 return lista
```

Ordenación por Inserción Para ordenar una lista con n elementos, la ordenación por inserción comienza por el segundo elemento. Se compara este segundo elemento con el primero y se coloca antes del primero si no es mayor que el primer elemento, y tras el primer elemento si es mayor que éste. En este punto, los dos primeros elementos están en el orden correcto. El tercer elemento se compara con el primero, y si es mayor que él, se compara con el segundo. Se coloca en la posición correcta entre los tres primeros elementos.

En general, en el paso j-ésimo, el elemento j-ésimo de la lista se inserta en la posición correcta de la lista formada por los j-1 elementos previamente ordenados. Para insertar el elemento j-ésimo se utiliza una búsqueda lineal. El pseudocódigo se puede ver en Algoritmo 5

Ejemplo 1.6 *La ordenación de 3,2,4,1,5 mediante el método de la inserción será:*

1. *La ordenación por inserción compara primero $a_1 = 2$ con $a_2 = 3$. Como 3>2 coloca 2 en la primera posición, quedando los dos primeros elementos ordenados:* 3,2,4,1,5 $\Rightarrow$ 2,3,4,1,5
2. *Posteriormente, se inserta el tercer elemento, 4, en la parte ya ordenada en las lsitas mediante las comparaciones 4>2 y 4>3. Como 4>3, se coloca en la tercera posición:* 2,3,4,1,5 $\Rightarrow$ 2,3, 4 ,1,5
3. *Ahora buscamos la posición correcta para el cuarto elemento,1. Como 1<2, lo insertamos en la primera posición:* 2,3,4,1,5 $\Rightarrow$ 1 ,2,3,4,5
4. *Finalmente, colocamos 5 en la posición correcta por comparaciones sucesivas con el resto de elementos. Como 5>4 va al final de la lista produciendo el orden correcto para la lista completa.*

Algorithm 5: Algoritmo de Ordenación por Inserción

Input: $lista = [a_1, a_2, \cdots, a_n]$
Output: Lista Ordenada

```
1  for j:=2 to n do
2      begin
3          i:=1
4          while a_j > a_i do
5              i:=i+1
6          end
7          m := a_j
8          for k:=0 to j-i-1 do
9              a_{j-k} := a_{j-k-1}
10         end
11         a_i := m
12     end
13 end
14 return lista
```

1.2. Complejidad Computacional

Para que un algoritmo proporcione una solución satisfactoria debe producir siempre la respuesta correcta y, además, ser eficiente. Una medida de eficiencia es el número de instrucciones que requiere el ordenador para resolver un problema utilizando un algoritmo para valores de entrada de un tamaño especificado.

Otra medida de eficacia sería la complejidad en espacio que está ligada a las estructuras de datos usadas en la implementación del algoritmo y que no vamos a tenerlas en cuenta en este curso.

1.2.1. Complejidad Temporal

La complejidad temporal de un algoritmo es la complejidad computacional que expresa el número de operaciones que realiza el algoritmo en relación al tamaño de los datos de entrada. Las operaciones utilizadas para medir la complejidad puede ser la comparación de enteros, la suma, multiplicación o división de enteros, así como cualquier otra operación básica.

La complejidad se describe en términos del número de operaciones repetidas, en lugar del tiempo de cálculo real (ya que estas medidas pueden cambiar en función del ordenador). Hay distintos modos de medir la complejidad: complejidad del caso peor y complejidad del caso medio.

Complejidad Caso Peor Por comportamiento de un algoritmo en el peor caso entendemos el mayor número de operaciones que hace falta para resolver el problema dado utilizando el algoritmo para unos datos de entrada de un determinado tamaño. Los análisis del peor caso nos dicen cuántas operaciones tienen que realizar los algoritmos para garantizar que producirán una solución.

Ejemplo 1.7 *Se desea describir la complejidad del Algoritmo 2. Para ello se observa que en cada paso del bucle se llevan a cabo dos comparaciones: una para ver si se ha alcanzado el final de la lista, y otra para comparar el elemento x con un término de la lista. Finalmente, fuera del bucle se hace una comparación más. Por tanto, si* $x = a_i$*, se hacen* $2i+1$ *comparaciones. El mayor número de comparaciones* $(2n+2)$ *se alcanza cuando el elemento no está en la lista. En este caso se hacen* $2n$ *comparaciones para determinar que* $x \neq a_i$*, una comparación para salir del bucle, y una comparación más fuera del bucle. Por tanto, una búsqueda lineal requiere como mucho* $O(n)$ *comparaciones.*

Complejidad del Caso Promedio En este tipo de análisis se busca el número promedio de operaciones realizadas para solucionar un problema considerando todas las posibles entradas de un tamaño determinado. El análisis de la complejidad del caso promedio es, generalmente, mucho más complicado que el análisis del caso peor.

Ejemplo 1.8 *Se desea describir el comportamiento en el caso promedio del algoritmo de búsqueda lineal suponiendo que el elemento x esta en la lista.*

Hay n tipos posibles de entradas cuando sabemos que x está en la lista. Vimos en el ejemplo anterior, que si x es el término i-ésimo de la lista, se necesitan $2i+1$ *comparaciones. Por tanto, el número promedio de comparaciones realizadas es igual a:*

$$\frac{3+5+7+\cdots+(2n+1)}{n} = \frac{1}{n}\sum_{i=1}^{n}(2i+1) = \frac{1}{n}\cdot\frac{(3+2n+1)n}{2} = n+2 \qquad (1.1)$$

ya que $\sum_{i=1}^{n}\frac{(a_i+a_n)n}{2}$*. Por tanto, la complejidad en caso promedio es* $\Theta(n)$*.*

1.2.2. Notaciones Asintóticas

Hemos dicho que deseamos determinar matemáticamente la cantidad de recursos que necesita el algoritmo como función de su tamaño. Sin embargo, no existe ningún ordenador que pueda comparar todas las medidas de tiempo de ejecución. Para ello, se utiliza la notación asintótica. Esta notación permite realizar simplificaciones sustanciales, aun cuando estemos interesados en medir algo más tangible que el tiempo de

ejecución, tal como el número de veces que se ejecuta una instrucción dada dentro de un programa.

Esta notación se denomina asintótica porque trata acerca del comportamiento de las funciones en el límite, es decir, para valores suficientemente grandes de un parámetro. Aunque los argumentos basados en la notación asintótica puedan no tener un valor práctico en casos reales, sirve para comparar algoritmos y escoger aquellos más eficientes.

Antes de definir formalmente la notación asintótica, vamos a ver cómo se pueden comparar funciones. Supongamos que tenemos dos funciones $f(n)$ y $g(n)$. Se dice que:

- $f(n)$ es asintóticamente menor que $g(n)$ cuando:

$$f(n) < g(n) \Leftrightarrow \lim_{n\to\infty} \frac{f(n)}{g(n)} = 0 \tag{1.2}$$

- $f(n)$ es asintóticamente mayor que $g(n)$ cuando:

$$f(n) > g(n) \Leftrightarrow \lim_{n\to\infty} \frac{g(n)}{fg(n)} = 0 \tag{1.3}$$

- $f(n)$ es asintóticamente igual que $g(n)$ cuando:

$$f(n) = g(n) \Leftrightarrow \lim_{n\to\infty} \frac{f(n)}{g(n)} = 1 \tag{1.4}$$

Así, $f(n)$ es asintóticamente menor o igual que $g(n)$, se dice que $g(n)$ es una cota superior de $f(n)$ asintóticamente. Es decir,

$$f(n) \in O(g(n)) \Leftrightarrow f(n) \leq g(n) \tag{1.5}$$

Sin embargo, lo que nos interesa es la cota superior más baja. Formalmente,

Definición 1.2 (Cota superior O) *Se dice que $g(n)$ es una cota superior de $f(n)$ si*

$$O(g(n)) = \{f(n) : \ \exists c > 0 \ \ y \ \ n_0 > 0 / 0 \leq f(n) \leq c \cdot g(n), \forall n \geq n_0\} \tag{1.6}$$

Es decir, a partir de n_0, $c \cdot g(n)$ siempre supera o iguala a $f(n)$, como se puede ver en la Figura 1.1

Ejemplo 1.9 *Para demostrar que una función $f(n) \in O(g(n))$, será necesario encontrar una (cualquiera) pareja de constantes $c >=$ y $n_o > 0$, de tal forma que se verifiquen las condiciones de la definición. Por ejemplo, para demostrar que $5n+2 \in O(n)$:*

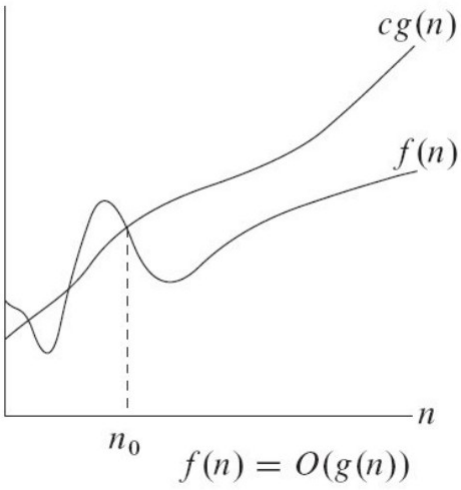


Figura 1.1: Cota superior O

- *Hay que encontrar* $c > 0$ *y* $n_0 > 0$ *tales que* $5n+2 \leq cn, \forall n \geq n_0$.
- *Cogiendo* $c = 6$, *tenemos que* $5n+2 \leq 6n \Rightarrow n \geq 2$.
- *Como para* $c = 6$ *se cumple para todo* $n \geq 2$, *luego podemos tomar* $n_0 = 2$, *y hemos encontrado una pareja de constantes (hay infinitas parejas más, pero basta con encontrar una.*

Análogamente, se define la cota inferior Ω:

Definición 1.3 (Cota inferior Ω) *Se dice que* $g(n)$ *es una cota inferior de* $f(n)$ *si*

$$\Omega(g(n)) = \{f(n) : \exists c > 0 \ \ y \ \ n_0 > 0 / 0 \leq c \cdot g(n) \leq f(n), \forall n \geq n_0\} \qquad (1.7)$$

Es decir, a partir de n_0, $f(n)$ siempre supera o iguala a $c \cdot g(n)$, como se puede ver en la Figura 1.2

Ejemplo 1.10 *Para demostrar que una función* $f(n) \in \Omega(g(n))$, *será necesario encontrar una (cualquiera) pareja de constantes* $c >=$ *y* $n_0 > 0$, *de tal forma que se verifiquen las condiciones de la definición. Por ejemplo, para demostrar que* $3n^2+2 \in \Omega(n)$*:*

- *Hay que encontrar* $c > 0$ *y* $n_0 > 0$ *takes que* $c \cdot n \leq 3n^2+2, \forall n \geq n_0$.
- *Cogiendo* $c = 5$, *tenemos que* $3n^2+2 \geq 5n \Rightarrow 3n^2-5n+2 \geq 0 \Rightarrow (n-2/3)(n-1) \geq 0$.
- *Por tanto, siempre será positiva para* $n \geq 1$. *Tomando* $c = 5$ *y* $n_0 = 1$, *se cumplen las condiciones de la definición y queda demostrado.*

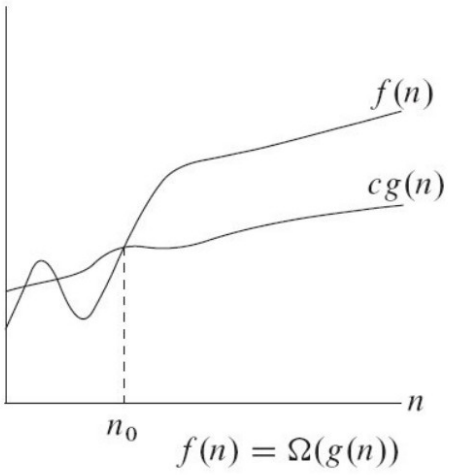


Figura 1.2: Cota inferior Ω

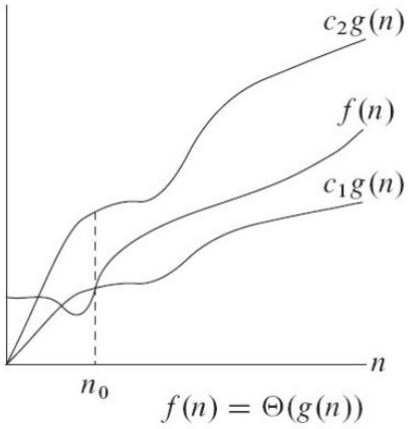


Figura 1.3: Cota ajustada Θ

Por último, definimos la cota ajustada Θ

Definición 1.4 (Cota ajustada Θ) *Se dice que $g(n)$ es una cota ajustada de $f(n)$ si*

$$\Theta(g(n)) = \{f(n) : \exists c_1 > 0, c_2 > 0 \;\; y \;\; n_0 > 0/0 \leq c_1 \cdot g(n) \leq f(n) \leq c_2 \cdot g(n), \forall n \geq n_0\} \tag{1.8}$$

Es decir, a partir de n_0, $f(n)$ siempre queda en medio de $c_1 \cdot g(n)$ y $c_2 \cdot g(n)$, como se puede ver en la Figura 1.3

Nótese que una función $g(n)$ es cota ajustada de otra $f(n)$ si es al mismo tiempo cota superior y cota inferior de $f(n)$. Es decir:

$$f(n) \in \Theta(g(n)) \Leftrightarrow \begin{cases} f(n) \in O(g(n)) \\ y \\ f(n) \in \Omega(g(n)) \end{cases} \tag{1.9}$$

Ejemplo 1.11 *Para demostrar que una función $f(n) \in \Theta(g(n))$, será necesario encontrar una (cualquiera) trío de constantes $c_1 > 0, c_2 > 0$ y $n_0 > 0$, de tal forma que se verifiquen las condiciones de la definición. Por ejemplo, para demostrar que $n^2/2 - 3n \in \Theta(n^2)$:*

- *Se busca $c_1 n^2 \leq n^2/2 - 3n \leq c_2 n^2$.*
- *Tomando $c_1 = 1/14$, tenemos que $n^2/14 \leq n^2/2 - 3n \Leftrightarrow n^2 \leq 7n^2 - 14 \cdot 3 \cdot n \Leftrightarrow 0 \leq 6n^2 - 14 \cdot 3 \cdot n \Leftrightarrow 0 \leq 6n(n-7) \Leftrightarrow n \geq 7$*
- *Por otra parte, si tomando $c_2 = 1/2$, tenemos que $n^2/2 - 3n \leq n^2/2 \Leftrightarrow n \geq 0$*
- *Por tanto, tomando $c_1 = 1/14, c_2 = 1/2$ y $n_0 = 7$ se cumplen las condiciones de la definición.*

Los comportamientos asintóticos de más frecuente aparición se pueden ordenar de menor a mayor crecimiento de la siguiente forma:

$$1 \; < \; log(n) \; < \; n \; < \; n.log(n) \; < \; \mathbf{n^2} \; < \; n^3 \; < \; ..... \; < \; 2^n \; < \; n!$$

Universidad
Rey Juan Carlos

Tema 2

Aritmética Modular

2.1. Aritmética Modular

En este tema se van a estudiar los conceptos básicos de aritmética modular y sus aplicaciones, ya que es una parte fundamenta en criptografía moderna en general, y de sistemas cifrados de clave pública en particular.

La parte de la matemática discreta que estudia los enteros y sus propiedades pertenece a la rama de las matemáticas conocida como teoría de números. La aritmética modular es una de las partes más importantes de la teoría de números. Sin embargo, para entender bien la aritmética modular es necesario tener conocimientos básicos de aritmética entera: la divisibilidad y el máximo común denominador.

2.2. Introducción a la Aritmética Entera

La aritmética entera hace referencia al conjunto de números enteros $\mathbb{Z}$, sus propiedades y las operaciones que se pueden realizar con ellos. En esa sección se estudiarán los conceptos básicos de divisibilidad con enteros y del máximo común denominador. Todo ello nos servirá para definir el algoritmo de Euclides y la identidad de Bézout.

2.2.1. Divisibilidad

Teorema 2.1 *Si a y b son enteros con $b > 0$, existe un único par de enteros q y r tal que:*

$$a = q \cdot b + r \quad 0 \leq r < b \tag{2.1}$$

Ejemplo 2.1 *Si $a = 9$ y $b = 4$, como $9 = 2 \times 4 + 1$, con $0 \leq 1 < 4$, se tiene que q=2 y r=1.*

Ejemplo 2.2 *Si $a = -9$ y $b = 4$, como $-9 = -3 \times 4 + 3$ con $0 \leq 3 < 4$, se tiene que $q = -3$ y $r = 3$.*

Definición 2.1 *Con la notación del teorema anterior, el entero q recibe el nombre de* cociente entero *o simplemente* cociente *y el también entero r el de* resto*. Si dividimos por b obtenemos que*

$$\frac{a}{b} = q + \frac{r}{b} \tag{2.2}$$

por lo que q es el mayor enetero no superior a a/b, recibe el nombre de *suelo de a/b* y se representa por $\lfloor \frac{a}{b} \rfloor$.

De forma análoga, al menor entero no inferior al cociente a/b lo denotamos por $\lceil \frac{a}{b} \rceil$ y recibe el nombre de *techo de a/b*.

Ejemplo 2.3 *Vamos a probar que si n es un cuadrado perfecto, al dividirlo entre 4 sólo puede darnos como resto 0 ó 1.*

Sea n=a^2, por el teorema anterior se tiene que $a = 4q + r$ con $r = 0, 1, 2, 3$. Por tanto, $n = a^2 = (4q + r)^2 = 16q^2 + 8qr + r^2$. Por tanto:

- *Si $r = 0$, se tiene que $n = 4(4q^2) + 0 \Rightarrow$ el resto es 0.*
- *Si $r = 1$, se tiene que $n = 4(4q^2 + 2q) + 1 \Rightarrow$ el resto es 1.*
- *Si $r = 2$, se tiene que $n = 4(4q^2 + 4q + 1) + 0 \Rightarrow$ el resto es 0.*
- *Si $r = 3$, se tiene que $n = 4(4q^2 + 6q + 2) + 1 \Rightarrow$ el resto es 1.*

Definición 2.2 *Si a y b son enteros y $a = qb$ para algún entero q, diremos que b divide a, que b es un divisor (o factor) de a, o que a es múltiplo de b y se representa por $b|a$. Si b no divide a a, lo denotaremos por $b \nmid a$*

Teorema 2.2 (Propiedades de la Divisibilidad)

1. $a|b$ y $b|c \Rightarrow a|c$
2. $a|b$ y $c|d \Rightarrow ac|bd$
3. $m \neq 0 \Rightarrow a|b$ si y sólo si $ma|mb$
4. $d|a$ y $a \neq 0 \Rightarrow |d| \leq |a|$
5. *Si c divide a $a_1, a_2, \cdots, a_k \Rightarrow$ c divide a $a_1u_1 + a_2u_2 + \cdots a_ku_k$ cualesquiera que sean los enteros $u_1, u_2, \cdots, u_k$*

Lema 2.1 *Si c es un divisor de a y b, c divide a $au + bv$ cualesquiera que sean los enteros u y v.*

Definición 2.3 *Un entero positivo p mayor que 1 se llama* primo *si los únicos divisores de p son 1 y p. Un entero positivo mayor que 1 que no es primo se denomina* compuesto

Teorema 2.3 (Teorema Fundamental de la Aritmética) *Todo entero positivo mayor que 1 se puede escribir de una única forma como un primo o como el producto de dos o más primos en el que los factores primos se escriben en orden no decreciente.*

2.2.2. Máximo Común Divisor

Si $d|a$ y $d|b$ decimos que d es un *divisor común* o *factor común* de a y b; por ejemplo, 1 es un divisor común a cualquier par de enteros a y b.

Teorema 2.4 (Máximo Común Divisor) *Dados dos enteros a y b, no ambos nulos, se denomina máximo común divisor de a y b, y se denota por mcd(a,b), al mayor de sus divisores comunes, que existe y es único.*

Teorema 2.5 (Propiedades del Máximo Común Divisor)

1. $mcd(a,b) = mcd(b,a) = mcd(-a,b) = mcd(a,-b) = mcd(-a,-b)$
2. $mcd(a,a) = mcd(a,0) = a$
3. $mcd(a_1,a_2,\cdots,a_k) = mcd(mcd(a_1,a_2),a_3,\cdots,a_k)$

Algoritmo de Euclides

Si podemos factorizar a y b en primos, entonces es muy fácil calcula el máximo común divisor. Por ejemplo:

$$a = 230895588646864 = 2^4 \cdot 157 \cdot 4513^3$$

$$b = 33107658350407876 = 2^2 \cdot 157 \cdot 2269^3 \cdot 4513$$

entonces, el mcd se calcula como

$$mcd(a,b) = 2^2 \cdot 157 \cdot 4513 = 283164$$

Sin embargo, factorizar es una operación costosa. Por ello, se utiliza el algoritmo de Euclides.

Antes de describir el algoritmo de Euclides, demostraremos cómo se usa para calcular el $mcd(91,287)$. Primero, divide 287, el mayor de los dos números, por 91, el menor, para obtener

$$287 = 91 \cdot 3 + 14$$

Cualquier divisor de 91 y 287 debe ser un divisor de $287 - 91 \cdot 3 = 14$. Además, cualquier divisor de 91 y 14 debe ser un divisor de $287 = 91 \cdot 3 + 14$. Por tanto, el máximo común divisor de 91 y 287 es el mismo que el máximo común divisor de 91 y 14. Esto significa que el problema de hallar el mcd(91,287) se ha reducido al problema de calcular el mcd(91,14). Ahora dividimos 91 por 14, para obtener

$$91 = 14 \cdot 6 + 7$$

Como cualquier divisor de 91 y 14 también divide a $91 - 14 \cdot 6 = 7$ y cualquier divisor común de 14 y 7 divide a 91, se sigue que mcd(91,14)=mcd(14,7). Se continúa dividiendo por 14 y 7, para obtener $14 = 7 \cdot 2$, Como 7 divide a 14, se sigue que

mcd(14,7)=7. Además, como mcd(287,91)=mcd(91,14)=mcd(14,7), el problema queda resuelto.

Es decir, sabiendo que para enteros, la división euclídea es la operación de, dados a y b, encontrar q y r con $0 \leq r < |b|$ de manera que:

$$a = q \cdot b + r$$

se tiene el siguiente lema:

Lema 2.2 *Dados dos enteros a y b se verifica que* $mcd(a,b) = mcd(b,r)$ *cualesquiera que sean los enteros q y r verificando* $a = bq + r$

De aquí se obtiene el algoritmo de Euclides:

Teorema 2.6 (Algoritmo de Euclides) *Sean a y b dos enteros (no ambos nulos) y tratemos de calcular* $d = mcd(a,b)$ *donde podemos suponer que* $a > b > 0$. *Utilizando el Teorema 2.1, obtenemos:*

$$a = q_1 b + r_1 \quad con \quad 0 \leq r_1 < b \tag{2.3}$$

Dividiendo ahora b entre r_1 *se obtiene*

$$b = q_2 r_1 + r_2 \quad con \quad 0 \leq r_2 < r_1 \tag{2.4}$$

Repitiendo el proceso obtenemos una sucesión de restos r_i *con* $b > r_1 > r_2 > \cdots \geq 0$.

Al tratarse de una sucesión de enteros positivos estrictamente decreciente, llegará un momento en el que necesariamente sea $r_{m+1} = 0$ *y en ese punto finalizamos el proceso:*

$$d = mcd(a,b) = mcd(b,r_1) = \cdots = mcd(r_m, r_{m+1}) = mcd(r_m, 0) = r_m \tag{2.5}$$

Es decir, para calcular el mcd de $r_0 = a$ y $r_1 = b$, se calcula r_2, r_3, r_4, etc. de la siguiente manera:

$$\begin{aligned} r_2 &= r_0 - q_1 r_1 \\ r_3 &= r_1 - q_2 r_2 \\ &\vdots \\ r_m &= r_{m-2} - q_{m-1} r_{m-1} \\ 0 &= r_{m-1} - q_m r_m \end{aligned}$$

Si d divide a a y b, entonces divide a r_2, r_3, r_4 y así sucesivamente. Por tanto:

$$d = mcd(a,b) = mcd(r_0,r_1) = mcd(r_1,r_2) = \cdots = mcd(r_{m-1},r_m) = mcd(r_m,0) = r_m$$

El Algoritmo 6 muestra el algoritmo de Euclides en pseudocódigo.

Algorithm 6: Euclides

Input: a, b; enteros positivos
Output: mcd(a,b)

```
1 x:=a y:=b
2 while y ≠ 0 do
3 |   r:=x mod y
4 |   x:=y
5 |   y:=r
6 end
7 return x
```

Ejemplo 2.4 *Usar el algoritmo de Euclides para calcular mcd(21,12):*

$$\begin{aligned} mcd(21,12) &= mcd(21\ (mod\ 12),12) \\ &= mcd(9,12) \\ &= mcd(12\ (mod\ 9),9) \\ &= mcd(3,9) \\ &= mcd(9\ (mod\ 3),3) \\ &= mcd(0,3) = 3 \end{aligned}$$

Ejemplo 2.5 *Usar el algoritmo de Euclides para calcular mcd(1 426 668 559 730,810 653 094 756):*

$$\begin{aligned} mcd(1426668559730,810653094756) &= mcd(810653094756,616015464974) \\ &= mcd(616015464974,194637629782) \\ &= mcd(194637629782,32102575628) \\ &= mcd(32102575628,2022176014) \\ &= mcd(2022176014,1769935418) \\ &= mcd(1769935418,252240596) \\ &= mcd(252240596,4251246) \\ &= mcd(4251246,1417082) \\ &= mcd(1417082,0) \\ &= 1417082 \end{aligned}$$

Teorema 2.7 (Identidad de Bézout) *Si a y b son enteros (no ambos nulos) existen enteros u y v tales que*

$$mcd(a,b) = au + bv \tag{2.6}$$

Los enteros u y v no son únicos.

Volviendo al algoritmo de Euclides, despejando r_m en la penúltima igualdad de la Ecuación 2.2.2), a continuación sustituyendo r_{m-1} por el resultado de despejarlo en la antepenúltima igualdad, y recorriendo todo el camino otra vez desde abajo hasta arriba, obtenemos una igualdad que expresa el máximo común divisor r_m en función de a y b. Se obtiene así una identidad de Bézout. El siguiente resultado proporciona un método eficiente para hallar el máximo común divisor de dos enteros no nulos y una identidad de Bézout entre ellos.

Teorema 2.8 (Algoritmo extendido de Euclides) *Consideramos las secuencias* $\{x_i\}_{i=0}^k$ *e* $\{y_i\}_{i=0}^k$ *definidas por:*

$$\begin{cases} x_0 = 1, x_1 = 0 \\ y_0 = 0, y_1 = 1 \end{cases} \tag{2.7}$$

y considerando $r_0 = a, r_1 = b$ *como* $r_{i-1} = q_i r_i + r_{i+1}$ *tenemos:*

$$\begin{cases} x_{i+1} = x_{i-1} - x_i q_i \\ y_{i+1} = y_{i-1} - y_i q_i \end{cases} \tag{2.8}$$

para $i = 1, 2, \cdots, m-1$. *Entonces, se verifica que*

$$r_i = x_i r_0 + y_i r_1 \quad i = 0, 1, \cdots, m \tag{2.9}$$

En particular, se tiene la identidad de Bézout

$$mcd(a,b) = r_m = x_m r_0 + y_m r_1 = x_m a + y_m b \tag{2.10}$$

Proof 2.1 *Lo probamos por inducción en* i :

- *Es obvio que* $r_0 = 1 \times r_0 + 0 \times r_1 = x_0 r_0 + y_0 r_1$ *y que* $r_1 = 0 \times r_0 + 1 \times r_1 = x_1 r_0 + y_1 r_1$
- *Para* $i \geq 2$, *suponemos cierto el resultado para los índices menores que* i *y lo probamos para i. Por la hipótesis de inducción, se verifica que*

$$\begin{aligned} x_i r_0 + y_i r_1 &= (x_{i-2} - x_{i-1} q_{i-1}) r_0 + (y_{i-2} - y_{i-1} q_{i-1}) r_1 = \\ &= (x_{i-2} r_0 + y_{i-2} r_1) - (x_{i-1} r_0 + y_{i-1} r_1) q_{i-1} = \\ &= r_{i-2} - r_{i-1} q_{i-1} = r_i \end{aligned}$$

Ejemplo 2.6 *Si tomamos los números b=249 y a=36, a partir de la Tabla 2.1.*

Se tiene que mcd(36,249)=3 y la identidad de Bézout es $3 = 249 \times (-1) + 36 \times 7$

[1] Vídeo explicativo del método extendido de la identidad de Bézout está disponible en https://bit.ly/2XpOTtu.

i	r_i	q_i	x_i	y_i
0	249		1	0
1	36	6	0	1
2	33	1	1	-6
3	3	11	-1	7
	0			

Tabla 2.1: Método extendido de la identidad de Bézout[1]

2.2.3. Función de Euler

Aunque se va a describir con más detalle en el siguiente capítulo, para definir la función de Euler vamos a definir el conjunto $\mathbb{Z}_n$ como el conjunto de enteros positivos menores a n, es decir, $\mathbb{Z}_n = \{0, 1, \cdots, n-1\}$.

El número de enteros de $\mathbb{Z}_n$ que pueden ser primos con n viene dada por la **función de Euler** $\phi(n)$. Dada la factorización en número primos de N, es fácil calcular el valor de $\phi(N)$. Si n tiene la factorización:

$$n = \prod_{i=1}^{k} p_i^{e_i} \tag{2.11}$$

entonces, la función de Euler viene dada por:

$$\phi(n) = \prod_{i=1}^{k} (p_i^{e_i} - p_i^{e_{i-1}}) = \prod_{i=1}^{k} p_i^{e_i}(p_i - 1) = n \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right) \tag{2.12}$$

Nótese que la siguiente afirmación es muy importante en criptografía: *Dada la factorización de n es fácil calcular el valor de* $\phi(n)$. Los casos más importantes en criptografía son:

1. Si p es primo, entonces $\phi(p) = p - 1$.
2. Si p y q son ambos primos y $p \neq q$ entonces $\phi(p \cdot q) = (p-1)(q-1)$

Ejemplo 2.7

$$\phi(60) = \phi(2^2 \cdot 3 \cdot 5) = 60 \cdot \left(1 - \frac{1}{2}\right) \cdot \left(1 - \frac{1}{3}\right) \cdot \left(1 - \frac{1}{5}\right) = 16$$

2.2.4. Mínimo Común Múltiplo

Teorema 2.9 (Mínimo Común Múltiplo) *Dados dos enteros a y b, se denomina mínimo común múltiplo de a y b y se denota por mcm(a,b) al menor de sus múltiplos*

comunes positivos, que existe y es único.

Teorema 2.10 *Sean a y b dos enteros positivos y sean d y m su mcd y su mcm respectivamente, se verifica entonces que*

$$dm = ab \tag{2.13}$$

En consecuencia, también se puede obtener el mínimo común múltiplo a partir del algoritmo de euclides.

Ejemplo 2.8 *Como d=mcd(21,12)=3, se tiene que*

$$m = mcm(21,12) = \frac{21 \cdot 12}{mcd(21,12)} = \frac{21 \cdot 12}{3} = 84$$

2.3. Introducción a la Aritmética Modular

La aritmética modular, es decir, la aritmética de las clases de congruencias, se encarga de simplificar problemas teóricos-numéricos sustituyendo cada entero por el resto de dividirlo entre un entero positivo fijo *n*. Esto produce el efecto de sustituir el conjunto infinito $\mathbb{Z}$ de número enteros por el conjunto $\mathbb{Z}_n$ que solo contiene *n* elementos.

2.3.1. Números Congruentes

Muchos problemas en los que se requieren enteros muy grandes pueden simplificarse con un técnica denominada **aritmética modular**, en la que se utilizan congruencias en vez de ecuaciones.

La idea fundamental de la aritmética modular es muy simple y se asemeja a la ”aritmética del reloj". Por ejemplo, para pasar del sistema de 24 hora al sistema de 12 horas, basta con restar 12 a la hora en formato de 24 horas. Más concretamente, las 13 en el formato 24 horas es la 1 en el sistema de 12 horas, ya que 13 módulo 12 es 1. Es decir, la idea básica es elegir un determinado entero *n* (dependiendo del problema), llamado *módulo* y sustituir cualquier entero por el resto de su división entre *n*. En general, los restos son pequeños y, por tanto, es fácil trabajar con ellos. Antes de entrar en la teoría general, veamos dos ejemplos sencillos.

Ejemplo 2.9 *Si contamos 100 días a partir de hoy, ¿en qué día de la semana caerá? Podemos resolver esta cuestión cogiendo un calendario y contando 100 días, pero un método más sencillo es utilizar el hecho de que los días de la semana se repiten en*

ciclos de 7. Como 100=14 × 7 + 2, dentro de 100 días será el mismo día de la semana que dentro de dos días y ésto es fácil de determinar. Aquí hemos tomado $n = 7$, y hemos reemplazado 100 por el resto de su división entre 7, es decir, por 2.

Ejemplo 2.10 *¿Es 22051946 un cuadrado perfecto? Esto se puede resolver calculando $\sqrt{22051946}$ y viendo si se obtiene un número entero, o alternativamente, elevando al cuadrado varios enteros y ver si puede obtenerse 22051946, pero es mucho más sencillo ver que este número no puede ser un cuadrado perfecto. Sabiendo que un cuadrado perfecto debe dar de resto 0 ó 1 cuando se divide por 4. Para trabajar sólo con dos dígitos podemos ver que:*

$$22051946 = 220519 \times 100 + 46 = 220519 \times 25 \times 4 + 46$$

nos da el mismo resto que 46, y como 46=$11 \times 4 + 2$, el resto es 2. Se sigue de ahí que 22051946 no es un cuadrado perfecto.

Formalmente,

Definición 2.4 (Módulo) *Sean a y b dos enteros y n un entero positivo, se define como módulo a la operación por la que se obtiene del resto positivo de dividir a entre b y se denota por $a \ (mod\ b) = n$.*

Ejemplo 2.11

$$18 \ (mod\ 7) = 4$$
$$-18 \ (mod\ 7) = 3$$

Como se ha visto en los ejemplos anteriores, en algunas situaciones sólo interesan los restos de las divisiones por enteros. Por ello, existen notaciones especiales para ellos. En concreto, existe una notación para indicar que dos enteros tienen el mismo resto cuando se dividen por el entero positivo *n*.

Definición 2.5 (Números congruentes) *Si a y b son enteros y n es un entero positivo, entonces a es congruente con b módulo n si n divide a $a - b$, o lo que es lo mismo, si a y b dan el mismo resto cuando se dividen entre n. Usaremos la notación $a \equiv b \ (mod\ \ n)$ para indicar que a es congruente con b módulo n. Si a y b no son congruentes módulo n, escribiremos $a \not\equiv b \ (mod\ \ n)$. De este modo, lo anterior se puede expresar como:*

$$\left.\begin{array}{ll} a = q \cdot n + r & 0 \leq r < n \\ b = q' \cdot n + r' & 0 \leq r' < n \end{array}\right\} a \equiv b \ \ (mod\ \ n) \Leftrightarrow r = r' \qquad (2.14)$$

Por conveniencia, definiremos el conjunto de posibles restos módulo n como :

$$\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}_n = \{0, \cdots, n-1\} \tag{2.15}$$

Ejemplo 2.12 *Determina si 17 es congruente con 5 módulo 6 y si 24 y 14 son congruentes módulo 6:*

Como 6 divide a 17-5=12, vemos que $17 \equiv 5 \ (mod \ 6)$. *No obstante, como 24-14=10, que no es divisible por 6, vemos que* $24 \not\equiv 14 \ (mod \ 6)$

Existen resultados muy importantes con los números primos y la operación módulo como:

Teorema 2.11 (Teorema Pequeño de Fermat) *Sea p un número primo y* $a \in \mathbb{Z}$, *entonces*

$$a^p \equiv a \ (mod \ p) \tag{2.16}$$

o lo que es lo mismo

$$a^{p-1} \equiv 1 \ (mod \ p) \tag{2.17}$$

Ejemplo 2.13 *Vamos a probar que* $a^{25} - a$ *es divisible entre 30,cualquiera que sea el entero a.*

Factorizando $30 = 2 \cdot 3 \cdot 5$, *vemos que es suficiente probar que* $a^{25} - a$ *es divisible por los primos 2, 3 y 5.*

Para $p = 5$, *tenemos:*

$$a^{25} = (a^5)^5 \equiv a^5 \equiv a \ (mod \ 5)$$

Por el teorema de Fermat, 5 divide a $a^{25} - a$ *para cualquier entero a.*

Análogamente, para $p = 3$

$$a^{25} = (a^3)^8 a \equiv a^8 a = a^9 = (a^3)^3 \equiv a^3 \equiv a \ (mod \ 3)$$

Es decir,3 divide a $a^{25} - a$ *para cualquier entero a.*

Por último, para $p = 2$,

$$a^{25} = (a^2)^{12} a \equiv a^{12} a = (a^2)^6 a \equiv a^6 a = (a^2)^3 a \equiv a^3 a = a^4 = (a^2)^2 \equiv a^2 \equiv a \ (mod \ 2)$$

Teorema 2.12 (Teorema Euler) *Si a y n son primos entre sí, se tiene que*

$$a^{\phi(n)} \equiv 1 \ (mod \ n) \tag{2.18}$$

Ejemplo 2.14 *Encontrar el resto de la división de* 2^{65} *entre 60.*

Dado que 23 es primo con 60, y $\phi(60) = 16$*, se tiene que* $23^{16} \equiv 1 \ (mod\ 60)$*. Por otra parte,*

$$23^{65} = 23^{16*4+1} = (23^{16})^4 \cdot 23 \equiv 23 = 23 \ (mod\ 60)$$

Teorema 2.13 *Sea n un entero positivo. Los enteros a y b son congruentes módulo n, si y sólo si, existe un entero k tal que* $a = b + kn$

Proof 2.2 *Si* $a \equiv b \ (mod\ n)$*, entonces* $n|(a-b)$*. Esto significa que hay un entero k tal que* $a - b = kn$*, por lo que* $a = b + kn$*. Recíprocamente, si hay un entero k tal que* $a = b + kn$*, entonces* $kn = a - b$*. Así, n divide a* $a - b$*, por lo que* $a \equiv b \ (mod\ n)$

2.3.2. Clases de Congruencias

A partir de la definición de congruencia se definen las siguientes relaciones de equivalencia:

Lema 2.3 (Relación de Equivalencia) *Para cualquier entero fijo* $n \geq 1$ *se verifican las propiedades:*

- *Reflexiva:* $a \equiv a \ \ (mod \ \ n)$ *para cualquier entero.*
- *Simétrica:* $a \equiv b \ (mod \ \ n) \Rightarrow b \equiv a \ (mod \ \ n)$.
- *Transitiva: si* $a \equiv b \ (mod \ \ n)$ *y* $b \equiv c \ (mod \ \ n)$ *, entonces* $a \equiv c \ (mod \ \ n)$.

Proof 2.3 *En efecto,*

- $n|(a-a) \ \ \forall a \in \mathbb{Z} \Leftrightarrow a \equiv a \ \ (mod \ \ b) \ \ \forall a \in \mathbb{Z}$
- $n|(a-b) \Rightarrow n|(b-a) \Leftrightarrow a \equiv b \ \ (mod \ \ n) \Rightarrow b \equiv a (mod \ \ n)$
- $\left.\begin{array}{l} n|(a-b) \\ n|(b-c) \end{array}\right\} \Rightarrow n|[(a-b)+(b-c)] = a-c \Leftrightarrow$

$$\Leftrightarrow \left.\begin{array}{l} a \equiv b \ (mod\ n) \\ b \equiv c \ (mod\ n) \end{array}\right\} \Rightarrow a \equiv c \ \ (mod \ \ n)$$

Estas tres propiedades definen una relación de equivalencia, por lo que para cada entero n, la congruencia módulo n es una relación de equivalencia en $\mathbb{Z}$. Al igual que una relación de orden "ordena un conjunto", una relación de equivalencia lo divide en subconjuntos disjuntos denominados **clases de equivalencia** de tal forma que cada una de las clases contiene a todos los elementos que están relacionados entre sí. La clase de equivalencia que comprende todos los enteros congruentes con un entero a módulo n se llama **clase de congruencia** de a módulo n. Formalmente,

Definición 2.6 (Clase de congruencia de a módulo n) *Cada elemento $a \in \mathbb{Z}$ define la clase de equivalencia:*

$$\begin{array}{rcl}[a] & = & \{x \in \mathbb{Z} : x \equiv a(mod\ n)\} = \\ & = & \{\cdots, a-2n, a-n, a, a+n, a+2n, \cdots\}\end{array} \tag{2.19}$$

quedando $\mathbb{Z}$ dividido en las n clases de equivalencia correspondiente a los n posibles restos de dividir un entero entre n:

$$[0],[1],[2],\cdots[n-1] \ \ ya\ que\ \ a \equiv b\ (mod\ n) \Leftrightarrow [a] = [b] \tag{2.20}$$

2.4. Aritmética en $\mathbb{Z}_n$

2.4.1. Suma y Multiplicación en $\mathbb{Z}_n$

Si a y b son elementos de $\mathbb{Z}_n$ (es decir, representan a las clases de congruencias $[a]$ y $[b]$ módulo n respectivamente), definimos su suma, diferencia y producto como las clases:

$$\begin{array}{l}[a]+[b] = [a+b] \\ [a]-[b] = [a-b] \\ [a][b] = [ab]\end{array}$$

Estas operaciones están bien definidas, en el sentido de que los resultados que se obtienen dependen de las clases $[a]$ y $[b]$, y no de los elementos a y b en particular que se hayan tomado como representantes de la clase. Es decir, debemos probar que si trabajamos con módulo n:

$$\left.\begin{array}{rl}[a] & = [a'] \\ [b] & = [b']\end{array}\right\} \Rightarrow \left\{\begin{array}{rcl}[a+b] & = & [a'+b'] \\ [a-b] & = & [a'-b'] \\ [ab] & = & [a'b']\end{array}\right.$$

Ejemplo 2.15 *Consideremos el conjunto $\mathbb{Z}_2 = 0,1$, donde $0 = [0]_2$ representa a cualquier entero par y $1 = [1]_2$ a cualquier número impar. Se puede decir que la suma de dos números impares es par, pero independientemente de los números impares que se hayan sumado ya que*

$$1+1 = [1]+[1] = [1+1] = [2] = [0] = 0$$

Teorema 2.14 *Sean n un entero positivo. Si $a \equiv b\ (mod\ n)$ y $c \equiv d\ (mod\ n)$, entonces:*

$$a+c \equiv c+d\ (mod\ \ n)\ y\ a \cdot c \equiv b \cdot d\ (mod\ \ n) \tag{2.21}$$

Proof 2.4 *Como $a \equiv b \pmod n$ y $c \equiv d \pmod n$, existirán dos enteros s y t tales que $b = a + s \cdot n$ y $d = c + t \cdot n$. Por tanto,*

$$b + d = (a + s \cdot n) + (c + t \cdot n) = (a + c) + (s + t) \cdot n$$

y

$$b \cdot d = (a + s \cdot n) \cdot (c + t \cdot n) = a \cdot c + n \cdot (a \cdot t + c \cdot s + s \cdot t \cdot n)$$

Así,

$$a + c \equiv c + d \ (mod \ \ n) \ y \ a \cdot c \equiv b \cdot d \ (mod \ \ n)$$

Se deduce de aquí que la suma y el producto de pares de clase de congruencias en $\mathbb{Z}_n$ están bien definidas. Por otra parte, la potencia de clases de congruencias está bien definida, pero no su exponenciación. En efecto, como $[a][b] = [ab]$, se deduce que $[a]^k = [a^k]$. Sin embargo, $[a]^{[b]} \neq [a^b]$. Como contraejemplo encontramos que, en $\mathbb{Z}_3$, como $[1] = [4], [2]^{[4]} = [2^4] = [16] = [1] \neq [2] = [2]^{[1]}$!!

Ejemplo 2.16 *Como $7 \equiv 2 \ \ (mod \ \ 5)$ y $11 \equiv 1 \ \ (mod \ \ 5)$, por el teorema anterior se sigue que:*

$$18 = 11 + 7 \equiv 2 + 1 = 3 \ (mod \ 5)$$

y que

$$77 = 7 \cdot 11 \equiv 2 \cdot 1 = 2 \ (mod \ 5)$$

Ejemplo 2.17 *Calculemos resto de la división de 28×33 entre 35.*

$$\left.\begin{array}{l} 28 \equiv -7 \ \ (mod \ \ 35) \\ 33 \equiv -2 \ \ (mod \ \ 35) \end{array}\right\} \Rightarrow 28 \times 33 \equiv (-7) \times (-2) \equiv 14 \ \ (mod \ \ 35)$$

Ejemplo 2.18 *(11+13) (mod 16) =24 (mod 16) = 8, ya que 24=1· 16 +8*

Ejemplo 2.19 *(11· 13) (mod 16) = 143 (mod 16) = 15, ya que 143=8· 16 +15*

La suma y multiplicación módulo *n* funciona igual que la aritmética sobre los reales o los enteros. En particular, tiene las siguientes propiedades:

Lema 2.4 *Propiedades de la suma y multiplicación:*

1. *La suma es cerrada: $\forall a, b \in \mathbb{Z}_n : a + b \in \mathbb{Z}_n$*
2. *La suma es asociativa: $\forall a, b.c \in \mathbb{Z}_n : (a + b) + c = a + (b + c)$*

3. *0 es la suma identidad:* $\forall a \in \mathbb{Z}_n : a+0 = 0+a = a$
4. *La suma inversa siempre existe:* $\forall a \in \mathbb{Z}_n : a+(N-a) = (N-a)+a = 0$
5. *La suma es conmutativa:* $\forall a,b \in \mathbb{Z}_n : a+b = b+a$
6. *La multiplicación es cerrada:* $\forall a,b \in \mathbb{Z}_n : a \cdot b \in \mathbb{Z}_n$
7. *La multiplicación es asociativa:* $\forall a,b.c \in \mathbb{Z}_n : (a \cdot b) \cdot c = a \cdot (b \cdot c)$
8. *1 es la multiplicación identidad:* $\forall a \in \mathbb{Z}_n : a \cdot 1 = 1 \cdot a = a$
9. *La suma y la multiplicación satisfacen la propiedad distributiva:* $\forall a,b,c \in \mathbb{Z}_n : (a+b) \cdot c = a \cdot c + b \cdot c$
10. *La multiplicación es conmutativa:* $\forall a,b \in \mathbb{Z}_n : a \cdot b = b \cdot a$

2.4.2. División en $\mathbb{Z}_n$

En esta sección se va a tratar de resolver el problema de la división entre clases de congruencias $[a]/[b]$, con $[a],[b] \in \mathbb{Z}_n$. O lo que es los mismo, es importante saber cuando, dados a y b, la ecuación

$$a \cdot x \equiv b \ (mod\ n) \tag{2.22}$$

tiene solución. Dicha ecuación se conoce como **congruencia lineal**.

Ejemplo 2.20 *Para la ecuación* $7x \equiv 3 \ (mod\ 143)$ *hay sólo una solución. Sin embargo, no hay soluciones para la ecuación* $11x \equiv 3 \ (mod\ 143)$. *Por otra parte, hay 11 soluciones para la ecuación* $11x \equiv 22 \ (mod\ 143)$

Afortunadamente es muy fácil saber cuando una ecuación tiene una, muchas o ninguna solución a partir del máximo común divisor de a y n ($mcd(a,n)$). En efecto:

- Si mcd(a,n)=1, entonces sólo hay una solución. Esta solución se calculará como $x \equiv b \cdot c \ (mod\ n)$, donde $a \cdot c \equiv 1 \ (mod\ n)$. En este caso, diremos que a y n son primos.
- Si $g = mcd(a,n) \neq 1$ y $mcd(a,n)$ divide b, entonces hay g soluciones.
- En otro caso no hay soluciones.

Formalmente, el siguiente resultado caracteriza cuándo una congruencia lineal tiene solución:

Teorema 2.15 *Sea* $a,b,n \in \mathbb{Z}$, *con n positivo, y sea* $d = mcd(a,m)$. *Entonces la congruencia lineal* $ax \equiv b \ (mod\ n)$ *tiene solución si, y sólo si, d divide a b y, en ese caso, el número de soluciones en* $\mathbb{Z}_n$ *es d*

Por otra parte, para calcular todas las soluciones de una congruencia lineal se tiene el siguiente resultado:

Lema 2.5 *Sea $ax \equiv b \ (mod\ n)$ una congruencia lineal. Si α es una solución de la misma, entonces todo $\beta \equiv \alpha \ (mod\ n)$ es también una solución de la congruencia.*

Proof 2.5 *Como α es solución de la congruencia, existe $\lambda \in \mathbb{Z}$ tal que*

$$a\alpha - b = \lambda n$$

Y como $\beta \equiv \alpha \ (mod\ n)$, existe $\mu \in \mathbb{Z}$ tal que

$$\beta = \alpha + \mu n$$

Por lo tanto,

$$a\beta - b = (\lambda + a\mu)n$$

En consecuencia,$a\beta \equiv b \ (mod\ n)$, por lo que β es una solución de la congruencia.

2.4.3. Congruencias lineales: Algoritmo de Euclides Extendido

Usando el algoritmo de Euclides se puede determinar cuando a tiene una inversa modulo n, y así resolver la congruencia lineal $ax \equiv b \ (mod\ n)$. Sin embargo, en muchos problemas de criptografía no solo interesará saber cuando tiene solución, sino qué solución tiene. Esto se puede hacer con el algoritmo de Euclides Extendido.

Partiendo de

$$r_{i-2} = q_{i-1}r_{i-1} + r_i \tag{2.23}$$

con $r_m = mcd(r_0, r_1)$. Si escribimos los términos $r_i, i \geq 2$ en términos de a y b:

$$\begin{aligned}
r_2 &= r_0 - q_1 r_1 = a - q_1 b \\
r_3 &= r_1 - q_2 r_2 = b - q_2(a - q_1 b) = -q_2 + (1 + q_1 q_2)b \\
&\vdots \quad\quad \vdots \\
r_{i-2} &= x_{i-2}a + y_{i-2}b \\
r_{i-1} &= x_{i-1}a + y_{i-1}b \\
r_i &= r_{i-2} - q_{i-1}r_{i-1} = a(x_{i-2} - q_{i-1}x_{i-1}) + b(y_{i-2} - q_{i-1}y_{i-1}) \\
&\vdots \quad\quad \vdots \\
r_m &= x_m a + y_m b
\end{aligned}$$

El algoritmo de Euclides Extendido toma como entrada a y b y devuelve r_m, x_m y y_m, de manera que

$$r_m = x_m a + y_m b \tag{2.24}$$

De aquí, podemos resolver nuestro problema original de determinar la inversas de a modulo n, cuando esa inversa existe. Para ello, primero se aplica el algoritmo extendido de Euclides a a y n para calcular d, x, y de manera que:

$$d = mcd(a,n) = ax + ny \tag{2.25}$$

Podemos resolver la ecuación $ax = 1 \ (mod \ n)$, ya que tenemos $d = xa + yn = xa \ (mod \ n)$. Por tanto, encontraremos la solución $x = a^{-1}$ cuando d=1.

Ejemplo 2.21 *Supongamos que queremos calcular la inversa de 7 módulo 19, Primero hacemos $r_0 = 19$ y $r_1 = 7$. A continuación calculamos:*

$$\begin{aligned} r_2 &= 5 = 19 - 2 \cdot 7 \\ r_3 &= 2 = 7 - 5 = 7 - (19 - 2 \cdot 7) = -19 + 3 \cdot 7 \\ r4 &= 1 = 5 - 2 \cdot 2 = (19 - 2 \cdot 7) - 2(-19 + 3 \cdot 7) = 3 \cdot 19 - 8 \cdot 7 \end{aligned}$$

Por tanto,

$$1 = -19 + 3 \cdot 7 \ (mod \ 19)$$

luego

$$7^{-1} = -8 = 11 \ (mod \ 19)$$

Ejemplo 2.22 *Consideramos la congruencia $10x \equiv 3 \ (mod \ 12)$. Aquí $a = 10, b = 3$ y $n = 12$. Como mcd(10,12)=2 y no divide a b=3, no hay soluciones.*

Ejemplo 2.23 *Si ahora consideramos la congruencia $7x \equiv 3 \ (mod \ 12)$. Aquí $a = 7, b = 3$ y $n = 12$. Como mcd(7,12)=1, sólo hay una solución $x = 3 \cdot c \ (mod \ 12)$, donde $7 \cdot c = 1 \ (mod \ 12)$. Para hallar c utilizamos el algoritmo de Euclides extendido:*

12		1	0
7	1	0	1
5	1	1	-1
2	2	-1	2
1	2	3	-5
0			

Como -5 (mod 12) = 7 , se tiene que $c = 7$, entonces $x = 3 \cdot 7 \ (mod \ 12) = 9$.

Ejemplo 2.24 *Por otra parte, si consideramos la congruencia* $10x \equiv 6 \ (mod\ 12)$. *Aquí* $a = 10, b = 6$ *y* $n = 12$. *Como mcd(10,12)=2 y 2 a b=6, hay 2 soluciones. Ahora, para calcularlas, consideramos la congruencia lineal que resulta de dividir todos los coeficientes por* $d = 2$*:* $5x \equiv 3 (mod\ 6)$. *En este caso, como mcd(5,6)=1, sólo habrá una solución que calculamos como* $x = b \cdot c \ (mod\ n)$, *donde* $a \cdot c = 1 \ (mod\ n)$. *Empezamos calculando c con el algoritmo de Euclides extendido:*

6		1	0
5	1	0	1
1	5	1	-1

Como $c = -1 \ (mod\ 6) = 5$, *se tiene que* $x_0 = 3 \cdot 5 \ (mod\ 6) = 3$. *El resto de soluciones se calcularán como* $x_i = x_0 + n \cdot i$. *Por tanto,* $x_1 = 3 \cdot 5 \ (mod\ 6) + 6 = 3 + 6 = 9$.

2.4.4. Módulo inverso multiplicativo *n*

Hemos visto que cuando queremos resolver una ecuación de la forma

$$ax = b \ (mod\ n) \tag{2.26}$$

se reduce el problema a examinar cuando un entero a módulo n tiene un inversa multiplicativa, es decir, cuando hay un número c tal que:

$$ac = ca = 1 \ (mod\ n) \tag{2.27}$$

Ese valor de c se escribe como a^{-1}. Claramente a^{-1} es la solución de la ecuación:

$$ax = 1 \ (mod\ n) \tag{2.28}$$

Por tanto, la inversa de a sólo existe cuando a y n son primos entre sí, es decir, cuando $mcd(a,n) = 1$.

En la práctica, para encontrar a^{-1} basta con encontrar una identidad de Bézout de la forma

$$1 = ax + ny \tag{2.29}$$

y tomar como a^{-1} el representante de x en $\mathbb{Z}_n$ (es decir, el resto de dividir x entre n. Por tanto, a partir del teorema extendido de euclides se puede obtener el inverso de un número $a^{-1} = y_m$.

Ejemplo 2.25 *Calculemos* 7^{-1} *en* $\mathbb{Z}_{37}$. *A partir de la tabla:*

Por tanto $7^{-1} = 16$ *en* $\mathbb{Z}_{37}$.

i	r_i	q_i	x_i	y_i
0	37		1	0
1	7	5	0	1
2	2	3	1	-5
3	1	2	-3	16
	0			

Un caso particular es cuando n es un primo p, ya que para todo valor distinto de cero en $a \in \mathbb{Z}_p$ siempre se obtiene una única solución a.

$$ax = 1 \ (mod \ p) \tag{2.30}$$

De esta afirmación se puede deducir que si p es primo, entonces cualquier elemento distinto de 0 en $\mathbb{Z}_p$ tiene una multiplicativa inversa. Formalmente,

Teorema 2.16 *Un elemento $r \in \mathbb{Z}_n$ es inversible si y sólo si, r y n son primos entre si, es decir, si $mcd(n,r) = 1$*

Unidades de $\mathbb{Z}_n$

Definición 2.7 (Unidades de $\mathbb{Z}_n$) *Un elemento r de $\mathbb{Z}_n$, decimos que es una unidad si es inversible, es decir, si existe otro elemento $s \in \mathbb{Z}_n$ tal que $sr = rs = 1$*

Teorema 2.17 *El inverso de un elemento unidad es único.*

Proof 2.6 *Supongamos que existen dos elementos inversos de r, s y s' y probemos que $s = s'$. En efecto,*

$$s = s \cdot 1 = s(rs') = (sr)s' = 1 \cdot s' = s'$$

Ejemplo 2.26 *Las unidades de $\mathbb{Z}_8$ son 1,3,5 y 7. En efecto, $1 \cdot 1 = 1$, $3 \cdot 3 = 1$, $5 \cdot 5 = 1$, $7 \cdot 7 = 1$, por lo que cada una de estas unidades es su propio inverso.*

En $\mathbb{Z}_9$ son 1,2,4,5,7 y 8. En efecto, $1 \cdot 1 = 1$, $2 \cdot 5 = 1$, $4 \cdot 7 = 1$, $8 \cdot 8 = 1$, por lo que cada una de estas unidades es su propio inverso.

2.4.5. Sistema de Congruencias Lineales: Teorema Chino del resto

Un sistema de congruencias lineales es un sistema de la forma:

$$\begin{aligned} a_1x &= b_1 \ (mod\ n_1) \\ a_2x &= b_2 \ (mod\ n_2) \\ &\vdots \\ a_kx &= b_k \ (mod\ n_k) \end{aligned} \tag{2.31}$$

Es decir, se trata de un sistema de k ecuaciones pero con una sola incógnita.

Para que el sistema tenga solución, deberán tenerla cada una de las ecuaciones del sistema. Del mismo modo que en el caso de las congruencias lineales, antes de resolver el sistema habrá que estudiar si cada ecuación tiene solución. Una vez comprobado, tratar de ver si existe alguna solución común a todas las ecuaciones, es decir, tratar de ver si el sistema tiene solución.

Si conseguimos eliminar los coeficientes a_i, todas las congruencias tendrán una única solución ya que $mcd(n_i, 1) = 1$. Una vez eliminados todos los coeficientes a_i, lo único que sabemos es que todas las ecuaciones tienen solución, pero desconocemos si existe alguna solución común a todas ellas.

La primera parte, eliminar los coeficientes a_i, ya sabemos hacerlo, se trata de resolver las k congruencias lineales, por lo que el estudio de la existencia de soluciones de un sistema lo haremos sobre un sistema que ya tiene resueltas todas sus congruencias, es decir, partiremos de un sistema de la forma

$$\begin{aligned} x &= a_1 \ (mod\ n_1) \\ x &= a_2 \ (mod\ n_2) \\ &\vdots \\ x &= a_k \ (mod\ n_k) \end{aligned} \tag{2.32}$$

Una forma sencilla de resolver esta ecuación es con el *Teorema Chino del Resto*. Este teorema data de hace más de 2000 años y actualmente se usa para mejorar protocolos de criptografía.

Empezamos suponiendo dos ecuaciones

Teorema 2.18 (Teorema Chino del Resto: dos ecuaciones) *Sean dos ecuaciones*

$$\begin{aligned} x &= a \ (mod\ n) \\ y &= b \ (mod\ m) \end{aligned} \tag{2.33}$$

entonces hay una única solución módulo $m \cdot n$, si y sólo si $mcd(n,m) = 1$

Ejemplo 2.27 *Dadas*

$$x = 4 \ (mod\ 7)$$
$$x = 3 \ (mod\ 5)$$

Como $mcd(5,7) = 1$, tenemos una única solución:

$$x = 18 \ (mod\ 35) = 18$$

Es fácil comprobar que es una solución puesto que 18 (mod 7) = 4 y 18 (mod 5) =3. Para encontrar esta solución, partimos de las ecuaciones:

$$x = 4 \ (mod\ 7)$$
$$x = 3 \ (mod\ 5)$$

De aquí, para algún u se tiene que :

$$x = 4 + 7u$$
$$x = 3 \ (mod\ 5)$$

Sustituyendo la primera ecuación en la segunda se tiene que:

$$\begin{aligned} 4+7u &= 3 \ (mod\ 5) \Leftrightarrow 7u = 3-4 \ (mod\ 5) \Leftrightarrow \\ 7u &= -1 \ (mod\ 5) \Leftrightarrow 7u = 4 \ (mod\ 5) \Leftrightarrow \\ 2u+5u &= 4 \ (mod\ 5) \Leftrightarrow 2u = 4 \ (mod\ 5) \end{aligned}$$

Como mcd(2,5)=mcd(7,5)=1, se puede resolver la ecuación anterior para u. Para ello, primero calculamos $2^{-1} \ (mod\ 5) = 3$. En efecto,

$$2 \cdot 3 = 6 = 1 \ (mod\ 5)$$

Por tanto,

$$2u = 4 \ (mod\ 5) \Leftrightarrow 2u \cdot 3 \equiv 4 \cdot 3 \ (mod\ 5) \Leftrightarrow u = 12 \ (mod\ 5) = \Leftrightarrow u = 2 \ (mod\ 5) = 2$$

Así, sustituyendo u en nuestra ecuación se obtiene el valor de x

$$x = 4 + 7u = 4 + 7 \cdot 2 = 18$$

Se puede generalizar el proceso de resolución del sistema con el siguiente algoritmo:

1. Calcular $t = m^{-1} \ (mod\ n)$
2. Calcular $u = (b-a) \cdot t \ (mod\ m)$
3. $x = a + u \cdot n$

Ejemplo 2.28 *En el caso anterior, podríamos haber resuelto el sistema haciendo*

$$\begin{aligned} t &= 5^{-1} \ (mod\ 7) = 3 \\ u &= (3-4) \cdot 3 \ (mod 5) = 2 \\ x &= 4 + 2 \cdot 7 = 18 \end{aligned}$$

Si lo hiciéramos al revés, tendríamos

$$\begin{aligned} t &= 7^{-1} \ (mod\ 5) = 3 \\ u &= (4-3) \cdot 3 \ (mod 7) = 3 \\ x &= 3 + 3 \cdot 5 = 18 \end{aligned}$$

Ahora, consideramos el caso general con más de dos ecuaciones.

Teorema 2.19 (Teorema Chino del Resto: Caso general) *Sean k ecuaciones:*

$$\begin{aligned} x &= a_1 \ (mod\ n_1) \\ x &= a_2 \ (mod\ n_2) \\ &\vdots \\ x &= a_k \ (mod\ n_k) \end{aligned} \tag{2.34}$$

sea $N = n_1 \cdot n_2 \cdot \cdots \cdot n_k$, el teorema chino del resto garantiza una única solución dada por

$$x = \sum_{i=1}^{k} a_i \cdot N_i \cdot y_i \ (mod N) \tag{2.35}$$

donde $N_i = N/n_i$ y $y_i = N_i^{-1} \ (mod\ n_i)$ si y sólo si n_i, n_j son primos entre sí $\forall i \neq j$.

Ejemplo 2.29 *Se quiere resolver el sistema:*

$$\begin{aligned} x &= 5 \ (mod\ 7) \\ x &= 3 \ (mod\ 11) \\ x &= 10 \ (mod\ 13) \end{aligned}$$

Se obtiene:

$$\begin{aligned} N &= 7\cdot 11\cdot 13 = 1001 \\ N_1 &= 1001/7 = 143\ ,\ y_1 = 143^{-1}\ (mod\ 7) = 5 \\ N_2 &= 1001/11 = 91\ ,\ y_2 = 91^{-1}\ (mod\ 11) = 4 \\ N_3 &= 1001/13 = 77\ ,\ y_3 = 77^{-1}\ (mod\ 13) = 12 \end{aligned}$$

Por tanto, la solución vendrá dada por:

$$\begin{aligned} x &= \sum_{i=1}^{k} a_i\cdot N_i\cdot y_i\ (modN) = \\ &= 5\cdot 143\cdot 5 + 3\cdot 91\cdot 4 + 10\cdot 77\cdot 12\ (mod\ 1001) = \\ &= 13907\ (mod\ 1001) = 894 \end{aligned}$$

Lema 2.6 *Consideremos la descomposición de n en factores primos $n = p_1^{e_1}\cdot\ldots\cdot p_k^{e_k}$, donde $p1,\cdots,p_k$ son primos diferentes. Para cualesquiera enteros a y b:*

$$a \equiv b\ (mod\ n) \Leftrightarrow a \equiv b\ (mod\ p_i^{e_i}) \quad \forall i = 1,\cdots,k \tag{2.36}$$

Ejemplo 2.30 *Vamos a resolver la congruencia $91x \equiv 459\ (mod\ 440)$:*

Al ser $mcd(91,440) = 1$ existe una única solución. Por otra parte $440 = 2^3\cdot 5\cdot 11$. Luego la congruencia es equivalente al sistema:

$$\begin{cases} 91x &\equiv 459\ (mod\ 2^3) \\ 91x &\equiv 459\ (mod\ 5) \\ 91x &\equiv 459\ (mod\ 11) \end{cases} \Leftrightarrow \begin{cases} 3x &\equiv 3\ (mod\ 8) \\ x &\equiv 4\ (mod\ 5) \\ 3x &\equiv 1\ (mod\ 11) \end{cases} \Leftrightarrow \begin{cases} x &\equiv 1\ (mod\ 8) \\ x &\equiv 4\ (mod\ 5) \\ x &\equiv 4\ (mod\ 11) \end{cases}$$

Para resolver el sistema calculamos:

$$\begin{aligned} N &= 8\cdot 5\cdot 11 = 440 \\ N_1 &= 440/8 = 55\ ,\ y_1 = 55^{-1}\ (mod\ 8) = 7 \\ N_2 &= 440/5 = 88\ ,\ y_2 = 88^{-1}\ (mod\ 5) = 2 \\ N_3 &= 440/11 = 40\ ,\ y_3 = 40^{-1}\ (mod\ 11) = 8 \end{aligned}$$

Por tanto, la solución vendrá dada por:

$$\begin{aligned} x &= \sum_{i=1}^{k} a_i\cdot N_i\cdot y_i\ (modN) = \\ &= 1\cdot 55\cdot 7 + 4\cdot 88\cdot 2 + 4\cdot 40\cdot 8\ (mod\ 440) = \\ &= 2369\ (mod\ 440) = 169 \end{aligned}$$

Si los coeficientes $n_1, \cdots, n_k$ no son primos dos a dos, tenemos el siguiente resultado

Teorema 2.20 (Teorema Chino de los Restos Generalizado) *Consideremos los enteros positivos $n_1, n_2, \cdots, n_k$ y seab $a_1, a_2, \cdots, a_k$ enteros cualquiera. El sistema de congruencias*

$$\begin{aligned} x &= a_1 \ (mod\ n_1) \\ x &= a_2 \ (mod\ n_2) \\ &\vdots \\ x &= a_k \ (mod\ n_k) \end{aligned} \tag{2.37}$$

admite solución si, y sólo si, mcd(n_i, n_j) divide a $(a_i - a_j)$ para cualesquiera $i \neq j$.

Cuando se verifica esta condición, la solución general constituye una única clase de congruencia módulo n, donde n es el mínimo común múltiplo de $n_1, \cdots, n_k$.

El siguiente ejemplo muestra cómo se resolverían este tipo de sistemas:

Ejemplo 2.31 *Consideremos las congruencias*

$$x \equiv 11 \ (mod\ 36), \quad x \equiv 7 \ (mod\ 40), x \equiv 32 \ (mod\ 75)$$

Como

$$\begin{aligned} mcd(36,40) = 4 | (a_1 - a_2) = 4 \\ mcd(35,75 = 3 | (a_1 - a_3) = -21 \\ mcd(40,75) = 5 | (a_2 - a_3) = -25 \end{aligned}$$

El sistema tiene solución. Ahora, transformamos las congruencias:

$$\begin{cases} x \equiv 11 \ (mod\ 2^2 \cdot 3^2) \Leftrightarrow \begin{cases} x \equiv 11 \ (mod\ 2^2) \Leftrightarrow x \equiv 3 \ (mod\ 2^2) \\ x \equiv 11 \ (mod\ 3^2) \Leftrightarrow x \equiv 2 \ (mod\ 3^2) \end{cases} \\ x \equiv 7 \ (mod\ 2^3 \cdot 5) \Leftrightarrow \begin{cases} x \equiv 7 \ (mod\ 2^3) \Leftrightarrow x \equiv 7 \ (mod\ 2^3) \\ x \equiv 7 \ (mod\ 5) \Leftrightarrow x \equiv 2 \ (mod\ 5) \end{cases} \\ x \equiv 32 \ (mod\ 3 \cdot 5^2) \Leftrightarrow \begin{cases} x \equiv 32 \ (mod\ 3) \Leftrightarrow x \equiv 2 \ (mod\ 3) \\ x \equiv 32 \ (mod\ 5^2) \Leftrightarrow x \equiv 7 \ (mod\ 5^2) \end{cases} \end{cases}$$

De este conjunto de seis congruencias en las que los módulos son potencias de los primos 2,3 y 5, seleccionamos las que involucran a la mayor potencia de cada primo para quedarnos con el sistema

$$x \equiv 2 \ (mod\ 9), \quad x \equiv 7 \ (mod\ 8), \quad x \equiv 7 \ (mod\ 25)$$

cuyos módulos son mutuamente primos entre sí, y podemos aplicarle los métodos anteriores, basados en el Teorema Chino del Resto, para encontrar la solución general

$$x \equiv 407 \ (mod\ 1800)$$

2.5. Conjuntos especiales en $\mathbb{Z}_n$

Los conjuntos que cumplen algunas de las propiedades vistas en el Lema 2.4 tienen nombres especiales. Así, definimos:

2.5.1. Grupos

Definición 2.8 (Grupos) *Un grupo es un conjunto con una operación que cumple las siguientes propiedades (ver Lema 2.4):*

1. *Cerrada*
2. *Es asociativa*
3. *Tiene una identidad*
4. *Cada elemento tiene su inversa*

Un grupo que es conmutativo se llama *abeliano*. Casi todos los grupos en criptografía son abelianos. Es decir, un conjunto con las propiedades 1,2 3 y 4 anteriores se llama grupo; mientras que un conjunto con las propiedades 1,2,3,4 y 5 se llama grupo abeliano (ver Lema 2.4).

Ejemplo 2.32 *($(\mathbb{R},+)$ y $(\mathbb{Z},+)$ son conjunto de enteros, reales o complejos respectivamente, bajo la suma es un grupo abeliano. Aquí la identidad es 0 y la inversa de x es $-x$, ya que $x+(-x)=0$*

Ejemplo 2.33 *$(\mathbb{Q}\setminus\{0\},\cdot)$ es el conjunto de los racionales distintos de 0, reales o complejos, bajo la multiplicación es un grupo abeliano. Aquí la identidad es 1 y la inversa de x es x^{-1}, ya que $x\cdot(x^{-1})=1$*

Un grupo se llama *multiplicativo*, y lo denotaremos por $(G,\cdot)$ si se tiende a escribir su operación de grupo de la misma manera que la multiplicación, es decir:

$$f = g\cdot h, \quad g^5 = g\cdot g\cdot g\cdot g\cdot g$$

Un grupo se llama *aditivo*, y lo denotaremos por $(G,+)$ si se tiende a escribir su operación de grupo de la misma manera que la suma, es decir:

$$f = g+h, \quad 5\cdot g = g+g+g+g+g$$

Un grupo se llama *cíclico* si tiene un elemento especial, denominado generador, del que se pueden obtener todos los demás elementos ya sea por la aplicación repetida de la operación del grupo, o por el uso de la operación inversa.

Ejemplo 2.34 *En el grupo $(\mathbb{Z},+)$ cada entero positivo puede obtenerse tras repetir la suma de 1 con sigo mismo. Por ejemplo, 7 se puede expresar como:*

$$7 = 1+1+1+1+1+1+1$$

Y cada número entero negativo puede obtenerse a partir de un entero positivo aplicando el operador de suma inversa. Por tanto, 1 es un generador de enteros bajo suma

Formalmente, si g es un generador del grupo cíclico G, escribiremos $G = \langle g \rangle$. Si G es multiplicativo, entonces cada elemento h de G se escribirá como:

$$h = g^x$$

Mientras que si G es aditivo, cada elemento h de G se escribirá como:

$$h = x\cdot g$$

siendo x un entero denominado logaritmo discreto de h en la base g.

2.5.2. Anillo

Definición 2.9 (Anillo) *Un anillo $(R,\cdot,+)$ es un conjunto con dos operaciones, normalmente denotadas por + y $\cdot$ para la suma y la multiplicación, que satisfacen las propiedades 1 a la 9 del Lema 2.4.*

Un anillo se llama *multiplicativo* si la multiplicación es conmutativa (propiedad 10 del Lema 2.4).

Esto puede parecer complicado, pero en criptografía solo necesitamos considerar anillos finitos, como el anillo conmutativo de números enteros módulo N. Es decir, el anillo $(\mathbb{Z}_n,\cdot,+)$ con multiplicación conmutativa.

2.5.3. Campo

Un anillo como $\mathbb{Z}_p$, con p primo, con esta propiedad se llama **campo**.

Definición 2.10 (Campo) *Un campo es un conjunto con dos operaciones $(G,\cdot,+)$ de manera que:*

- $(G,+)$ *es un grupo abeliano con la identidad denotada por 0.*
- $(G\setminus 0,\cdot)$ *es un grupo abeliano.*
- $(G,\cdot,+)$ *satisface la propiedad distributiva.*

Por tanto, un campo en un anillo conmutativo donde cada elemento distinto de 0 tiene una inversa multiplicativa.

Ejemplo 2.35 *Los números racionales reales o complejos son campos infinitos.*

Ejemplo 2.36 *Si confeccionamos las tablas de la suma y el producto en $\mathbb{Z}_4$*

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

×	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

observamos que $2\cdot 1 = 2\cdot 3$ y esto no implica la igualdad de 1 y 3. Es decir, en $\mathbb{Z}_4$ no se verifica la propiedad cancelativa del producto.

2.5.4. Conjunto de elementos invertibles

A la vista de la tabla del producto en $\mathbb{Z}_4$ nos damos cuenta de que aunque el producto no tiene elemento inverso, existen elementos que sí lo tienen, por ejemplo el 3 ($3\times 3 = 1$, es decir 3 es autoinverso. Cabe entonces preguntarse cuándo va a tener inverso un elemento en $\mathbb{Z}_n$

Definición 2.11 (Conjunto de elementos invertibles) *Definimos el conjunto de elementos invertibles en $\mathbb{Z}_n$ por:*

$$(\mathbb{Z}_n)* = \{x \in \mathbb{Z}_n : mcd(x,n) = 1\} \tag{2.38}$$

*El * en A* para cualquier anillo se refiere al subconjunto más grande de A que forma un grupo bajo multiplicación. Por tanto, el conjunto $(\mathbb{Z}_n)*$ es un grupo con respecto a la multiplicación y tiene tamaño $\phi(n)$.*

En el caso especial cuando n es un primo p tenemos:

$$(\mathbb{Z}_p)* = \{1, \cdots, p-1\}$$

ya que todo elemento no cero de $\mathbb{Z}_p$ es primo con p. Para un campo arbitrario F, el conjunto F* es igual al conjunto $F \setminus \{0\}$. Para facilitar la notación, definiremos:

$$\mathbb{F}_p = \mathbb{Z}_p = \{0, \cdots, p-1\}$$
$$\mathbb{F}*_p = (\mathbb{Z}_p)* = \{1, \cdots, p-1\}$$

El conjunto $\mathbb{F}_p$ es un campo finito de características p. Sin embargo, es importante resaltar que los enteros módulo n son sólo campo cuando n es primo.

Los enteros modulo a primo p no son los únicos campos finitos. Existen otros tipo de campo finito que son muy importante en criptografía. En concreto, en lugar de considerar p número primo se puede considerar p un número primo y el conjunto de polinomios en X cuyos coeficientes son reducibles módulo p. Esto se denota como $\mathbb{F}_p[X]$. Nótese que $\mathbb{F}_p[X]$ es un anillo con las definiciones naturales de suma y multiplicación.

Ejemplo 2.37 *Un caso particular es cuando $p = 2$. Por ejemplo, en $\mathbb{F}_2[X]$ se tiene:*

$$(1+X+X^2)+(X+X^3) = 1+X^2+X^3$$
$$(1+X+X^2)\cdot(X+X^3) = X+X^2+X^4+X^5$$

2.5.5. Criterios de divisibilidad

Teorema 2.21 *Sea P(X) un polinomio con coeficientes enteros y sea $n \geq 1$,*

$$a \equiv b \ (mod\ n) \Rightarrow P(a) \equiv P(b) \ (mod\ n) \tag{2.39}$$

Una aplicación directa del teorema anterior es la obtención de *criterios de divisibilidad*.

Sea $N = a_n a_{n-1} \cdots a_1 a_0$ con $0 \leq a_i \leq 9$ la expresión decimal de un entero $N; P(x) = a_n x^n + \cdots a_1 x + a_0$

Entonces:

- *Criterio de divisibilidad por 3*: Un número es divisible por 3 si, y sólo si, lo es el número formado por la suma de sus cifras.
- *Criterio de divisibilidad por 9*: Un número es divisible por 9 si, y sólo si, lo es el número formado por la suma de sus cifras.
- *Criterio de divisibilidad por 11*: Un número es divisible por 1 si, y sólo si, lo es el número resultante de sumar las cifras que ocupan lugar par y restarle la suma de las cifras que ocupan lugar impar.

Ejemplo 2.38 *Sea* $N = P(10) = a_n 10^n + \cdots a_1 10 + a_0$*:*

- $10 \equiv 1 \ (mod\ 3) \Rightarrow N = P(10) \equiv P(1) = \sum_{i=0}^{n} a_i \ (mod\ 3)$
- $10 \equiv 1 \ (mod\ 9) \Rightarrow N = P(10) \equiv P(1) = \sum_{i=0}^{n} a_i \ (mod\ 9)$
- $10 \equiv -1 \ (mod\ 11) \Rightarrow N = P(10) \equiv P(-1) = a_0 - a_1 + \cdots + (-1)^n a_n \ (mod\ 11)$

Como n divide a m si, y sólo si, $m \equiv 0 \ (mod\ n)$, se sigue que los problemas sobre divisibilidad son equivalentes a los problemas sobre congruencias y, estos últimos son, a veces, más fáciles de resolver. Por ejemplo:

Ejemplo 2.39 *Probar, mediante congruencias, que* $3^{2n+5} + 2^{4n+1}$ *es divisible por 7 cualquiera que sea el entero* $n \geq 1$.

Trabajando módulo 7, se tiene que

$$3^{2n+5} + 2^{4n+1} = 3^5 \cdot 3^{2n} + 2 \cdot 2^{4n} = 243 \cdot 9^n + 2 \cdot 16^n \equiv 5 \cdot 2^n + 2 \cdot 2^n = 7 \cdot 2^n \equiv 0$$

Es decir, 7 divide a $3^{2n+5} + 2^{4n+1}$

Otra aplicación nos permite detectar si un polinomio entero puede tener raíces enteras:

Teorema 2.22 *Sea P(x) un polinomio con coeficientes enteros. Dado que*

$$\left.\begin{array}{c} P(\bar{x}) = 0, \quad \bar{x} \in \mathbb{Z} \\ \bar{x} \equiv a \ (mod\ n) \end{array}\right\} \Leftrightarrow P(a) \equiv 0 \ (mod\ n) \tag{2.40}$$

Si para algún $n > 1$ *se verifica que* $P(a) \not\equiv 0 \ (mod\ n) \quad \forall a \in \mathbb{Z}_n$*, el polinomio carece de raíces enteras.*

Ejemplo 2.40 *Para* $n = 4$ *el polinomio* $P(x) = x^5 - x^2 + x - 3$ *verifica que*

$$\left.\begin{array}{c} P(0) = -3 \equiv 1 \neq 0 \ (mod\ 4) \\ P(1) = -2 \equiv 2 \neq 0 \ (mod\ 4) \\ P(2) = 27 \equiv 3 \neq 0 \ (mod\ 4) \\ P(3) = 234 \equiv 2 \neq 0 \ (mod\ 4) \end{array}\right\} \Leftrightarrow P(x) \ \textit{carece de raíces enteras}$$

Teorema de Lagrange

Teorema 2.23 (Teorema de Lagrange) *Si $(G,\cdot)$ es un grupo de orden (tamaño) k= #G, entonces para todo $a \in G$ se tiene que $a^k = 1$.*

Por tanto, si $x \in (\mathbb{Z}_n)*$ entonces:

$$x^{\phi(n)} = 1 \ (mod\ n) \tag{2.41}$$

ya que $\#(\mathbb{Z}_n) = \phi(n)$

Universidad Rey Juan Carlos

Tema 3

Combinatoria

3.1. Análisis Combinatorio

El análisis combinatorio es la técnica que permite saber cuántos elementos hay en un conjunto sin realmente tener que conocerlos. Contar el número de objetos que verifican ciertas propiedades es una parte importante de la combinatoria, pues se necesita para resolver problemas de muy diversos tipos. Por ejemplo, una contraseña de un sistema informático consiste en seis, siete u ocho caracteres. Cada uno de dichos caracteres debe ser un dígito o una letra del alfabeto. Cada contraseña debe contener al menos un dígito. ¿Cuántas posibles contraseñas existen? En esta tema introduciremos las técnicas necesarias para resolver, entre otras, a esta cuestión.

3.2. Técnicas Básicas

Al hablar de enumeración se quiere *contar* los elementos de un conjunto asignando un número natural a cada uno de ellos. Ahora bien, si no disponemos de una lista de sus elementos, sino que el conjunto viene definido a través de unas propiedades, es necesario desarrollar técnicas diferentes a las ya conocidas, capaces de contar sus elementos. Las técnicas básicas de conteo son el principio de adición o de la suma, el principio de distribución y el principio de multiplicación. Antes de desarrollar estas técnicas necesitamos las siguientes definiciones:

Definición 3.1 *Cardinal: Se define el cardinal de un conjunto, y lo denotaremos por $|A|$ al número de elementos del conjunto finito A.*

Definición 3.2 *Conjunto unión: Dados dos conjuntos A y B, se define el conjunto unión y se denota por $A \cup B$ como el conjunto constituido por todos los elementos que pertenecen a A o a B (o simultáneamente a ambos):*

$$x \in A \cup B \Leftrightarrow \begin{cases} x \in A & \acute{o} \\ x \in B \end{cases}$$

Definición 3.3 *Conjunto intersección: Dados dos conjuntos A y B, se define el conjunto intersección y se denota por $A \cap B$ como el conjunto de los elementos que pertenecen simultáneamente a ambos conjuntos:*

$$x \in A \cap B \Leftrightarrow \begin{cases} x \in A & y \\ x \in B \end{cases}$$

*Si la intersección de dos conjuntos es vacía, diremos que dichos conjuntos son **disjuntos**:*

$$A \text{ y } B \text{ disjuntos } \Leftrightarrow A \cap B = \varnothing$$

Ejemplo 3.1 *Si A= {1,2,3} y B={2,4,6}, tenemos que:*

$$A \cup B = \{1,2,3,4,6\} \text{y } A \cap B = \{2\}$$

Lema 3.1 *Si dos conjuntos A y B son disjuntos, se verifica que*

$$|A \cup B| = |A| + |B|$$

Las técnicas básicas de conteo son el principio de adición o de la suma, el principio de distribución y el principio de multiplicación.

3.2.1. Principio de Adición

El principio de adición es la técnica más sencilla para calcular el cardinal de un conjunto A. La idea consiste en descomponer A en subconjuntos disjuntos dos a dos cuyos cardinales sean más fáciles de calcular, y sumarlos.

Ejemplo 3.2 *Supongamos que lanzamos dos dados distintos y queremos calcular de cuántas formas podemos conseguir que la suma de los puntos sea 7 u 8. Para ello, calculamos primero cuántas formas hay de conseguir 7 de suma. los más rápido es describirlas:*

(1,6),(2,5),(3,4),(4,3),(5,3),(6,1)

Hay 6 formas. A continuación calculamos cuántas formas hay de conseguir 8 de suma. Hay cinco, que son:

(2,6)(3,5),(4,4),(5,3),(6,2)

Por tanto, en total hay 6+5=11 lanzamientos con los que podemos conseguir 7 u 8 de suma.

Usando una notación más matemática, lo que nos piden en el ejemplo anterior es calcular el cardinal de la unión $S_7 \cup S_8$, es decir $|S_7 \cup S_8|$, donde S_i representa el conjunto de lanzamiento cuya suma de puntos vale i. Y lo que hemos hecho para calcular $|S_7 \cup S_8|$ es calcular $|S_7| + |S_8|$.

Teorema 3.1 *Principio de Adición: Si $A_1, A_2, \cdots, A_n$ son conjuntos disjuntos dos a dos, es decir*

$$A_i \cup A_j = \varnothing \; si \;\; i \neq j, 1 \leq i, j \leq n$$

se verifica que

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = |A_1| + |A_2| + \cdots + |A_n|$$

Ejemplo 3.3 *Supongamos que lanzamos cuatro monedas distintas y queremos calcular cuántas formas hay de conseguir al menos dos caras.*

Partimos el conjunto A cuyo cardinal queremos calcular en varios subconjuntos $A_1, \cdots, A_n$ dos a dos disjuntos. En nuestro ejemplo, el conjunto A está formado por todos los lanzamientos de cuatro monedas en los que se obtienen dos, tres o cuatro caras. Por tanto, partimos A en los siguientes tres subconjuntos:

- $A_2 =$ *{lanzamientos en los que se obtienen exactamente dos caras}.*
- $A_3 =$*{lanzamientos en los que se obtienen exactamente tres caras}.*
- $A_4 =$*{lanzamientos en los que se obtienen exactamente cuatro caras}.*

Es evidente que un mismo lanzamiento no puede estar en dos de estos subconjuntos, por lo que son subconjuntos dos a dos disjuntos.

Para calcular el cardinal de cada uno de los subconjuntos A_i, lo más rápido es contar sus elementos:

- $A_2 =$ *{(c,c,x,x),(c,x,c,x),(c,x,x,c),(x,c,c,x),(x,c,x,c),(x,x,c,c)}.*
- $A_3 =$*{(c,c,c,x),(c,c,x,c),(c,x,c,c),(x,c,c,c)}.*
- $A_4 =$*{(c,c,c,c)}.*

Por tanto, $|A_2| = 6, |A_3| = 4, |A_4| = 1$. *Sumando dichos cardinales,* $|A| = |A_2 \cup A_3 \cup A_4| = |A_2| + |A_3| + |A_4| = 6+4+1 = 11$. *Por lo que hay once formas de conseguir al menos dos caras.*

3.2.2. Principio de Multiplicación

La herramienta que más se usa para calcular cardinales de conjuntos es el principio de multiplicación. Su principal aplicación es contar el número de listas ordenadas de longitud *r*, esto es, listas formadas por *r* objetos en los que importa el orden de colocación.

Ejemplo 3.4 *En el menú del día de la facultad ofrecen 2 primeros platos, 3 segundos y 2 postres. ¿Cuántos posibles menús podemos elegir?*

Un menú es una lista ordenada de tres objetos: un primer plato, un segundo y un postre. Para elegir el primer plato tenemos dos opciones, digamos ensalada o macarrones. Una vez seleccionado el primer plato, e independientemente de dicha selección, podemos elegir cualquiera de los tres segundos, por ejemplo, pollo, bacalao o huevos. Esto hace un total de $2 \cdot 3 = 6$ *elecciones posibles para los dos primeros platos. Para cada una de estas 6 elecciones, podemos escoger entre dos postres: fruta o yogur, dando un total de* $6 \cdot 2 = 12$ *posibles menús.*

Representaremos cada lista de r objetos ordenados como si fuese un vector de r componentes. En el ejemplo:

$$\left(\frac{primer\ plato}{2\ elecciones}, \frac{segundo\ plato}{3\ elecciones}, \frac{postre}{2\ elecciones}\right)$$

Este tipo de representación nos permite escribir cuántas elecciones hay en cada posición, por lo que basta multiplicarlas todas para obtener el número total de posibilidades. Es decir $2 \cdot 3 \cdot 2 = 12$.

El principio de multiplicación generaliza el ejemplo anterior:

Teorema 3.2 *Principio de multiplicación: Se quiere hacer una lista ordenada de longitud r. Para el primer lugar de la lista se puede elegir entre* n_1 *objetos, para el segundo entre* n_2*, y así sucesivamente hasta el lugar r-ésimo, para el que se puede elegir entre* n_r *objetos. Entonces, el número total de listas de r objetos ordenados es:*

$$n_1 \cdot n_2 \cdot \cdots \cdot n_r$$

Ejemplo 3.5 *¿Cuántos números de cuatro cifras distintas se pueden formar con los dígitos* $\{1,2,3,4,5,6,7\}$?

Si no nos dejan repetir las cifras, a la hora de poner el primero número tendríamos 7 posibles elecciones. Sin embargo, a la hora de elegir el segundo tendríamos solo 6 elecciones, y así sucesivamente. Así,

$$\left(\frac{unidades\ de\ millar}{7\ elecciones}, \frac{centenas}{6\ elecciones}, \frac{decenas}{5\ elecciones}, \frac{unidades}{4\ elecciones}\right)$$

Aplicando el principio de multiplicación, obtenemos $7 \cdot 6 \cdot 5 \cdot 4 = 840$ *números con las cuatro cifras distintas.*

Si por el contrario se pueden repetir cifras, entonces en el segundo lugar podemos colocar también el mismo número que hayamos colocado en primer lugar. Es decir,

$$\left(\frac{unidades\ de\ millar}{7\ elecciones}, \frac{centenas}{7\ elecciones}, \frac{decenas}{7\ elecciones}, \frac{unidades}{7\ elecciones}\right)$$

Aplicando el principio de multiplicación, obtenemos $7 \cdot 7 \cdot 7 \cdot 7 = 2401$ *números con las cuatro cifras.*

Por último, si queremos saber cuántos son pares:

$$\left(\frac{\textit{unidades de millar}}{\textit{7 elecciones}}, \frac{\textit{centenas}}{\textit{7 elecciones}}, \frac{\textit{decenas}}{\textit{7 elecciones}}, \frac{\textit{unidades}}{\textit{3 elecciones}}\right)$$

Aplicando el principio de multiplicación, obtenemos $7 \cdot 7 \cdot 3 \cdot 7 = 1029$ *números con las cuatro cifras pares.*

Hemos sido capaces de resolver este ejemplo usando solo el principio de multiplicación. Sin embargo, a menudo se nos plantean problemas más complicados para cuya resolución necesitamos combinar el principio de multiplicación con el de adición:

Ejemplo 3.6 *Cada usuario de un servidor de Internet tiene una palabra clave para acceder a dicho servidor. Dicha palabra clave está formada por cuatro caracteres,a elegir entre 26 mayúsculas y 10 dígitos* $\{0,1,\cdots,9\}$. *Además, la clave debe tener al menos un dígito. Por ejemplo 1ATE es un clave válida pero CATE no. ¿Cuántas palabras clave se pueden formar?*

Cada palabra clave es una lista ordenada de longitud cuatro, por lo que el principio de multiplicación parece, a priori, una buena forma de abordar el problema propuesto.

Para el primer carácter de una clave tenemos 36 elecciones (26 caracteres y 10 dígitos). Como no nos exigen que sean distintos, para el segundo carácter volvemos a tener otras 36 elecciones. Y lo mismo ocurre con el tercer carácter. Sin embargo, el número de elecciones para el cuarto carácter depende de las elecciones anteriores.

En efecto, si lo tres primeros caracteres son todos letras, entonces necesariamente el cuarto ha de ser un dígito y sólo tendríamos 10 elecciones. Sin embargo, si ya hay un dígito, seguiríamos teniendo 36 posibles elecciones. Por tanto, no podemos aplicar el principio de multiplicación de manera automática. Sin embargo, la locución "al menos"del enunciado nos sugiere usar el principio de adición.

El conjunto de palabras clave con al menos un dígito se particiona en cuatro subconjuntos: el de las claves que tienen exactamente un dígito, el de las que tienen exactamente cuatro dígitos, el que tiene exactamente 3 dígitos y el que tiene cuatro. Claramente estos conjuntos son disjuntos dos a dos, por lo que podemos aplicar el principio de adición.

Si denotamos a estos conjuntos C_1, C_2, C_3 *y* C_4 *respectivamente, entonces tendremos que calcular:*

$$|C_1| + |C_2| + |C_3| + |C_4|$$

Calculamos la suma por separado:

Para C_1, hay cuatro posibles combinaciones en función de donde está colocado el dígito ($\{(d,c,c,c),(c,d,c,c),(c,c,d,c),(c,c,c,d)\}$). Así, para la primera combinación, hay 10 elecciones para el primer carácter, 26 para el segundo, 26 para el tercero y 26 para el cuarto. Es decir, por el principio de multiplicación hay $10 \cdot 26 \cdot 26 \cdot 26 = 175760$ *posibles claves. Si en vez de elegir esta disposición elegimos otra cualquiera de las tres restantes, cambiará el orden de los elementos. Por tanto, cambiará el orden de los factores pero no el producto final. Así,* $|C_1| = 4 \cdot 10 \cdot 26 \cdot 26 \cdot 26 = 4 \cdot 10 \cdot 26^3$.

Análogamente, hay 6 posibles combinaciones con dos dígitos y dos letras en $C_2 = (\{(d,d,c,c),(d,c,d,c),(d,c,c,d),(c,d,c,d),(c,c,d,d),(c,d,d,c)\})$. *Independientemente del orden, habrá 10 posibles elecciones para el primer dígito, 10 elecciones para el segundo dígito, 26 elecciones para el primer carácter y 26 elecciones para el segundo carácter. Por tanto, por el principio de multiplicación* $|C_2| = 6 \cdot 10 \cdot 10 \cdot 26 \cdot 26 = 6 \cdot 10^2 \cdot 26^2$.

Del mismo modo, hay 4 elementos en $C_3 = \{(d,d,d,c),(d,d,c,d),(d,c,d,d),(c,d,d,d)\}$. *Cada dígito puede obtenerse de un conjunto de 10 elementos y el carácter de un conjunto de 26 elementos. De nuevo por el principio de multiplicación:* $|C_3| = 4 \cdot 10 \cdot 10 \cdot 10 \cdot 26 = 4 \cdot 10^3 \cdot 26$.

Y finalmente, solo hay una forma de tener una clave con cuatro dígitos, pudiendo escoger cada dígito del conjunto de 10 elementos antes mencionado, por tanto, $|C_4| = 10^4$.

Por tanto, el número total de claves es $|C_1| + |C_2| + |C_3| + |C_4| = 4 \cdot 10 \cdot 26^3 + 6 \cdot 10^2 \cdot 26^2 + 4 \cdot 10^3 \cdot 26 + 10^4$

3.2.3. Principio de Distribución

Un resultado inmediato del principio de adición es el denominado principio de distribución. Uno de los ejemplos típicos en matemáticas de cómo una idea sencilla aplicada con habilidad puede resolver elegamente un problema, es el *principio de distribución de Dirichlet* o *principio del palomar*

Lema 3.2 *Principio del Palomar: Si $m+1$ palomas ocupan m nidos, entonces algún nido contiene más de una paloma.*

Ejemplo 3.7 *En un grupo de 13 personas hay al menos dos que cumplen los años el mismo mes.*

Ejemplo 3.8 *Para sacar dinero del cajero automático, los clientes de una entidad bancaria tienen que teclear una clave de cuatro dígitos. Si dicha entidad tiene más de*

10000 clientes, entonces hay al menos dos clientes con la misma clave. En efecto, por el principio de multiplicación, el número de claves de cuatro dígitos es $10 \cdot 10 \cdot 10 \cdot 10 = 10000$.

Esto se puede generalizar al *principio de distribución*.

Teorema 3.3 *Principio de Distribución: Si queremos repartir n objetos en m cajas y* $r \cdot m < n$*, al menos una caja ha de recibir más de r objetos.*

Ejemplo 3.9 *En cualquier conjunto de 38 enteros positivos menores de 1000 siempre hay dos cuya diferencia es menor de 27. En efecto, si dividimos los 999 enteros positivos menores de 1000 en grupos de 27 empezando por 1:* $[1,27],[27,54],\cdots,[972,999]$. *Como 999=37· 27, hay 37 grupos. Por tanto, al elegir 38 números, habrá dos que estén en el mismo grupo, por lo que su diferencia será menor de 27.*

3.2.4. Principio de inclusión y exclusión

Acabamos de ver que, por el principio de adición, si dos conjuntos son disjuntos se verifica que $|A \cup B| = |A| + |B|$. Sin embargo, no sabemos nada sobre el cardinal de la unión cuando los conjuntos no son disjuntos.

Ejemplo 3.10 *Supongamos que en el caso en el que lanzamos dos dados distintos, nos piden calcular de cuántas formas podemos conseguir que la suma de los puntos sea múltiplo de 4 o de 6.*

Si calculamos primero cuántas formas hay de conseguir un múltiplo de 4, vemos que hay nuevo. Llamemos A al conjunto de todas ellas:

$$A = \{(1,3),(2,2),(3,1),(2,6),(3,5),(4,4),(5,3),(6,2),(6,6)\}$$

A continuación, calculamos cuántas formas hay de conseguir un múltiplo de 6; hay seis. Llamemos B al conjunto de todas ellas:

$$A = \{(1,5),(2,4),(3,3),(4,2),(5,1),(6,6)\}$$

Vemos que para obtener el número de tiradas cuya suma de puntos sea múltiplo de 4 o de 6 no podemos aplicar directamente el principio de adición y sumar cardinales: $|A| + |B| = 9 + 6$*, pues contaríamos dos veces al* $(6,6)$*. Por tanto, no podemos aplicar el principio de adición.*

Dados que $A \cap BA$ y $A \cap BB$, los elementos de $A \cap B$ se han contado tanto al contar los elementos de A como al contar los elementos de B. Sin embargo, solo debemos hacerlo una vez. Por eso, si los conjuntos no son disjuntos se va a verificar:

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Para el caso de tres conjuntos se verifica:

$$\begin{aligned} |A \cup B \cup C| &= |A \cup (B \cup C)| = |A| + |B \cup C| - |A \cap (B \cup C)| = \\ &= |A| + |B| + |C| - |B \cap C| - |(A \cap B) \cup (A \cap C)| = \\ &= |A| + |B| + |C| - |B \cap C| - (|A \cap B| + |A \cap C| - |A \cap B \cap C|) = \\ &= |A| + |B| + |C| - (|A \cap B| + |A \cap C|) + |B \cap C|) + |A \cap B \cap C| \end{aligned}$$

Si llamamos

$$\begin{aligned} \alpha_1 &= |A| + |B| + |C| \\ \alpha_2 &= |A \cap B| + |A \cap C| + |B \cap C| \\ \alpha_3 &= |A \cap B \cap C| \end{aligned}$$

Podemos generalizar el resultado para obtener el siguiente teorema

Teorema 3.4 *Principio de Inclusión-ExclusiónSi $A_1, A_2, \cdots, A_n$ son conjuntos finitos y denotamos por α_i a la suma de los cardinales de las intersecciones de i conjuntos*

$$\begin{aligned} \alpha_1 &= |A_1| + |A_2| + \cdots + |A_n| \\ \alpha_2 &= |A_1 \cap A_2| + |A_1 \cap A_3| + \cdots + |A_{n-1} \cap A_n| \\ \cdots & \\ \alpha_n &= |A_1 \cap A_2 \cap \cdots \cap A_n| \end{aligned}$$

se verifica que

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \alpha_1 - \alpha_2 + \cdots (-1)^{n+1} \alpha_n$$

Ejemplo 3.11 *En un encuentro de informática hay participantes de España, Francia e Inglaterra. Sabiendo que 9 de ellos hablan francés, otros 9 inglés y otros 9 castellano; que hay 4 que hablan francés e ingles, 3 que hablan francés y castellano, 4 que hablan inglés y castellano y 1 que hablan los tres idiomas, ¿cuántos participantes hay en el encuentro? ¿hay alguno que sólo hable castellano?*

Indicando por F, I y C a los conjuntos de participantes que habla francés, inglés y castellano respectivamente, los datos que tenemos son los siguientes:

$$|F| = |I| = |C| = 9; |F \cap I| = 4; |F \cap C| = 3; |I \cap C| = 4; |F \cap I \cap C| = 1$$

Teniendo en cuenta que cualquier de los participantes habla alguno de los tres idiomas, el principio de inclusión y exclusión nos dice que:

$$|F \cup I \cup C| = |F| + |I| + |C| - |F \cap I| - |F \cap C| - |I \cap C| + |F \cap I \cap C| = 9+9+9-4-3-4+1 = 17$$

por lo que hay 17 participantes de los cuales hablará sólo castellano los que no sepan francés ni inglés. Dado que

$$|F \cup I| = |F| + |I| - |F \cap I| = 9+9-4 = 14$$

hay 14 participantes que hablan francés o inglés por lo que el resto, es decir, 3 sólo saben castellano.

3.3. Variaciones, Permutaciones y Combinaciones

En esta sección abordamos el siguiente problema: dado un conjunto con n objetos, ¿cuál es el número de colecciones de k objetos que se pueden formar eligiéndolos entre los n del conjunto? La respuesta va a depender del orden y de los elementos repetidos. Las posibles respuestas dan lugar a cuatro tipo de colecciones: *variaciones, variaciones con repetición, combinaciones combinaciones con repetición*. Como caso particular de las variaciones estudiaremos también las permutaciones.

En la tabla 3.1 encontrará un resumen de los distintos tipos de colecciones.

	# objetos	# obj/grupo	Repetición	Orden	Formula
$V_{(n,k)}$	n	k	No	Si	$\frac{n!}{(n-k)!}$
$VR_{(n,k)}$	n	k	Si	Si	n^k
P_n	n	n	No	Si	$n!$
$PR^n_{n_1,\cdots n_k}$	n	$n_1,\cdots n_k$	Si	Si	$\frac{n!}{n_1!n_2!\cdots n_k!}$
PC_n	n	n	No	Circ.	$(n-1)!$
$C_{(n,k)}$	n	k	No	No	$\frac{n!}{(n-k)!k!}$
$CR_{(n,k)}$	n	k	Si	No	$\frac{(n+k-1)!}{(n-1)!k!}$

Tabla 3.1: Tabla resumen combinatoria

3.3.1. Variaciones

Definición 3.4 *Variaciones $V_{(n,k)}$: Una variación de orden k de n objetos ($k \leq n$) (o una variación de n elementos tomados de k en k) es una lista ordenada de k objetos distintos elegidos entre n objetos dados.*

El número $V_{(n,k)}$ de variaciones de orden k de n elementos es:

$$V_{(n,k)} = n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

Ejemplo 3.12 *¿Cuántos números de seis cifras pueden escribirse sabiendo que deben comenzar por 1 y no tener cifras repetidas?*

Teniendo en cuenta que todos comienzan por 1, basta con escribir las otras cinco cifras del número y que ninguna de estas puede estar repetida ni ser un 1: Es decir, con las cifras 0,2,3,4,5,6,7,8 y 9 debemos formar todas las variaciones sin repetición de longitud 5:

$$V_{9,5} = \frac{9!}{(9-5)!} = \frac{9!}{4!} = 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = 15120$$

por lo que existen 15120 números con las características pedidas

Si se pueden repetir los elementos, se tiene:

Definición 3.5 *Variaciones con repetición $VR_{(n,k)}$: Una variación con repetición de orden k de n es una lista ordenada de objetos no necesariamente distintos elegidos entre n objetos dados. También se suele llamar variación con repetición de n elementos tomados de k en k.*

$$VR_{(n,k)} = n^k \tag{3.1}$$

Ejemplo 3.13 *¿Cuántos posible números puedo representar en un buffer de 3 bytes?*

Un byte corresponden a 8 bits, luego tengo 24 bits. Cada bit puede contener un 0 ó un 1. Como importa el orden y se puede repetir, es una variación con repetición:

$$VR_{2,24} = 2^{24}$$

3.3.2. Permutaciones

Si $k = n$, entonces la variación se conoce como *permutación*

Definición 3.6 *Permutaciones P_n: Una permutación de n objetos distintos es una lista ordenada de los n objetos. Es decir, es una variación de orden n de n objetos.*

El número P_n de permutaciones de n objetos distintos es :

$$P_n = n! \tag{3.2}$$

Ejemplo 3.14 *Las letras de la palabra CESA pueden ser permutadas de*

$$4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$$

formas diferentes.

Supongamos ahora que tratamos de contar el número de formas en que se pueden permutar las letras de la palabra CASA.

Como se trata de permutar las letras, estamos hablando de permutaciones, pero hay que observar que si las letras que se permutan son las dos A-es, la palabra resultante es la misma, por lo que sólo obtendremos la mita de las ordenaciones que en el caso de la palabra CESA, que tiene las cuatro letras diferentes.

Definición 3.7 *Permutaciones con repetición $P_n^{n_1,\cdots n_k}$: El número de permutaciones con repetición de un conjunto de n elementos donde existe un grupo de n_1 elementos repetidos, otro de n_2 elementos etc. viene dado por :*

$$PR^n_{n_1,\cdots n_k} = \frac{n!}{n_1!n_2!\cdots n_k!} \tag{3.3}$$

Ejemplo 3.15 *Las letras de la palabra CASA pueden ser permutadas de*

$$PR^4_{1,2,1} = \frac{4!}{1!2!1!} = \frac{4\cdot 3\cdot 2\cdot 2\cdot 1}{2\cdot 1} = \frac{24}{2} = 12 \tag{3.4}$$

formas diferentes

Definición 3.8 *Permutaciones circulares PC_n: Lista ordenada de n objetos donde no importa el orden sino la posición relativa.*

$$PC_n = (n-1)! \tag{3.5}$$

Ejemplo 3.16 *Si son 5 comensales que hay que colocar alrededor de una mesa circular, entonces las distintas permutaciones circulares se obtienen fijando uno de ellos y permutando los 4 restantes de todas las maneras posibles. Es decir de (5-1)!=4!= 24 formas distintas.*

3.3.3. Combinaciones

Si no importa el orden, entonces la técnica de conteo es la combinación:

Definición 3.9 *Combinaciones $C_{(n,k)}$: Una combinación de orden k de n objetos ($k \leq n$) es una colección de k objetos distintos elegidos entre n objetos dados sin importar el orden. También se suele llamar combinación de n objetos tomados de k en k.*

Una combinación de orden k de n objetos no es más que un subconjunto de k elementos elegidos en un conjunto de n.

$$C_{(n,k)} = \frac{V_{(n,k)}}{k!} = \frac{n!}{(n-k)!k!} \tag{3.6}$$

Definición 3.10 *Combinaciones con repetición: Una combinación con repetición de orden k de n objetos tomados de k en k, es una colección de k objetos no necesariamente distintos elegidos entre n objetos dados sin importar el orden:*

$$CR_{(n,k)} = C_{(n+k-1,k)} = \frac{(n+k-1)!}{(n-1)!k!} \tag{3.7}$$

3.4. Coeficientes Binomiales

Tal y como se observó anteriormente, el número de k-combinaciones de un conjunto de n elementos se denota a menudo como $\binom{n}{k}$. Este número se conoce también como **coeficiente binomial** porque aparece como coeficiente en el desarrollo de la expresión $(a+b)^n$. En esta sección veremos el **Teorema del Binomio**, que proporciona el desarrollo de la potencia anterior en términos de potencia de a y b y unos ciertos coeficientes, que son los coeficientes binomiales.

3.4.1. Teorema del binomio

El teorema del binomio, debido a Newton, es una de las fórmulas más útiles que se pueden encontrar en ramas de las matemáticas muy distintas entre sí. En su versión más elemental, nos proporciona una fórmula para desarrollar la potencia n-ésima de un binomio, esto es, para desarrollar la expresión $(a+n)^n$, donde n es un entero positivo. Los coeficientes que aparecen en este desarrollo se llaman coeficientes binomiales, y resulta ser los números combinatorios $C(n,k)$.

Definición 3.11 *Números combinatorios: Al número $C_{(n,k)}$ de combinaciones de n elementos tomados de k en k se le denomina número combinatorio o coeficiente de binomios. Normalmente se denota como $C_{(n,k)} = \binom{n}{k}$.*

Los binomios tienen las siguientes propiedades:

Lema 3.3 *Para todo n y k con $k \leq n \Rightarrow \binom{n}{k} = \binom{n}{n-k}$*

Proof 3.1 *Decidir qué k objetos de un conjunto de n elementos van a formar parte de un subconjunto, es lo mismo que decidir qué $n-k$ objetos no van a formar parte. Por tanto, el número de formar de elegir un subconjunto de k objetos entre n (que es el lado izquierdo de la igualdad) es el mismo que el número de formas de elegir un subconjunto de $n-k$ objetos entre n (que es el lado derecho de la igualdad).*

Lema 3.4 *Fórmula de Pascal: Para todo n y k con* $k \leq n \Rightarrow \binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.

Proof 3.2 *El lado izquierdo de la igualdad es el número de subconjuntos de k elementos elegidos entre los de un conjunto T, de n elementos. Sea a un elemento de T y* $S = T - \{a\}$. *Obsérvese que hay* $\binom{n}{k}$ *subconjuntos de k elementos en T. Sin embargo, un subconjunto de k elementos de T bien contiene al elemento a junto con otros* $k-1$ *elementos de S; o bien contiene k elementos de S y no contiene al elemento a. Los subconjuntos que no contienen a a se forman eligiendo k elementos del conjunto S. Como en este conjunto hay* $n-1$ *elementos, hay* $\binom{n-1}{k}$ *subconjuntos de la segunda clase. Por otra parte, los conjuntos que sí contienen a constan de a y otros* $k-1$ *elementos de S. Se forman entonces eligiendo* $k-1$ *elementos de S y habrá* $\binom{n-1}{k-1}$ *subconjuntos de la primera clase. Como estas dos clases son disjuntas, el número total será la suma de ambas.*

Coloquemos los números combinatorios $C(n,k)$ en filas de manera que en la fila n-ésima estén los números $C(n,0), C(n,1), \cdots C(n,n)$. Empezando por la fila cero, esto nos dará una distribución de los números combinatorios en forma de triángulo, que algunos llaman **Triángulo de Pascal**

$$\begin{array}{ccccccc}
 & & & \binom{0}{0} & & & \\
 & & \binom{1}{0} & & \binom{1}{1} & & \\
 & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & \\
\binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} \\
 & & & \cdots & & &
\end{array}$$

La fórmula de Pascal nos dice que cada número combinatorio que aparece en fila se obtiene sumando los dos números de la fila superior que está justo encima de él:

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$$

De este modo se pueden calcular los primeros números combinatorios. Por ejemplo, las seis primeras filas del triángulo de Pascal son

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
 & & & & & \ldots & & & & &
\end{array}
$$

Teorema 3.5 *Cálculo de los números combinatorios:Si k y n son enteros positivos tales que* $1 \leq k \leq n$ *se verifica que*

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}$$

Proof 3.3 *Por inducción,*

- *La fórmula es cierta para* $n = 1$ *ya que si* $n = 1$ *ha de ser necesariamente* $k = 1$ *y* $\binom{1}{1} = 1$ *ya que un conjunto de un sólo elemento sólo tiene un subconjunto de un elemento.*
- *Si se verifica para n, vamos a probar que también es cierto para* $n+1$*. Es decir, suponiendo que* $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ *, quiero demostrar que* $\binom{n+1}{k} = \frac{(n+1)!}{k!(n+1-k)!}$*. En efecto, por la fórmula de Pascal se tiene que*

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$$

y dado que estamos suponiendo cierta la propiedad para n, tenemos que

$$\begin{aligned}
\binom{n+1}{k} &= \frac{n!}{(n-k+1)!\cdot(k-1)!} + \frac{n!}{(n-k)!\cdot k!} = \\
&= \frac{n!\cdot k}{(n-k+1)!\cdot(k-1)!\cdot k} + \frac{n!\cdot(n-k+1)}{(n-k)!\cdot k!\cdot(n-k+1)} = \\
&= \frac{n!\cdot k}{(n-k+1)!\cdot k!} + \frac{n!\cdot(n-k+1)}{k!\cdot(n-k+1)!} \\
&= \frac{n!\cdot k + n!\cdot(n-k+1)}{k!\cdot(n-k+1)!} = \frac{n!\cdot(k+n-k+1)}{k!\cdot(n-k+1)!} = \\
&= \frac{n!\cdot(n+1)}{k!\cdot(n-k+1)!} = \frac{(n+1)!}{k!\cdot(n-k+1)!}
\end{aligned}$$

Si $k = 0$ *o* $k = n+1$*, los valores* $\binom{n}{0}$*=1 y* $\binom{n}{n-1} = 0$ *aseguran la validez de la demostración.*

Teorema 3.6 *Teorema del Binomio: Para cualquier entero positivo:*

$$\begin{aligned}(a+b)^n &= \binom{n}{0}\cdot a^n + \binom{n}{1}\cdot a^{n-1}\cdot b + \binom{n}{2}\cdot a^{n-2}\cdot b^2 + \cdots \binom{n}{n}\cdot b^n \\ &= \sum_{k=0}^{n}\binom{n}{k}\cdot a^{n-k}\cdot b^k\end{aligned}$$

Lema 3.5

$$\forall n \geq 0 \Rightarrow \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n-1} + \binom{n}{n} = 2^n \tag{3.8}$$

Proof 3.4 *Aplicando el teorema de binomio con $x = 1$ e $y = 1$, se tiene que:*

$$2^n = (1+1)^n = \sum_{k=0}^{n}\binom{n}{k}1^k 1^{n-k} = \sum_{k=0}^{n}\binom{n}{k}$$

Ejemplo 3.17 *Calcular 7^6 con el teorema del binomio:*

$$\begin{aligned}7^6 = (5+2)^6 &= \binom{6}{0}\cdot 2^6 + \binom{6}{1}\cdot 2^5\cdot 5 + \binom{6}{2}\cdot 2^4\cdot 5^2 + \binom{6}{3}\cdot 2^3\cdot 5^3 + \\ &\quad \binom{6}{4}\cdot 2^2\cdot 5^4 + \binom{6}{5}\cdot 2\cdot 5^5 + \binom{6}{6}\cdot 5^6\end{aligned}$$

Ejemplo 3.18 *Calcular $(x+y)^8$ con el teorema del binomio:*

$$\begin{aligned}(x+y)^8 &= \binom{8}{0}\cdot x^8 + \binom{8}{1}\cdot x^7\cdot y + \binom{8}{2}\cdot x^6\cdot y^2 + \binom{8}{3}\cdot x^5\cdot y^3 + \\ &\quad \binom{8}{4}\cdot x^4\cdot y^4 + \binom{8}{5}\cdot x^3\cdot y^5 + \binom{8}{6}\cdot x^2\cdot y^6 + \binom{8}{7}\cdot x\cdot y^7 + \binom{8}{8}\cdot y^8 = \\ &\quad x^8 + 8\cdot x^7\cdot y + 28\cdot x^6\cdot y^2 + 58\cdot x^5\cdot y^3 + 70\cdot x^4\cdot \\ &\quad y^4 + 58\cdot x^3\cdot y^5 + 28\cdot x^2\cdot y^6 + 8\cdot x\cdot y^7 + y^8\end{aligned}$$

Universidad Rey Juan Carlos

Tema 4

Teoría de Grafos

4.1. Teoría de Grafos

La teoría de grafos es una de las partes de las Matemáticas que más se han desarrollado en las últimas décadas gracias a sus diversas aplicaciones en múltiples disciplinas. La construcción de redes de comunicación, la confección de horarios, el diseño de circuitos integrados, problemas de emparejamiento, de vigilancia, etc.

4.1.1. Tipos de Grafos

Cuando pensamos, por ejemplo, en el diseño de una red de ordenadores y de las posibles conexiones entre ellos, podemos imaginar un dibujo formado por puntos (ordenadores) y líneas que los unen entre ellos (las conexiones entre ellos). Es decir, los grafos son estructuras discretas que constan de vértices y de aristas que conectan entre sí esos vértices. Hay varios tipos de grafos, que se diferencian entres sí por el tipo y el número de aristas que pueden conectar cada par de vértices.

Vamos a presentar los diferentes tipos de grafos mostrando la forma en que se pueden utilizar cada uno de ellos para modelar una red informática. Supongamos que una red consta de ordenadores y de líneas telefónicas que conectan esos ordenadores. Podemos representar cada ordenador mediante un punto y cada línea telefónica mediante un segmento, tal y como se muestra en la Fig. 4.1

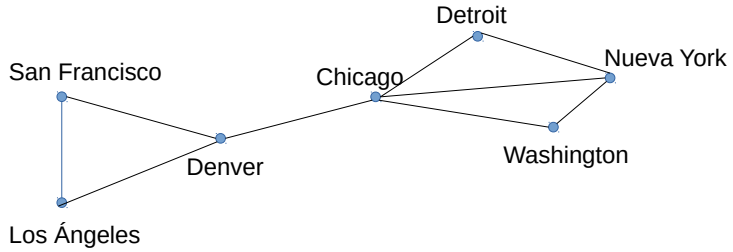


Figura 4.1: Ejemplo Grafo Simple

Definición 4.1 (Grafo Simple) *Un grafo simple $G = (V,E)$ consta de V, un conjunto no vacío de vértices, y de E, un conjunto de pares no ordenados de elementos distintos de V. A estos pares se les llama aristas.*

Cuando hay mucho tráfico de información, puede haber líneas telefónicas múltiples entre los ordenadores de la red, como se muestra en la Fig. 4.2. Los grafos simples no bastan para modelar esta situación. En lugar de grados simples, emplearemos multigrafos, que constan de vértices y de aristas no dirigidas entre esos vértices, pero admitiendo la existencia de aristas múltiples entre pares de vértices. Todo grado simple es un multigrafo, pero no todos los multigrafos son grafos simples. No podemos usar simplemente un par de vértices para especificar la arista de un grafo si hay aristas múltiples, esto lleva a la siguiente definición:

Definición 4.2 (Multigrafo) *Un multigrafo $G = (V,E)$ consta de un conjunto V de vértices, un conjunto E de aristas y una función f de E en $\{\{u,v\}|u,v \in V, u \neq v\}$. Se dice que las aristas e_1 y e_2 son aristas múltiples o paralelas si $f(e_1) = f(e_2)$*

Definición 4.3 (Multiplicidad de una arista) *, m_{ij}, es el número de veces que dicha arista aparece en el multigrafo.*

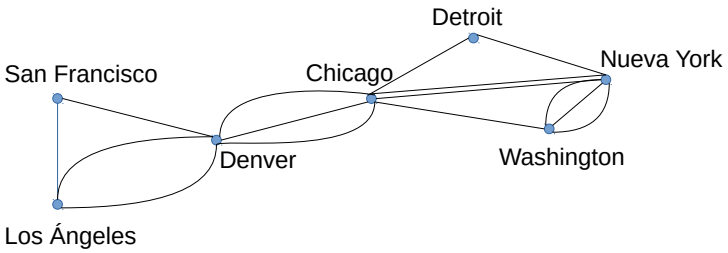


Figura 4.2: Ejemplo Grafo Múltiple

Definición 4.4 (p-grafo) *Un p-grafo, $p \in \mathbb{N}$, es un multigrafo en el que ninguna arista tiene una multiplicidad mayor que p, y existe alguna arista con multiplicidad p.*

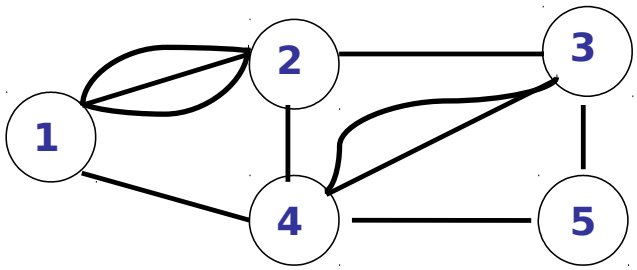


Figura 4.3: Ejemplo de 3-grafo

Una red informática puede contener una línea telefónica que conecte un ordenador consigo mismo. No podemos usar multigrafos para para representar estas redes, ya que no se admiten **bucles** (i.e, aristas que conectan un vértice consigo mismo) en un mutligrafo. En lugar de multigrafos, utilizaremos pseudografos (ver Fig. 4.4, que son más generales que los multigrafos, ya que una arista de un pseudografo puede conectar un vértice consigo mismo).

Definición 4.5 (Pseudografo) *Un pseudografo $G = (V,E)$ consta de un conjunto V de vértices, un conjunto E de aristas y una función f de E en $\{\{u,v\}|u,v \in V\}$. Una arista e es un bucle, o lazo, si $f(e) = \{u,v\} = \{u\}$ para algún $u \in V$.*

Puede que las líneas telefónicas de una red informática no operen en las dos direcciones. Utilizaremos grafos dirigidos (ver Fig. 4.5) para modelar redes de este tipo. Las aristas de un grafo dirigido son pares ordenados. Se admiten los bucles, pares ordenados con sus dos elementos iguales, pero no se admiten aristas múltiples en la misma dirección entre dos vértices.

Definición 4.6 (Grafo dirigido o digrafo) *Un grafo dirigido (V,E) consta de un conjunto V de vértices y de un conjunto E de aristas, que son pares ordenados de elementos de V.*

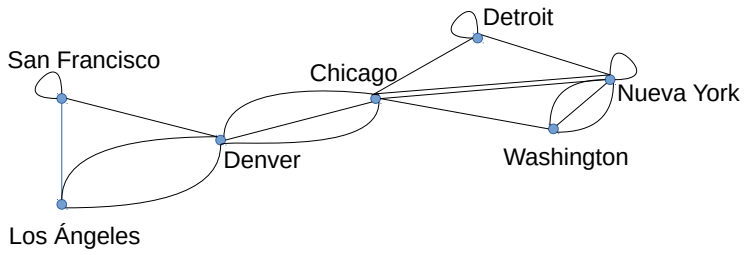


Figura 4.4: Ejemplo Pseudografo

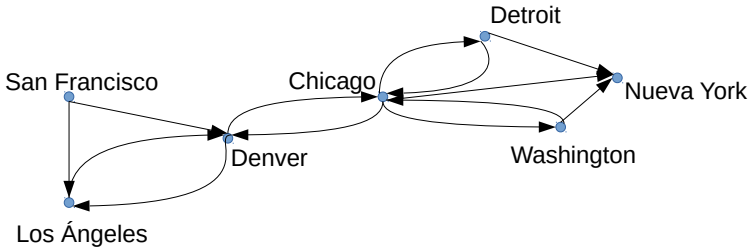


Figura 4.5: Ejemplo Grafo Dirigido

Utilizamos una flecha apuntando desde u hacia v para indicar la dirección de la arista $\{u,v\}$.

Finalmente, puede haber líneas múltiples en la red informática, de modo que haya varias líneas unidireccionales desde cada nodo en dirección a un mismo ordenador, y quizá más de una línea de vuelta a cada ordenador como en la Fig. 4.6. Utilizaremos multigrafos dirigidos, que pueden tener aristas dirigidas múltiples desde un vértice a un segundo vértice (que, eventualmente, puede coincidir con el primero), para representar este tipo de redes. Formalmente,

Definición 4.7 (Multigrafo dirigido) *Un multigrafo dirigido G=(V,E) consta de un conjunto V de vértices, un conjunto E de aristas y una función f de E en $\{\{u,v\}|u,v \in V\}$. Se dice que las aristas e_1 y e_2 son aristas múltiples si $f(e_1) = f(e_2)$*

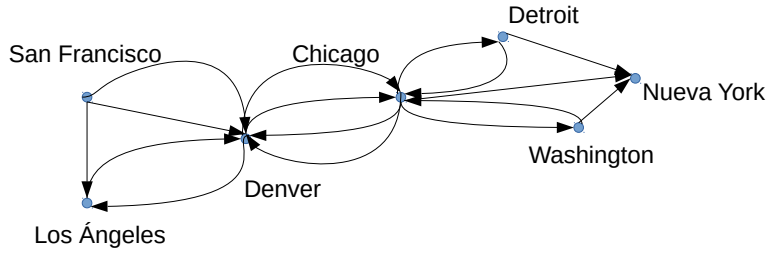


Figura 4.6: Ejemplo Multigrafo Dirigido

Hay ocasiones en que un grafo tiene la propiedad de que su conjunto de vértices se puede dividir en dos subconjuntos disjuntos tales que cada arista conecta un vértice de

uno de esos subconjuntos con un vértice de otro subconjunto. Por ejemplo, consideremos el grafo que representa matrimonios entre las personas de un pueblo. En él, cada persona se representa mediante un vértice y cada matrimonio se presenta mediante una arista.

Definición 4.8 (Grafos Bipartitos) *Se dice que un grafo simple G es bipartito si su conjunto de vértices V se puede dividir en dos conjuntos disjuntos V_1, V_2 tales que cada arista del grafo conecta un vértice de V_1 con un vértice de V_2, de manera que no haya ninguna arista que conecte entre sí dos vértices de V_1 ni tampoco dos vértices de V_2.*

Ejemplo 4.1 *En la Fig. 4.7, el grafo G es bipartito, mientras que H no lo es. En efecto, en G los vértices de $V_1 = \{a,b,d\}$ conectan con los vértices de $V_2 = \{c,e,f,g\}$.*

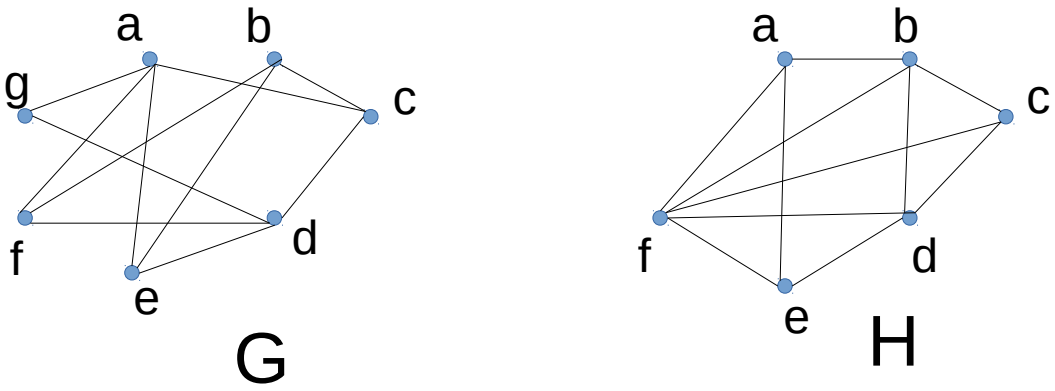


Figura 4.7: Ejemplo grafo bipartito

Definición 4.9 (Subgrafo) *Sea el grafo $G = (V,E)$, el subgrafo generado por el conjunto de vértices es el grafo $G_B = (B,E_B)$, formado por los vértices de B y las aristas de E cuyos extremos pertenecen a B*

$$E_B = \{e \in E / e = (i,j), i,j \in B\}$$

Ejemplo 4.2 *Tomando los vértices $B = \{1,2,4,5\}$ y las aristas $E_B = \{(1,2),(1,4),(2,4),(4,5)\}$, G_B es subgrafo de G.*

4.2. Conceptos básicos

Definición 4.10 (Vértices adyacentes) *Dos vértices a y b son adyacentes si están unidos por una arista, es decir, si $\{a,b\} \in A$. En tal caso, se dice que a y b son los extremos de la arista $\{a,b\}$.*

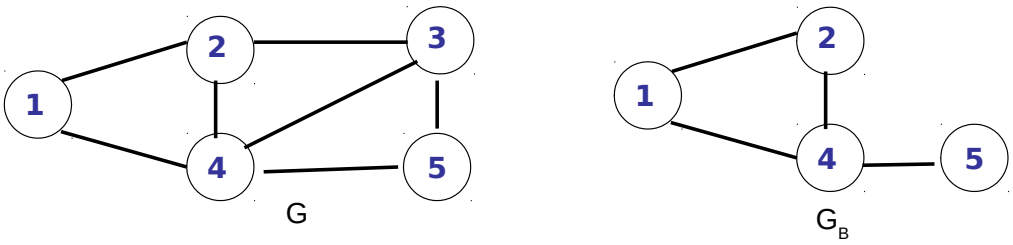


Figura 4.8: Ejemplo subgrafo

Definición 4.11 (Aristas Incidentes) *Se dice que la arista e es incidente con los vértices u y v, si u es adyacente a v. También se dice que e conecta a u y v, o que u y v son los extremos de la arista e.*

Definición 4.12 (Grado de un vértice de un grafo no dirigido) *: Número de aristas incidentes con él, exceptuando los bucles, cada uno de los cuales contribuye con dos unidades al grado del vértice. El grado del vértice se denota por* $\delta(v)$.

Ejemplo 4.3 *¿Cuáles son los grados de los vértices de los grafos G y H que se muestran en la Fig. 4.9*

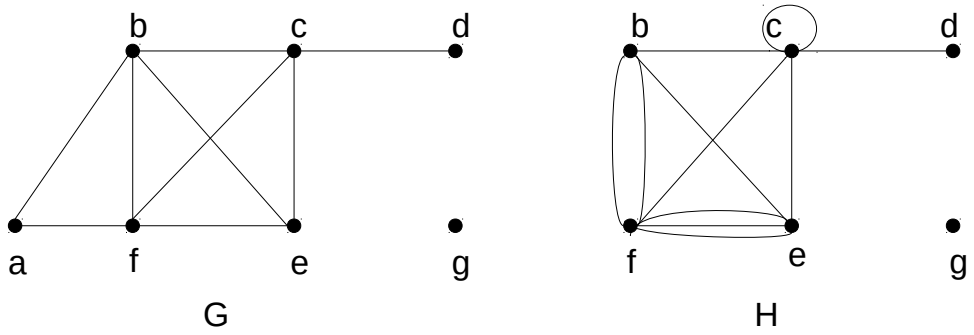


Figura 4.9: Ejemplo grados grafo

En G, $\delta(a) = 2$, $\delta(b) = \delta(c) = \delta(f) = 4$, $\delta(d) = 1$, $\delta(e) = 3$ *y* $\delta(g) = 0$.

En H, $\delta(b) = 4$, $\delta(c) = \delta(f) = 6$, $\delta(d) = 1$ *y* $\delta(e) = 5$.

Teorema 4.1 (El Teorema de los Apretones de Manos) *Sea G=(V,E) un grafo no dirigido con e aristas, entonces:*

$$2e = \sum_{v \in V} \delta(v)$$

Ejemplo 4.4 *¿Cuántas aristas hay en un grafo con diez vértices, cada uno de los cuales tiene grado seis?*

Como la suma de los grados de los vértices es $6 \cdot 10 = 60$*, se sigue que* $2e = 60$*. Por tanto,* $e = 30$*.*

El teorema del apretón de manos nos dice que la suma de los grados de los vértices de un grafo no dirigido es un número par. Como consecuencia, se enuncia el siguiente teorema:

Teorema 4.2 *Todo grafo no dirigido tiene un número par de vértices de grado impar.*

Proof 4.1 *Sean* V_1 *y* V_2 *el conjunto de vértices de grado par y el conjunto de vértices de grado impar, respectivamente, de un grafo no dirigido* $G = (V, E)$*. Entonces,*

$$2e = \sum_{v \in V} \delta(v) = \sum_{v \in V_1} \delta(v) + \sum_{v \in V_2} \delta(v)$$

Como $\delta(v)$ *es par si* $v \in V_1$*, el primer sumando es par, puesto que esa suma es* $2e$*. Por tanto, el segundo sumando es también par. Como todos los términos que se suman en ese segundo sumando son impares, tiene que haber un número par de ellos. Por tanto, hay un número par de vértices de grado impar.*

Definición 4.13 (Vértice inicial y vértice final) *Si* (u, v) *es una arista del grafo dirigido G, se dice que u es adyacente a v y que v es adyacente desde u. Al vértice u se le llama vértice inicial de* (u, v) *y a v se le llama vértice final o terminal de* (u, v)*. Los vértices incial y final de un bucle coinciden.*

Definición 4.14 (Grado de entrada y de salida de un vértice) *: En un grafo dirigido, el grado de entrada de un vértice v, denotado por* $\delta^-(v)$*, es el número de aristas que tienen a v como vértice final. El grado de salida de un vértice v, denotado por* $\delta^+(v)$*, es el número de aristas que tienen a v como vértice inicial (nótese que un bucle contribuye con una unidad tanto al grado de entrada como al grado de salida del vértice correspondiente).*

Ejemplo 4.5 *Halla los grados de entrada y de salida de cada vértice del grafo dirigido G que se muestra en Fig. 4.10*

Los grados de entrada son: $\delta^-(a) = 2, \delta^-(b) = 2$ *,* $\delta^-(c) = 2, \delta^-(d) = 2, \delta^-(e) = 3, \delta^-(f) = 0$*. Los grados de salda son;* $\delta^+(a) = 4, \delta^+(b) = 1, \delta^+(c) = 1, \delta^+(d) = 2, \delta^+(e) = 2, \delta^+(f) = 0$*.*

Como cada arista tiene un vértice inicial y un vértice final, la suma de los grados de entrada y la suma de los grados de salida de todos los vértices del grafo dirigido coinciden. Ambas sumas son iguales al número de aristas que tiene el grafo, como se enuncia en el siguiente teorema:

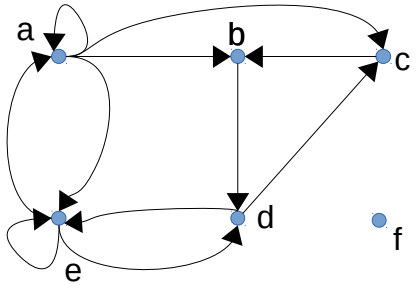


Figura 4.10: Ejemplo grados grafo dirigido

Teorema 4.3 *Sea G=(V,E) un grafo dirigido, entonces*

$$\sum_{v \in V} \delta^-(v) = \sum_{v \in V} \delta^+(v) = |E|$$

4.2.1. Familias de grafos

Definición 4.15 (Grafos completos (K_n)) *Los grafos que tienen todas las aristas posibles, es decir, aquellos en que cada uno de sus vértices está unido con todos los demás vértices, se llaman completos. Se denotan K_n, donde n es el número de vértices del grafo. En la Fig. 4.11 se muestran los grafos K_n para $n = 1,2,3,4,5$*

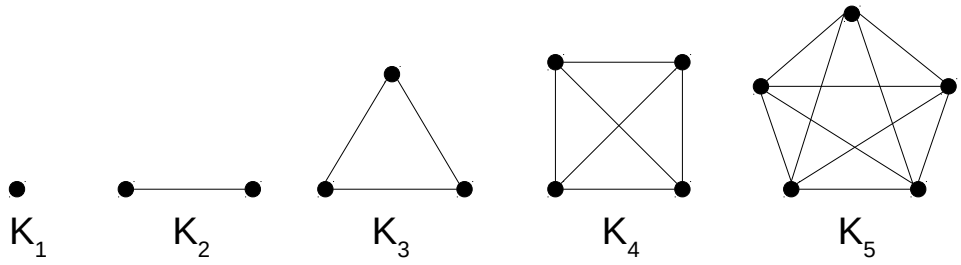


Figura 4.11: Grafos K_n para $n = 1,2,3,4,5$

Definición 4.16 (Ciclos (C_n)) *El ciclo C_n $n \geq 3$, consta de n vértices $v_1, v_2, \cdots, v_n$ y aristas $\{v_1, v_2\}, \{v_2, v_3\}, \cdots, \{v_{n-1}, v_n\}$ y $\{v_n, v_1\}$. En la Fig. 4.12 se muestran los grafos C_n para $n = 3,4,5,6$*

Definición 4.17 (Ruedas (W_n)) *Obtenemos la rueda W_n cuando añadimos un vértice adicional al ciclo C_n, para $n \geq 3$, y conectamos ese nuevo vértice con cada uno de los n vértices de C_n mediante una nueva arista. En la Fig. 4.13 se muestran los grafos W_n para $n = 3,4,5,6$*

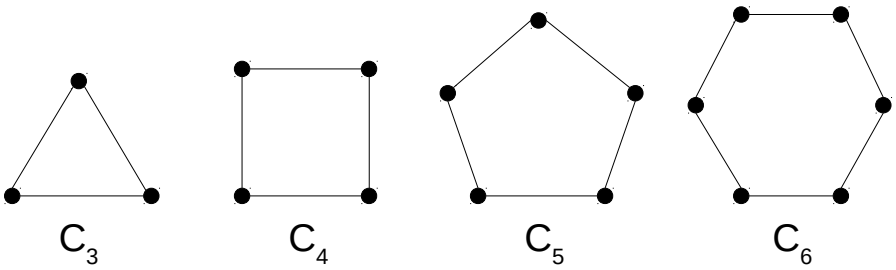


Figura 4.12: Grafos C_n para $n = 3,4,5,6$

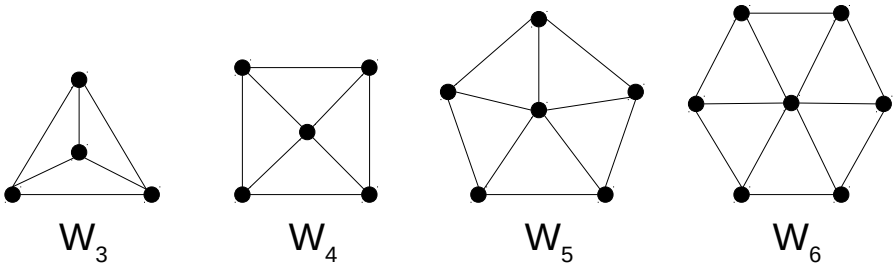


Figura 4.13: Grafos W_n para $n = 3,4,5,6$

Definición 4.18 (n-Cubos (Q_n)) *El cubo n-dimesional, o n-cubo, denotado por Q_n, es el grafo cuyos vértices representan las 2^n cadenas de bits de logitud n. Dos vérties son adyacentes si, y sólo si, las cadenas de bits a las que representan difieren exactamente en un bit. En la Fig. 4.14 se muestran los grafos Q_n para $n = 1,2,3$*

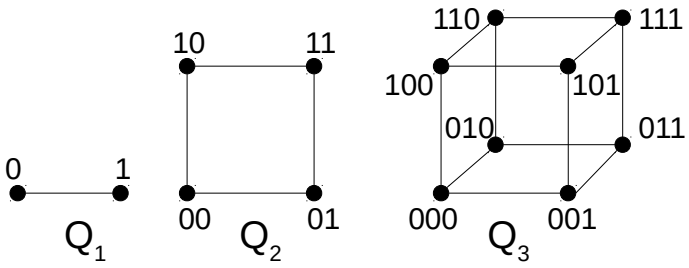


Figura 4.14: Grafos Q_n para $n = 1,2,3$

4.3. Representación de Grafos e Isomorfismos

Hay muchas maneras útiles de representar los grafos. A veces dos grafos tienen exactamente la misma forma, en el sentido de que hay biyección entre sus conjuntos de vértices que preserva las aristas. Es tal caso, decimos que los dos grafos son isomorfos. El determinar si dos grafos son isomorfos o no es un problema importante en la teoría de grafos que estudiaremos en esta sección.

Vértices	Vértices adyacentes
a	b,c,e
b	a
c	a,d,e
d	c,e
e	a,c,d

Tabla 4.1: Lista de adyacencia para el grafo simple *G*

Vértices	Vértices adyacentes
a	b,c,d,e
b	b,d
c	a,c,e
d	
e	b,c,d

Tabla 4.2: Lista de adyacencia para el grafo dirigido *H*

Tabla 4.3: Listas de adyacencia

4.3.1. Represesentación de Grafos

Listas de adyacencia Una forma de representar grafos sin aristas mútiples es utilizar listas de adyacencia, que especifican los vértices que son adyacentes a cada uno de los vértices del grafo.

Ejemplo 4.6 *Utiliza listas de adyacencias para describir los grafos de Fig. 4.15*

En la tabla 4.3 se puede ver las listas de adyacencia correspondientes.

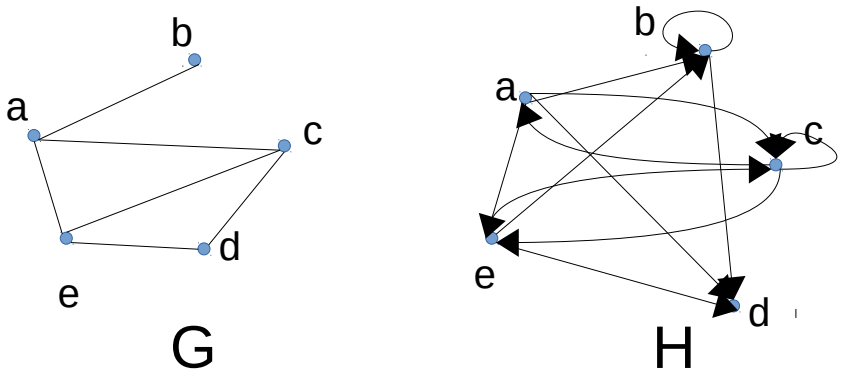


Figura 4.15: Ejemplo listas adyacencia

Matrices de Adyacencia Ejecutar algoritmos para grafos utilizando la representación de los grafos por medio de listas de aristas, o por medio de listas de adyacencia, puede ser complicado si el grafo tiene muchas aristas. Para simplificar los cálculos, conviene representar los grafos por medio de matrices. Presentaremos dos tipos de matrices que se utilizan con frecuencia para representar grafos. Uno se basa en la adyacencia de vértices y el otro se basa en la incidencia entre vértices y aristas.

Supongamos que $G = (V,E)$ es un grafo simple con $|V| = n$. Supongamos que los vértices de G se enumeran de manera arbitraria como $v_1, v_2, \cdots, v_n$. La **matriz de adyacencia A** o $\mathbf{A_G}$ de G con respecto a este listado de los vértices es la matriz booleana $n \times n$ que tiene un 1 en l posición (i, j) si v_i, v_j son adyacentes, y tiene un 0 en la posición (i, j) si no lo son. En otras palabras, la matriz de adyacencia se define como

$$\mathbf{A} = [a_{ij}] \text{ tal que } \quad a_{ij} = \begin{cases} 1 & \text{si } \{v_i, v_j\} \quad \text{es una arista de } G \\ 0 & \text{en caso contrario} \end{cases}$$

Nótese que la matriz de adyacencia de un grafo depende del orden elegido para los vértices. Por tanto, hay $n!$ matrices de adyacencia distintas para un grafo de n vértices.

La matriz de adyacencia de un grafo simple es simétrica, esto es, $a_{ij} = a_{ji}$, puesto que ambos elementos son 1 si dichos vértices son adyacentes y ambos son 0 en caso contrario. Además, como un grafo simple no tiene bucles, cada elemeno $a_{ii} = 0 \quad \forall \quad i = 1, 2, \cdots n$.

Las matrices de adyacencia pueden usarse también para representar grafos no dirigidos con bucles y aristas múltiples. Un bucle en el vértice a_i se representa por medio de un 1 en la posición (i,i) de la matriz de adyacencia. Cuando hay aristas múltiples, la matriz de adyacencia deja de ser booleana, ya que el elemento en la posición (i, j) de esta matriz es igual al número de aristas asociadas con $\{a_i, a_j\}$. Todos los grafos no dirigidos, incluyendo multigrafos y pseudografos tienen matrices de adyacencia simétricas.

Análogamente, la matriz de adyacencia de un grafo dirigido tiene un 1 en su posición (i, j) si hay una arista que va de a_i a a_j. Esta matriz no tiene porqué ser simétrica. Por otra parte, también se pueden utilizar matrices de adyacencia para representar multigrafos dirigidos. De nuevo estas matrices ya no son booleanas si hay aristas múltiples en el mismo sentido conectando dos vértices. En la matriz de adyacencia de un multigrafo dirigido, a_{ij} es el numero de aristas asociadas con (v_i, v_j)

Ejemplo 4.7 *Utiliza una matriz de adyacencia para representar el grafo de fig:4.16*

Para el grafo simple G, ordenando los vértices de la forma a, b, c, d, se tiene la matriz

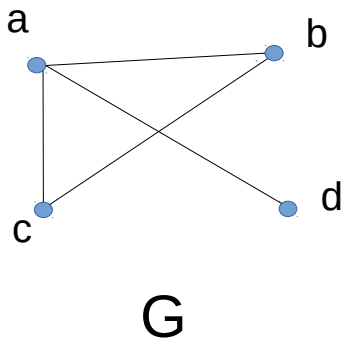


Figura 4.16: Ejemplo matriz de adyacencia en grafo simple

de adyacencia:

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

Ejemplo 4.8 *Dibuja el grafo con la matriz de adyacencia*

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

El grafo correspondiente será

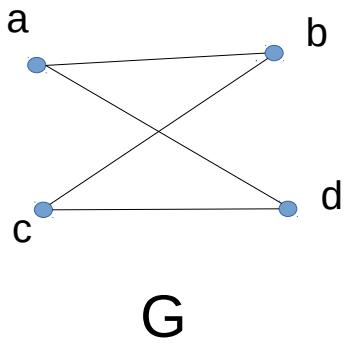


Figura 4.17: Solución construcción grafo a partir de matriz de adyacencia

Ejemplo 4.9 *Utiliza una matriz de adyacencia para representar el pseudografo que se muestra en Fig. 4.18*

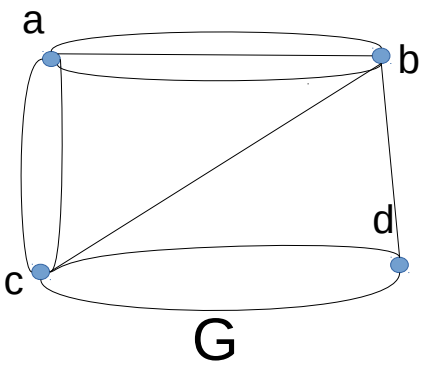


Figura 4.18: Solución construcción grafo a partir de matriz de adyacencia

Para el pseudografo G, ordenando los vértices de la forma a,b,c,d, se tiene la matriz de adyacencia:

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

Matrices de Incidencia Otra representación de grafos que se usa con frecuencia es la que emplea **matrices de incidencia**. Sea G=(V,E) un grafo no dirigido. Supongamos que $v_1, v_2, \cdots v_n$ son los vértices y $e_1, e_2, \cdots, e_m$ las aristas de G. Entonces, la matriz de incidencia con respecto a este ordenamiento de V y de E es la matriz $\mathbf{M} = [m_{ij}]$ de $n \times m$ dada por:

$$m_{ij} = \begin{cases} 1 & \text{si la arista } e_j \quad \text{es incidente con} \quad v_i \\ 0 & \text{en caso contrario} \end{cases}$$

Ejemplo 4.10 *Representa por medio de una matriz de incidencia el grafo que se muestra en 4.19*

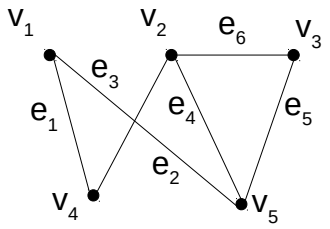


Figura 4.19: Grafo ejemplo matriz incidencia

La matriz de incidencia es:

	e_1	e_2	e_3	e_4	e_5	e_6
v_1	1	1	0	0	0	0
v_2	0	0	1	1	0	1
v_3	0	0	0	0	1	1
v_4	1	0	1	0	0	0
v_5	0	1	0	1	1	0

Tabla 4.4: Matriz de incidencia el grafo 4.19

Las matrices de incidencia pueden usarse también para representar aristas múltiples y bucles. Las aristas múltiples se representan en la matriz de incidencia mediante columnas con todos sus elementos idénticos, puestos que dichas aristas son incidentes con el mismo par de vértices. Los bucles se representan por medio de una columna con un único elemento igual a 1, que corresponde al vértices con el que es incidente el bucle.

Ejemplo 4.11 *Representa por medio de una matriz de incidencia el grafo que se muestra en 4.20*

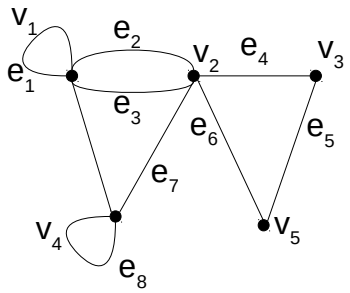


Figura 4.20: Grafo ejemplo matriz incidencia

La matriz de incidencia es:

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8
v_1	1	1	1	0	0	0	0	0
v_2	0	1	1	1	0	1	1	0
v_3	0	0	0	1	1	0	0	0
v_4	0	0	0	0	0	0	1	1
v_5	0	0	0	0	1	1	0	0

Tabla 4.5: Matriz de incidencia el grafo 4.19

4.3.2. Isomorfismos de Grafos

A menudo queremos saber si es posible o no dibujar dos grafos de la misma forma. Disponemos de una terminología muy útil para los grafos que tienen la misma estructura.

Definición 4.19 (Grafos Isomorfos) *Los grafos simples $G_1 = (V_1, E)$ y $G_2 = (V_2, E)$ son isomorfos si hay una función biyectiva f de V_1 en V_2 con la propiedad de que, para cada par de vértices $u, v \in V_1$, u y v son adyacentes en G_1 si y sólo si, $f(u)$ y $f(v)$ son adyacentes en G_2. Se dice que está función f es un isomorfismo.*

Recordamos que una función es biyectiva si es al mismo tiempo inyectiva y sobreyectiva; es decir, si todos los elementos del conjunto de salida tienen una imagen distinta en el conjunto de llegada, y a cada elemento del conjunto de llegada le corresponde un elemento del conjunto de salida.

En otras palabras, cuando dos grafos simples son isomorfos, hay una función biyectiva entre los vértices de los dos grafos que preserva la relación de adyacencia. El isomorfismo de grafos simples es una relación de equivalencia.

Ejemplo 4.12 *Demuestra que los grafos $G = (V, E)$ y $H = (W, F)$ que se muestran en la Fig. 4.21 son isomorfos.*

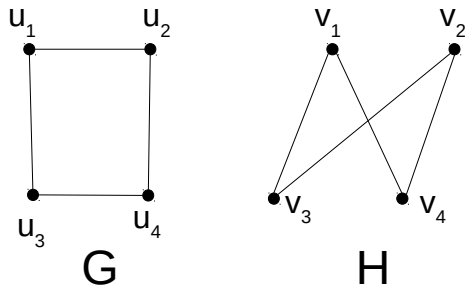


Figura 4.21: Grafo ejemplo isomorfismos

La función f con $f(u_1) = v_1$, $f(u_2) = v_4$, $f(u_3) = v_3$ y $f(u_4) = v_2$, es una función biyectiva entre V y W. Para ver que esta función preserva la adyacencia, nótese que los pares adyacentes en G son $\{u_1, u_2\}, \{u_1, u_3\}, \{u_2, u_4\}$ y $\{u_3, u_4\}$, y cada uno de los pares $\{f(u_1) = v_1, f(u_2) = v_4\}, \{f(u_1) = v_1, f(u_3) = v_3\}, \{f(u_2) = v_4, f(u_4) = v_2\}$ y $\{f(u_3) = v_3, f(u_4) = v_2\}$ son adyacentes en H.

A menudo es díficil determinar si dos grafos simples son o no isomorfos, ya que hay muchas $(n!)$ posibles combinaciones de funciones biyectivas. Sin embargo, con frecuencia se puede demostrar que dos grafos no son isomorfos demostrando que no

comparten alguna propiedad que dos grafos isomorfos deberían tener en común. A tales propiedades se les llama **invariantes** bajo isomorfismo de grafos simples. Por ejemplo, dos grafos simples isomorfos tienen que tener el mismo número de vértices, el mismo número de aristas y los grados de los vértices tienen que coincidir.

Ejemplo 4.13 *En el grafo 4.22, tanto G como H tienen cinco vértices y seis aristas. Sin embargo, H tiene un vértice de grado uno (el vértice e), mientras que G no tiene vértices de grado uno. Por tanto, no son isomorfos.*

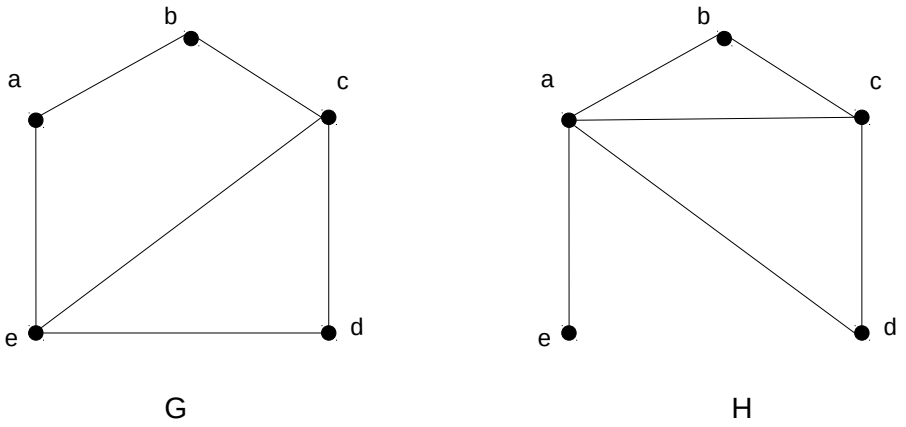


Figura 4.22: Grafo ejemplo isomorfismos

Ejemplo 4.14 *En el grafo 4.23, tanto G como H tienen ocho vértices y diez aristas. Ambos tienen también cuatro vértices de grado dos y cuatro de grado tres. Podríamos pensar que son isomorfos. Sin embargo, no es cierto. En efecto, como $\delta(a) = 2$ en G, a deberia corresponder a uno de los vértices t, u, x o y de H, ya que son los vértices de grado dos de H. Sin embargo, cada uno de estos cuatro vértices de H es adyacente a otro vértice de grado dos de H, lo que no es cierto para a en el grafo G.*

4.4. Grafos Eulerianos y Hamiltonianos

¿Podemos movernos por las aristas de un grafo comenzando en un vértice y volviendo a él después de haber pasado por cada arista del grafo exactamente una vez? Análogamente, ¿podemos desplazarnos por las aristas de un grafo comenzando en un vértice y volviendo a él después de haber visitado cada vértice del grafo exactamente una vez? Aunque estas preguntas parecen similares, la primera de ellas, que pregunta si el grafo contiene lo que se llama un circuito euleriano, puede resolverse fácilmente para cualquier grafo, mientras que la segunda cuestión, la de si el grafo contiene o no lo que se

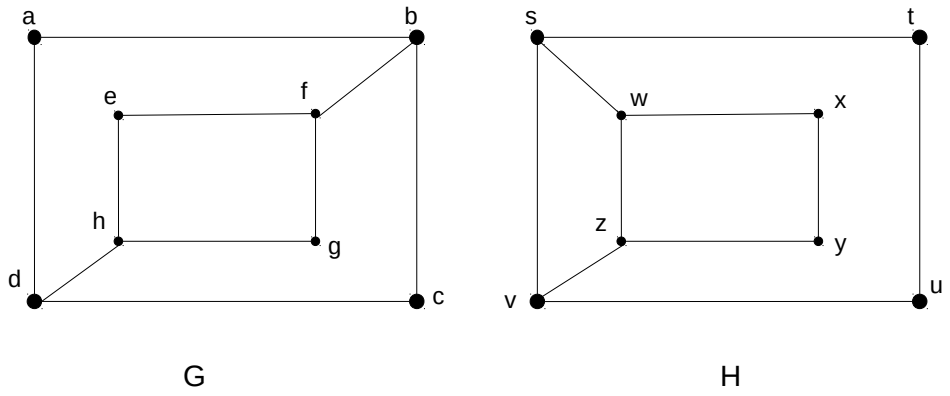


Figura 4.23: Grafo ejemplo isomorfismos

llama un circuito hamiltoniano, es bastante difícil de resolver. En esta sección analizaremos ambas preguntas y estudiaremos las dificultadess que se presentan a la hora de resolverlas.

4.4.1. Caminos

Hay muchos problemas que se pueden representar por medio de caminos que se forman al ir recorriendo las aristas de un grafo. Por ejemplo, el problema de determinar si se puede enviar o no un mensaje entre dos ordenadores usando enlaces intermedios puede estudiarse utilizando un modelo de grafos.

De manera informal, un **camino** es una secuencia de aristas que comienza en un vértice del grafo y recorre ciertas aristas del grafo siempre conectando pares de vértices adyacentes. Formalmente,

Definición 4.20 (Camino de un Grafo:) *Sea n un entero no negativo, y sea G un grafo no dirigido. Un camino de longitud n de u a v en G es una secuencia de aristas* $a_1, a_2, \cdots, a_n$ *de G tal que* $f(a_1) = \{x_0, x_1\}, f(a_2) = \{x_1, x_2\}, \cdots f(a_n) = \{x_{n-1}, x_n\}$, *donde* $x_0 = u$ *y* $x_n = v$.

Si el grafo es simple, denotamos este camino por su secuencia de vértices $x_0, x_2, \cdots, x_n$ (ya que enumerar estos vértices determina el camino de forma única).

Definición 4.21 (Circuito) *El camino es un circuito si comienza y termina en el mismo vértice. Es decir, si* $u = v$, *y tiene longitud mayor que cero.*

Un camino o circuito es simple si no contiene la misma arista más de una vez.

Cuando no es necesario distinguir entre aristas múltiples, denotaremos el camino $a_1, a_2, \cdots, a_n$, donde $f(a_i) = \{x_{i-1}, x_i\}$ para $i = 1, 2 \cdots, n$. Esta notación identifica un camino mediante los vértices por los que pasa. Puede haber más de un camino que pase por esa secuencia de vértices. Nótese que un camino de longitud cero consiste en un único vértice.

Ejemplo 4.15 *En el grafo simple que se muestra en Fig. 4.24, a, d, c, f, e es un camino simple de longitud 4, ya que $\{a,d\}, \{d,c\}, \{c,f\}, \{f,e\}$ son aristas. Sin embargo, d, e, c, a no es un camino puesto que $\{e,c\}$ no es una arista. Nótese que b, c, f, e, b es un circuito de longitud 4. Por otra parte, el camino a, b, e, d, a, b, que tiene longitud 5, no es simple, ya que contiene dos veces la arista $\{a,b\}$*

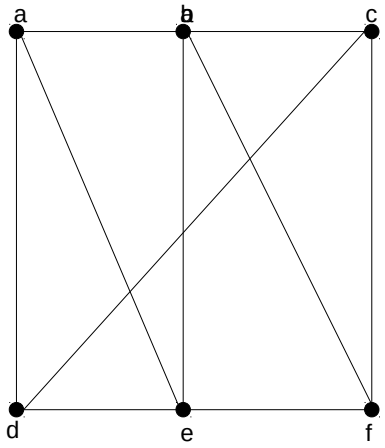


Figura 4.24: Grafo ejemplo camino

4.4.2. Conexión en grafos no dirigidos

¿Cuándo tiene una red informática la propiedad de que dos ordenadores cualesquiera pueden compartir información si pueden enviarse mensajes a través de uno o más ordenadores intermedios? Si se utiliza un grafo para representar esta red informática, con los vértices representando ordenadores y las aristas representando enlaces de comunicación, la pregunta se convierte en la siguiente: ¿Bajo qué condiciones existe siempre un camino entre dos vértices cualesquiera del grafo?

Definición 4.22 *Se dice que un grafo no dirigido es conexo si hay un camino entre cada par de vértices distintos del grafo.*

Por tanto, cada dos ordenadores de la red pueden comunicarse entre sí si y sólo si, el grafo de esta red es conexo.

Ejemplo 4.16 *El grafo G_1 de la Fig. 4.25 es conexo, ya que para cada par de vértices distintos hay un camino entre ellos. Sin embargo, el grafo G_2 no es conexo al no haber ningún camino entre a y d.*

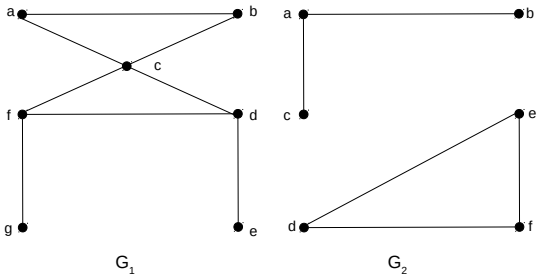


Figura 4.25: Grafo ejemplo conexo

Teorema 4.4 *Hay un camino simple entre cada par de vértices distintos de un grafo no dirigido conexo.*

4.4.3. Conexión en grafos dirigidos

Si en lugar de no dirigido estamos trabajando con grafos dirigidos, se distinguen dos nociones de conexión:

Definición 4.23 *Se dice que un grafo dirigido es fuertemente conexo si hay un camino de a a b y un camino de b a a para cualesquiera dos vértices de a y b del grafo.*

Para que un grafo dirigido sea fuertemente conexo tiene que haber una secuencia de aristas dirigidas desde cualquier vértice del grafo a cualquier otro vértice. Un grafo dirigido puede no ser fuertemente conexo, pero estar formado por una sola pieza. Es decir:

Definición 4.24 *Se dice que un grafo dirigido es débilemente conexo si hay un camino entre cada dos vértices del grafo no dirigido subyacente.*

Esto es, un grafo dirigido es débilmente conexo si, y sólo si, hay siempre un camino entre dos vértices cuando se ignoran las direcciones de las aristas. Claramente, cualquier grafo dirigido fuertemente conexo es también débilmente conexo.

Ejemplo 4.17 *¿Son fuertemente conexos los grafos dirigidos G y H de la Fig. 4.26 ¿Son débilmente conexos?*

G es fuertemente conexo porque hay un camino entre cada dos vértices de este grafo dirigido. Por tanto, G es también débilmente conexo. El grafo H no es fuertemente conexo. No hay ningún camino dirigido entre a y b en este grafo. Sin embargo, H es débilmente conexo, ya que hay un camino entre cada dos vértices en el grafo no dirigido subyacente a H.

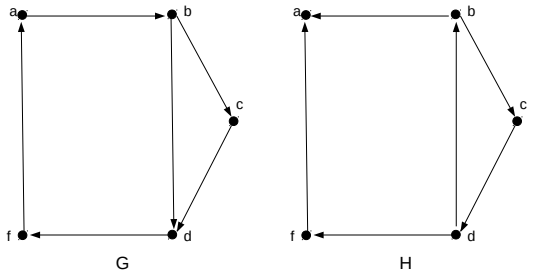


Figura 4.26: Grafo Fuermente Conexo

A aquellos subgrafos de un grafo dirigido *G* que son fuertemente conexos, pero que no están contenidos en ningún subgrafo fuertemente conexo mayor, se les llama **componentes fuertemente conexas** o **componentes fuertes** de *G*.

Ejemplo 4.18 *El grafo H de la Fig. 4.26 tiene tres componentes fuertemente conexas. La primera consta del vértice a. La segunda consta del vértice e. La tercera es el grafo formado por los vértices b, c y d y por las aristas* (b,c), (c,d) *y* (d,b).

Hay muchas maneras en las que los caminos y circuitos pueden ayudarnos a determinar si dos grafos son o no isomorfos. Por ejemplo, la existencia de un circuito simple de una longitud concreta es un invariante útil que se puede emplear a la hora de mostrar que dos grafos son isomorfos.

Ejemplo 4.19 *Determina si los grafos G y H de 4.27 son o no isomorfos.*

Tanto G como H tienen seis vértices y ocho aristas. Ambos tienen cuatro vértices de grados tres y dos vértices de grado dos. Por tanto, los tres invariantes (número de vértices, número de aristas y grado de los vértices) coinciden en los dos grafos. Sin embargo, H contiene un circuito simple de longitud tres (v_1, v_2, v_6, v_1)*, mientras que G no contiene ningún circuito simple de longitud tres. Como la existecia de un circuito simple de longitud tres es un invariante bajo isomorfismo.*

Ejemplo 4.20 *Determina si los grafos G y H de 4.28 son o no isomorfos.*

Tanto G como H tienen cinco vértices y seis aristas. Ambos tienen dos vértices de grado tres y tres vértices de grado dos, y ambos contienen un circuito simple de longitud

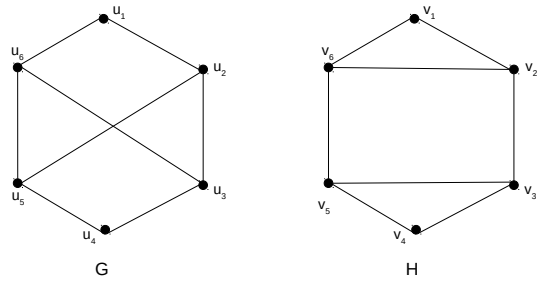


Figura 4.27: Grafo ejemplo caminos e isomorfismo

tres, un circuito simple de longitud cuatro y un circuito simple de longitud cinco. Como todos estos invariantes coinciden, G y H pueden ser isomorfos. En efecto existe f de la forma $f(u_1) = v_3, f(u_4) = v_2, f(u_3) = v_1, f(u_2) = v_5$, y $f(u_5) = v_4$. Se puede demostrar que f es un isomorfismo demostrando que f mantiene las adyacencias o bien demostrando que, para ordenaciones adecuadas de los vértices, las matrices de adyacencia G y de H son iguales.

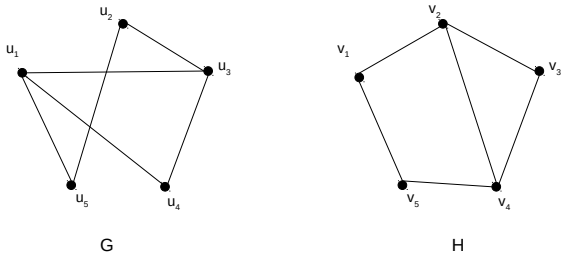


Figura 4.28: Grafo ejemplo camino e isomorfismo

4.4.4. Número de caminos entre vértices

Por otra parte, el número de caminos que hay entre dos vértices de un grafo se puede determinar usando su matriz de adyacencia.

Teorema 4.5 *Sea G un grafo y sea* **A** *su matriz de adyacencia con respecto a la ordenación $v_1, v_2, \cdots, v_n$ (se admiten aristas dirigidas o no dirigidas, aristas múltiples o bucles). El número de caminos distintos de longitud r de v_i a v_j, siendo r un entero positivo, es igual al elemento en la posición (i, j) de la matriz $\mathbf{A}^r$.*

Ejemplo 4.21 *Cuántos caminos de longitud 4 hay entre a y d en el grafo simple de la Fig. 4.29*

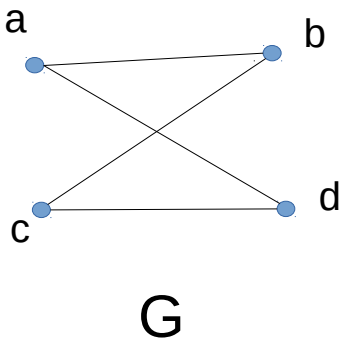


Figura 4.29: Grafo ejemplo camino e isomorfismo

La matriz de adyacencia de G, ordenado los vértices de la forma a,b,c,d es:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Por tanto, el número de caminos de lontigud 4 entre a y d es el elemento (1,4) de $\mathbf{A}^4$. Como

$$\mathbf{A}^4 = \begin{pmatrix} 8 & 0 & 0 & 8 \\ 0 & 8 & 8 & 0 \\ 0 & 8 & 8 & 0 \\ 8 & 0 & 0 & 8 \end{pmatrix}$$

hay examente 8 caminos de longitud 4 entre a y d.

4.4.5. Caminos y circuitos Eulerianos

La ciudad prusiana de Königsberg (hoy en díua Kaliningrado) estaba dividida en cuatro partes por los dos brazos en los que se bifurca el río Pregel (ver Fig. 4.30). Esta cuatro partes eran las dos regiones a orilla del río Pregel, la isla de Kneiphof y la región que quedaba entre ambos brazos del Pregel. Siete puentes conectaban entre sí estas regiones en el siglo XVIII.

Los habitantes de Königsberg solían dar largos paseos por la ciudad los domingos. Hubo quien se preguntó si sería posible comenzar el paseo en algún sitio de la ciudad, atravesar todos los puentes sin cruzar ninguno dos veces y regresar al punto de partida.

El matemático suizo Leonhard Euler resolvió este problema. Euler estudió el problema

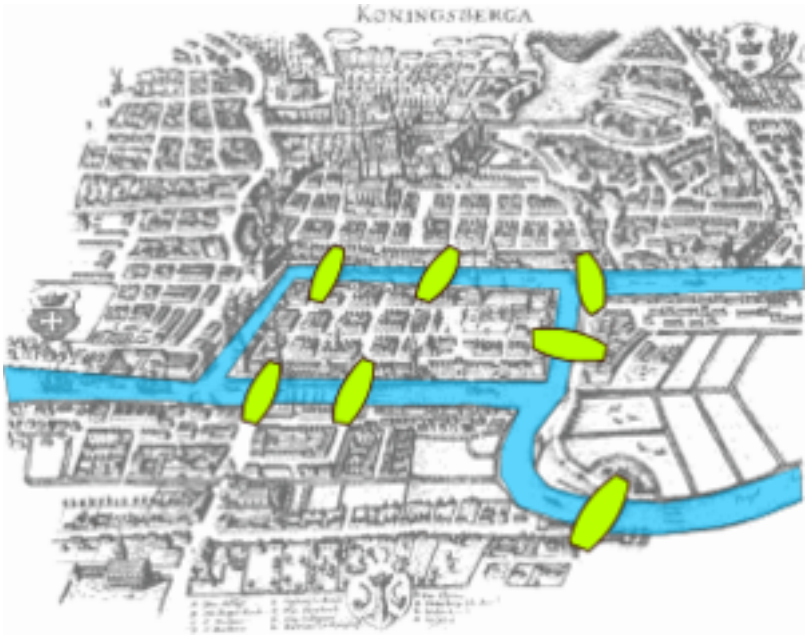


Figura 4.30: Los sieta puentes de Königsberg

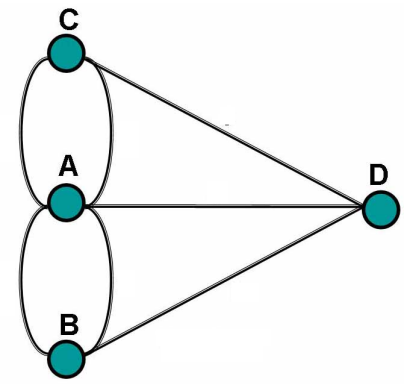


Figura 4.31: Modelo en forma de multigrafo de Königsberg

usando el multigrafo que se obtiene si se representan las cuatro regiones mediante vértices y los siete puentes mediante aristas (ver Fig. 4.31).

El problema de recorrer todos los puentes sin cruzar ninguno de ellos más de una vez puede replantearse en términos de este modelo. La pregunta se convierte entonces en la de si hay o no algún circuito simple en el multigrafo que contenga a todas las aristas del grafo.

Definición 4.25 (Circuito Euleriano) *Un circuito euleriano de un grafo G es un circuito simple que contiene a todas las aristas de G.*

Definición 4.26 (Camino Euleriano) *Un camino euleriano de un grafo no dirigido G es un camino simple que contiene a todas las aristas de G.*

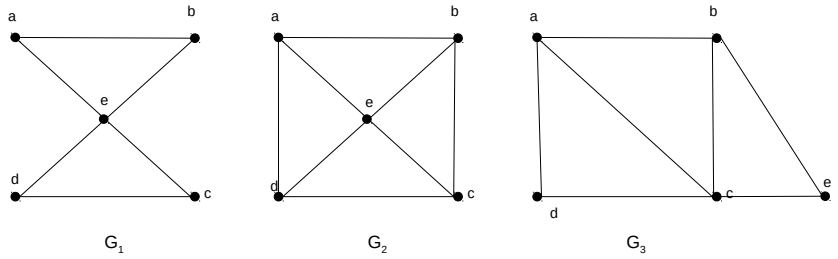


Figura 4.32: Grafos no dirigidos

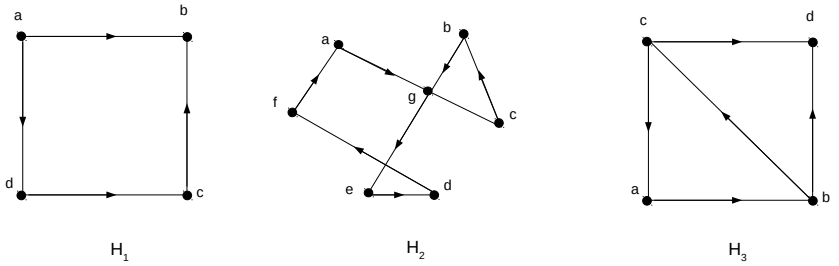


Figura 4.33: Grafos dirigidos

Ejemplo 4.22 *¿Cuáles de los grafos no dirigidos en 4.32 contienen un circuito euleriano? Entre aquellos que no lo contienen, ¿Cuáles contienen un camino euleriano?*

El grafo G_1 contiene un circuito euleriano, por ejemplo, a,e,c,d,e,b,a. Ni G_2 ni G_3 contienen un circuito euleriano. No obstante, G_3 contiene un camino euleriano, a saber, a,d,c,e,b,c,a,b. El grafo G_2 no contiene nigún camino euleriano.

Ejemplo 4.23 *¿Cuáles de los grafos dirigidos en 4.33 contienen un circuito euleriano? Entre aquellos que no lo contienen, ¿Cuáles contienen un camino euleriano?*

El grafo H_2 contiene un circuito euleriano, por ejemplo a,g,c,b,g,e,d,f,a. Ni H_1 ni H_3 contienen un circuito euleriano. El grafo H_3 contiene un camino euleriano, a saber, c,a,b,c,d,b, pero H_1 no.

Definición 4.27 (Grafo Euleriano) *:Un grafo $G=(V,E)$ es euleriano si contiene un circuito euleriano.*

Nótese o G es conexo, o bien hay una componente conexa que contiene todas las aristas de G y el resto son vértices aislados.

Teorema 4.6 *Un grafo $G(V;E)$ no dirigido es euleriano si y sólo si todas las aristas están en la misma componente conexa y todos los vértices tienen grado par.*

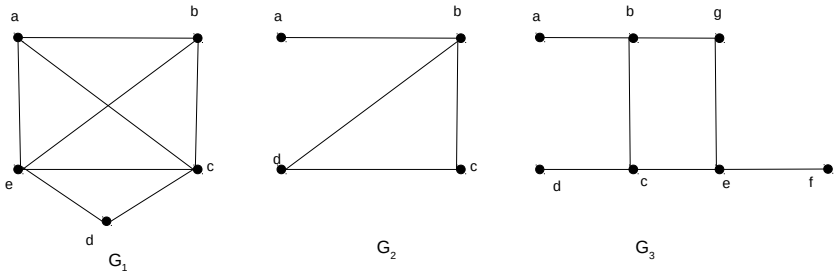


Figura 4.34: Circuitos Hamiltonianos

Teorema 4.7 *Un grafo $G(V,E)$ no dirigido de manera que todas sus aristas están en la misma componente conexa, admite un camino euleriano si y sólo si contiene exactamente dos vértices de grado impar.*

Ya podemos resolver el problema de los puentes de Könisberg. Dado que el multigrafo que representa a estos puentes tiene cuatro vértices de grado impar, no contiene ningún circuito euleriano. No hay ninguna forma de comenzar en un punto, cruzar el puente exactamente una vez y regresar al punto de partida.

4.4.6. Caminos y circuitos Hamiltonianos

Hemos obtenido condiciones necesarias y suficientes para la existencia de caminos y circuitos que contienen a cada una de las aristas de un multigrafo exactamente una vez. ¿Podemos hacer lo mismo con caminos y circuitos simples que pasen exactamente una vez por cada vértice el grafo?

Definición 4.28 (Camino Hamiltoniano) *Se dice que un camino $x_0, x_1, \cdots, x_n$ del grafo $G = (V,E)$ es un camino hamiltoniano si $V = \{x_0, x_1, \cdots, x_n\}$ y $x_i \neq x_j$ para $0 \leq i < j \leq n$.*

Definición 4.29 (Circuito Hamiltoniano) *Se dice que un circuito $x_0, x_1, \cdots, x_n, x_0$ $(n > 1)$ del grafo $G = (V,E)$ es un circuito hamiltoniano si $x_0, x_1, \cdots, x_n$ es un camino hamiltoniano.*

Ejemplo 4.24 *¿Cuál de los grafos simple de 4.34 contienen un circuito o un camino hamiltoniano?*

G_1 contiene un circuito hamiltoniano a,b,c,d,e,a. No hay circuitos hamiltonianos en G_2 ya que cualquier circuito que pase por todos los vértices tiene que contener dos veces la arista $\{a,b\}$. Sin embargo G_2 contiene un camino hamiltoniano, que es a,b,c,d. El grafo G_3 no contiene ni un circuito hamiltoniano ni un camino hamiltoniano, ya que

cualquier camino que pase por todos los vértices tiene que contener más de una vez una de las aristas $\{a,b\},\{e,f\},\{c,d\}$

Dada la cercanía entre las definiciones de circuitos eulerianos y ciclos hamiltonianos, y dado que el criterio de Euler nos proporciona un modo sencillo de decidir si un grafo admite o no un circuito euleriano, se podría esperar que hubiese un modo sencillo de decidir si un grafo admite o no un ciclo hamiltoniano. Sin embargo, la situación para el caso hamiltoniano es dŕasticamente diferentes a la del caso euleriano.

En efecto, *no se conoce ninguna condición necesaria y suficiente (sencilla de aplicar) para que un grafo sea hamiltoniano. Y tampoco se conoce ningún algoritmo eficiente para buscar un ciclo hamiltoniano en un grafo hamiltoniano.*

No obstante, sí se conocen teoremas que dan condiciones suficientes para que un grafo sea hamiltoniano:

Teorema 4.8 (Teorema de Dirac) *Sea G un grafo simple con n vértices para* $n \geq 3$ *tal que todos los vértices de G tienen grado mayor o igual que* $n/2$*. Entonces G contiene un circuito hamiltoniano*

Ejemplo 4.25 *El grafo completo* K_6 *es un grafo con 6 vértices en el que cada uno de ellos está unido con los 5 vértices restantes. Por tanto,* $\delta(v) = 5$ *para todo v. Como* $5 \geq 3$*, el teorema anterior nos asegura que el grafo completo* K_6 *es hamiltoniano. De hecho, todo grafo completo* K_n *con* $n \geq 3$ *es hamiltoniano.*

Parte II

Álgebra

Universidad Rey Juan Carlos

Tema 5

Cálculo Matricial

5.1. Cálculo Matricial

En este tema introducimos las matrices y las operaciones con ellas, pues constituyen el lenguaje adecuado para abordar cuestiones de naturaleza lineal. Entre estas, la más elemental es la discusión de sistemas de ecuaciones lineales, para lo que previamente es imprescindible caracterizar las matrices invertibles en términos de su determinante y analizar la noción de rango de una matriz.

5.1.1. Ejemplos

Presentamos a continuación ejemplos sencillos que ponen de manifiesto cómo el lenguaje matricial es el adecuado para plantear y resolver algunas cuestiones elementales de naturaleza lineal.

Las primeras ecuaciones que uno aprende a resolver son de la forma $ax = b$, donde a y b denotan número reales dados, y x, otro número real, que se pretende calcular. A veces, problemas de la vida cotidiana o de las ciencias experimentales nos abocan a abordar ecuaciones con más de una incógnita. Por ejemplo, si $a_{11}, a_{12}, a_{21}, a_{22}, b_1$ y b_2 son números reales, pretendemos decidir si existen otros x e y tales que

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases}$$

y, en caso de que existan, averiguar si son únicos o no y, finalmente, calcular todos ellos. Las matrices, que introduciremos a continuación, nos permiten tratar el sistema de ecuaciones como la ecuación $ax = b$.

5.2. Definiciones

Definición 5.1 (Matriz:) *Una matriz de orden $m \times n$ es una tabla de números del cuerpo $\mathbb{K}$ formada por m filas y n columnas. El elemento o coeficiente de una matriz A que está en la fila i y en la columna j se denota a_{ij}, y diremos que es su coeficiente (i, j). Escribiremos: $A = (a_{ij})$ con $1 \leq i \leq m, 1 \leq j \leq n$*

Ejemplo 5.1

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 7 & 5 \end{pmatrix}$$

es una matriz de orden 2×3 (es decir, tiene 2 filas y 3 columnas) y $a_{13} = 3$ es el elemento que se encuentra en la primera fila y tercera columna.

Dos matrices son iguales si tienen igual orden e iguales elementos en cada una de sus posiciones.

Ejemplo 5.2 *Las matrices:*

$$A = \begin{pmatrix} a & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \; y \; B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & b \end{pmatrix}$$

son iguales si y sólo si a=1 y b=6. Por otra parte, las matrices:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \; y \; B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

no son iguales puesto que tienen orden distinto.

De una matriz con una única fila diremos que es una **matriz fila** y de una matriz con una única columna diremos que es una **matriz columna**. Una matriz es **cuadrada** si tiene igual número de filas que de columnas, esto es: si es de orden $n \times n$ para algún n.

Al conjunto de todas las matrices de orden $m \times n$ con coeficientes en el cuerpo $\mathbb{K}$ lo denotaremos por $\mathcal{M}_{m\times n}(\mathbb{K})$. Para el caso en que $m = n$, escribiremos simplemente $\mathcal{M}_n(\mathbb{K})$.

Ejemplo 5.3 $\mathcal{M}_3(\mathbb{R})$ *designará el conjunto de todas las matrices cuadradas de orden 3 con coeficientes en el cuerpo* $\mathbb{R}$ *de números reales.*

$\mathcal{M}_{2\times 3}(\mathbb{Q})$ *es el conjunto de todas las matrices de orden* 2×3 *con coeficientes en el cuerpo* $\mathbb{Q}$ *de números racionales.*

Definición 5.2 (Matriz traspuesta:) *La traspuesta de una matriz* $A = (a_{ij})$ *es la matriz* $A^t = (b_{ij})$ *definida por* $b_{ij} = a_{ji}$*. En palabras, las filas de la matriz traspuesta son las columnas de la matriz dada.*

Definición 5.3 (Matriz simétrica y Matriz antisimétrica:) *Una matriz que coincide con su traspuesta se llama simétrica, mientras que si coincide con la opuesta de su traspuesta se llama antisimétrica. Para que se dé cualquiera de las dos situaciones anteriores es necesario que la matriz sea cuadrada.*

Definición 5.4 (Submatriz:) *Dada una matriz A, llamaremos submatriz de A a cada matriz que se obtenga de A suprimiendo algunas de sus filas o columnas.*

Ejemplo 5.4 *La matriz A es submatriz de B*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \; ; \; B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

Definición 5.5 (Matriz diagonal:) *Dada una matriz cuadrada $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$, los elementos $a_{11}, a_{22}, \cdots, a_{nn}$ constituyen su **diagonal principal**. Se dice que A es una **matriz diagonal** si todos los elementos fuera de sus diagonal principal son cero (esto es: $a_{ij} = 0 \quad \forall i \neq j$):*

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

Definición 5.6 (Matriz triangular:) *A es **triangular superior** si todos los elementos por debajo de su diagonal son 0 ($a_{ij} = 0, \forall i > j$) y **triangular inferior** si todos los elementos por encima de su diagonal son 0 ($a_{ij} = 0, \forall i < j$).*

Triangular superior *Triangular inferior*

$$A_S = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \quad A_I = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Definición 5.7 (Matriz identidad de orden n:) *Matriz donde $m = n$ y los coeficientes a_{ij} son nulos si $i \neq j$ y valen 1 para $i = j$. Habitualmente, para definir la matriz identidad se utiliza la denominada delta de Krönecker dada por:*

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Así, $I_n = (\delta_{ij})$

Ejemplo 5.5 *Una matriz identidad de orden 5 sería:*

$$I_5 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Definición 5.8 (Matriz nula:) *Se denota por 0 y se llama nula la matriz cuyos coeficientes son todos nulos.*

Definición 5.9 (Matrices escalonadas reducidas:) *Sea $A \in \mathcal{M}_{mxn}(\mathbb{K})$ una matriz. LLamaremos pivote de una fila (o columna) de A al primer elemento no nulo de dicha fila (o columna), si es que hay alguno. La matriz A se dice que es* ***escalonada por filas*** *si verifica las siguientes condiciones:*

1. *Si A tiene filas compuestas enteramente por ceros (filas nulas), éstas están agrupadas en la parte inferior de la matriz.*
2. *El pivote de cada fila no nula es 1.*
3. *El pivote de cada fila no nula está a la derecha del de la fila anterior,*
4. *Los elementos que aparecen en la misma columna que el pivote de una fila y debajo de él, son todos cero.*

A es ***escalonada reducida por filas*** *si además de ser escalonada verifica que: los elementos que aparecen en la misma columna que el pivote de una fila son todos cero.*

Ejemplo 5.6 *Consideremos las siguientes matrices en las que se han recuadrado los pivotes de cada fila:*

$$A = \begin{pmatrix} \boxed{2} & 0 & 0 & 5 \\ 0 & \boxed{1} & 0 & -2 \\ 0 & 0 & \boxed{1} & 4 \end{pmatrix} \quad ; \quad B = \begin{pmatrix} \boxed{1} & 0 & 0 & 5 \\ 0 & \boxed{1} & 0 & -2 \\ 0 & \boxed{1} & 1 & 4 \end{pmatrix}$$

$$C = \begin{pmatrix} \boxed{1} & 0 & 0 & 5 \\ 0 & \boxed{1} & 1 & -2 \\ 0 & 0 & \boxed{1} & 4 \end{pmatrix} \quad ; \quad D = \begin{pmatrix} \boxed{1} & 0 & 0 & 5 \\ 0 & \boxed{1} & 0 & -2 \\ 0 & 0 & \boxed{1} & 4 \end{pmatrix}$$

A no es escalonada por filas puesto que el pivote de la primera fila no es 1. B no es escalonada por filas puesto que los pivotes de la segunda y tercera fila no están uno a la derecha del otro. C es escalonada por filas pero no reducida, ya que en la tercera columna deberían ser cero todos los elementos salvo el pivote de la tercera fila. Por último D es escalonada reducida por filas.

De forma análoga se definiría la matriz escalonada y escalonada reducida por columnas.

Nuestro siguiente objetivo será transformar de alguna forma cada matriz en una escalonada reducida. Para ello, establecemos primero qué tipos de transformaciones consideramos válidas. Para simplificar, utilizaremos el nombre de escalares para referirnos a los elementos del cuerpo que se esté considerando.

Llamaremos **transformaciones elementales** de filas a cada una de las de los siguientes tipos:

1. Tipo I: Intercambiar la posición de dos filas.
2. Tipo II: Multiplicar todos los elementos de una fila por un escalar no nulo.
3. Tipo III: Sumar a una fila otra multiplicada por un escalar.

Diremos que una fila (respectivamente columna) de una matriz es **combinación lineal** del resto de filas si se obtiene mediante las transformaciones elementales de tipo II y III del resto de filas. Las filas que no son combinación lineal del resto se dicen **independientes.**

Definición 5.10 (Matrices equivalentes:) *Diremos que dos matrices A y B son equivalentes por filas* $(A \sim_f B)$ *si podemos transformar la primera en la segunda mediante una cantidad finita de operaciones elementales por filas. Se puede decir que dos matrices son equivalentes, si son equivalentes por filas.*

Lema 5.1 *Para cualesquiera matrices A, B y C en* $\mathcal{M}_{mxn}(\mathbb{K})$*, se verifica:*

1. $A \sim_f A$. *Es decir, una matriz es equivalente consigo misma.*
2. $A \sim_f B \Longleftrightarrow B \sim_f A$. *Es decir, si A es equivalente a B, entonces B es equivalente a A.*
3. *Si* $A \sim_f B$ *y* $B \sim_f C$, *entonces* $A \sim_f C$. *Es decir, si A es equivalente a B y B es equivalente a C; entonces A es equivalente a C.*

Lema 5.2 *Sean A y B dos matrices escalonadas reducidas por filas. Si A y B son equivalentes por filas, entonces A=B.*

Teorema 5.1 *Cada matriz es equivalente por filas a una única matriz escalonada reducida por filas.*

Definición 5.11 (Forma normal de Hermite:) *Dada una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ llamaremos* forma normal de Hermite *por filas (o forma escalonada reducida por filas) de A a la única matriz escalonada reducida por filas que se obtiene de A por transformaciones elementales de filas. De forma similar, se define la forma normal de Hermite por columnas de una matriz, y se prueba su existencia y unicidad.*

Definición 5.12 (Escalamiento de matrices:) *Toda matriz no nula es equivalente por filas a una matriz escalonada reducida. Para probarlo, basta aplicar a la matriz A un algoritmo similar al de Gauss-Jordan:*

1. *Se hace $k = 1$*
2. *Si todas las filas a partir de la k-ésima inclusive son nulas, hemos terminado.*
3. *Si no, se busca en ellas un coeficiente no nulo $a_{i_k j_k}$ con j_k mínimo.*
4. *Se divide la fila i_k-ésima por su primer coeficiente no nulo, que es $a_{i_k j_k}$, y se intercambia con la k-ésima fila.*
5. *A cada fila distinta de la nueva k-ésima, le restamos ésta multiplicada por el coeficiente j_k-ésimo de aquélla.*
6. *Se hace $k = k+1$ y se vuelve al paso 1.*

El algoritmo va cambiando la matriz fila a fila. Como cada modificación que se hace en el proceso es una operación elemental, obtenemos una matriz equivalente por filas a la de partida. En el paso 1 se mira si en realidad hay algo que hacer. Si lo hay, en el paso 2 se busca una fila que deba ir la k-ésima atendiendo a su primer coeficiente no nulo. En el paso 3 se hace 1 ese coeficiente, y se coloca la fila efectivamente en el lugar k-ésimo. En el paso 4 se hacen nulos todos los coeficientes que hay por encima y por debajo de ése no nulo en su misma columna. Y en el paso 5 se recomienza.

Ejemplo 5.7 *Vamos a encontrar un matriz escalonada reducida equivalente por filas a la matriz:*

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 7 & 5 \\ 2 & 6 & 8 \end{pmatrix}$$

De nuevo, aligeramos la descripción de cómo se aplica el algoritmo. Como el coeficiente $a_{11} \neq 0$ ya vale 1, restamos a las filas segunda y tercera la primera multiplicada por 3 y 2, respectivamente, y así se obtiene la matriz equivalente por filas:

$$A_1 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 6 \end{pmatrix}$$

Si ahora restamos a la primera y a la tercera fila de A_1 el doble de la segunda, obtenemos otra matriz equivalente por filas:

$$A_2 = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$$

A continuación, dividimos la tercera fila entre 2:

$$A_3 = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

Finalmente, sumamos a la primera fila el triple de la tercera, y a la segunda le restamos el doble, y se obtiene la matriz equivalente por filas:

$$A_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

que es una matriz escalonada reducida.

Definición 5.13 (Matriz ampliada:) *Si $Ax^t = b^t$ es la escritura matricial de un sistema lineal, donde $A = (a_{ij})$ es la matriz de coeficientes de las incógnitas del sistema. Para tener en cuenta también los términos independientes en las manipulaciones con filas, consideramos la matriz ampliada:*

$$\widetilde{A} = \left(\begin{array}{ccccc|c} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} & b_1 \\ \cdots & \ddots & \cdots & \ddots & \cdots & \cdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} & b_i \\ \cdots & \ddots & \cdots & \ddots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} & b_m \end{array}\right)$$

Definición 5.14 (Rango de una matriz:) *Se llama rango (por filas) de una matriz A, y se denota rg(A), al número de filas no nulas de la única matriz escalonada reducida equivalente por filas a A, que es igual al máximo número de filas independientes de A.*

Ejemplo 5.8 *Pretendemos calcular el rango de la matriz :*

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 7 & 2 \\ 2 & 6 & 0 \end{pmatrix}$$

- *En primer lugar, buscamos una matriz escalonada reducida equivalente por filas de A. Restamos a las filas segunda y tercera el triple y el doble de la primera, respectivamente, lo que nos proporciona una matriz equivalente por filas:*

$$A' = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 2 & -2 \end{pmatrix}$$

Restando el doble de la segunda fila a las filas primera y tercera se tiene:

$$A'' = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

que es escalonada reducida y tiene 2 filas no nulas. Por tanto, rg(A)=2.

- *Según la definición anterior, en la matriz A hay dos filas independientes pero no 3. El proceso desarrollado en el punto anterior nos permite hallar una relación de dependencia entre las filas de A. Si llamamos f_1, f_2, f_3 a las filas de A, en la primera etapa hemos obtenido que $f_1' = f_1, f_2' = f_2 - 3f_1$ y $f_3' = f_3 - 2f_1$ son las de A'. En la segunda etapa hemos comprobado que la tercera fila $f_3'' = f_3' - 2f_2'$ de A" es la fila nula. En consecuencia:*

$$\begin{aligned} f_3' = 2f_2' \Leftrightarrow \\ f_3 - 2f_1 = 2(f_2 - 3f_1) \Leftrightarrow \\ 4f_1 - 2f_2 + f_3 = 0 \end{aligned}$$

y por tanto $f_3 = 2f_2 - 4f_1$ nos da una relación entre las filas de A.

- *Calculemos ahora el rango de A utilizando una matriz equivalente por filas que no es escalonada reducida, pero en la que es inmediato determinar el número de filas independientes. Recordamos que la matriz*

$$A' = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 2 & -2 \end{pmatrix}$$

es equivalente por filas a A. Claramente sus dos últimas filas son proporcionales y las dos primeras independientes. Por tanto, $2 \leq rg(A') < 3$ y así, rg(A)=rg(A')=2.

En consecuencia, se tienen definidos dos rangos: uno por filas y otro por columnas. Naturalmente, la pregunta crucial es qué relación existe entre ambos, y la respuesta es la mejor posible:

Proposición 5.1 *El rango por filas de una matriz coincide con su rango por columnas, o en otras palabras, el máximo número de filas independientes es igual al máximo número de columnas independientes.*

La igualdad de los rangos por filas y columnas que acabamos de establecer se expresa simplemente usando la matriz traspuesta:

$$rg(A) = rg(A^t)$$

Deducimos, además, que el rango es menor o igual que el número de columnas y que el número de filas. Cuando el rango de una matriz coincide con el menor de esos dos números se dice que la matriz tiene *rango máximo*

Definición 5.15 (Matriz inversa:) *Una matriz cuadrada A de orden n se llama invertible si existe otra C tal que $CA = AC = I_n$. Entonces C es única, se llama matriz inversa de A y se denota A^{-1}.*

Está claro que no toda matriz cuadrada tiene inversa, como muestra el siguiente ejemplo:

Ejemplo 5.9 *La matriz:*

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

no se puede tener inversa puesta que al multiplicar A por cualquier otra matriz cuadrada de orden 2 se tiene:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$$

que en ningún caso puede ser la identidad

Ejemplo 5.10 *La matriz*

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

es invertible y su inversa es

$$A^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

ya que

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

y

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

Proposición 5.2 *Dadas $A, B \in \mathcal{M}_{m\times n}(\mathbb{K})$ se verifica:*

1. *Si A y B son invertibles, entonces AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$*
2. *Si $A_1, \cdots, A_n$ son invertibles, entonces $A_1 \cdots A_n$ es invertible y $(A_1 \cdots A_n)^{-1} = A_n^{-1} \cdots A_1^{-1}$*
3. *Si A invertible, entonces A^t es invertible y $(A^t)^{-1} = (A^{-1})^t$*

Proposición 5.3 *Una matriz cuadrada es invertible si y sólo si tiene rango máximo*

5.3. Operaciones Matrices

5.3.1. Suma de Matrices

Dadas dos matrices de igual orden $m \times n, A = (a_{ij})$ y $B = (b_{ij})$, se define su suma como la matriz de orden también $m \times n$ dada por:

$$A + B = (a_{ij} + b_{ij}) = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

Esto es, la suma de matrices A y B es la matriz que en la posición ij tiene al elemento $a_{ij} + b_{ij}$ suma de los correspondientes elementos de A y B. Nótese que la suma de matrices sólo está definida para matrices de igual orden.

Ejemplo 5.11

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

Esta operación verifica las siguientes propiedades:

- **Asociativa**:$A+(B+C)=(A+B)+C$; $\forall A,B,C \in \mathcal{M}_{m\times n}(\mathbb{K})$
- **Conmutativa**: $A+B=B+A$; $\forall A,B \in \mathcal{M}_{m\times n}(\mathbb{K})$
- **Elemento Neutro**: $\exists 0 \in \mathcal{M}_{m\times n}(\mathbb{K}) / A+0=0+A=A$; $\forall A \in \mathcal{M}_{m\times n}(\mathbb{K})$
- **Elementos simétricos:**
  $\forall A \in \mathcal{M}_{m\times n}(\mathbb{K}); \exists -A \in \mathcal{M}_{m\times n}(\mathbb{K}) / A+(-A)=(-A)+A=0.$

.

5.3.2. Producto de un escalar por una matriz

Dada una matriz $A=(a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ y dado un escalar $\alpha \in \mathbb{K}$, se define su producto como la matriz de orden $m \times n$:

$$\alpha A = A\alpha = (\alpha a_{ij}) = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{21n} \\ \cdots & \cdots & \ddots & \cdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{pmatrix}$$

Ejemplo 5.12 *Si consideramos la matriz:*

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

entonces:

$$2A = \begin{pmatrix} 2 & 0 & 10 \\ 6 & 10 & -2 \\ 4 & 2 & 0 \end{pmatrix} \quad (-0{,}5)A = \begin{pmatrix} -0{,}5 & 0 & -2{,}5 \\ -1{,}5 & -2{,}5 & 0{,}5 \\ -1 & -0{,}5 & 0 \end{pmatrix}$$

Propiedades del producto por escalares

1. Distributiva respecto de la suma de escalares:
   $(\alpha+\beta)A = \alpha A + \beta A$; $\quad \forall \alpha,\beta \in \mathbb{K}; \forall A \in \mathcal{M}_{m\times n}(\mathbb{K})$
2. Distributiva respecto de la suma de matrices:
   $\alpha(A+B) = \alpha A + \alpha B$; $\quad \forall \alpha \in \mathbb{K}; \forall A,B \in \mathcal{M}_{m\times n}(\mathbb{K})$
3. Pseudoasociativa:
   $(\alpha\beta)A = \alpha(\beta A)$; $\quad \forall \alpha,\beta \in \mathbb{K}; \forall A \in \mathcal{M}_{m\times n}(\mathbb{K})$
4. Ley de identidad: $1A = A$; $\quad \forall A \in \mathcal{M}_{m\times n}(\mathbb{K})$

5.3.3. Producto de matrices

Dadas matrices A y B en las siguientes condiciones:

$$\begin{aligned} A &= (a_{ik}) \in \mathcal{M}_{m\times p}(\mathbb{K})(m \text{ filas y } p \text{ columnas}) \\ B &= (b_{kj}) \in \mathcal{M}_{p\times n}(\mathbb{K})(p \text{ filas y } n \text{ columnas}) \end{aligned}$$

se define su producto como la matriz:

$$AB = (c_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})(m \text{ filas y } n \text{ columnas})$$

donde

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj}$$

Así, pues el producto de dos matrices A y B sólo está definido cuando el número de columnas de A es igual al número de filas de B. En tal caso, la matriz resultante tiene tantas filas como A y tantas columnas como B. El elemento c_{ij} que ocupa el lugar (i,j) en AB se obtiene a partir de la fila i-ésima de A y la columnas j-ésima de B (que han de tener igual número de elementos, en este caso p).

$$\begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{ip} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mp} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ \vdots & & \vdots & & \vdots \\ b_{p1} & \cdots & b_{pj} & \cdots & b_{pn} \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1n} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \cdots & \boxed{c_{ij}} & \cdots & c_{in} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \cdots & c_{mj} & \cdots & c_{mn} \end{pmatrix}$$

Ejemplo 5.13 *Consideremos el producto:*

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

luego el producto está definido y el resultado será una matriz de orden 2×4. *Denotando a la matriz producto por:*

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \end{pmatrix}$$

el elemento c_{11} *se obtiene de la primera fila de A y la primera columna de B:*

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

Así pues: $c_{11} = 2 \cdot 0 + 3 \cdot 3 + 1 \cdot 1 = 9$. *De igual forma* $c_{12} = 2 \cdot 1 + 3 \cdot 5 + 1 \cdot 1 = 18$. *Y calculándolos de esta forma se obtiene:*

$$\begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 2 \\ 3 & 5 & 0 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 9 & 18 & 4 & 10 \\ 3 & 7 & 4 & 7 \end{pmatrix}$$

Siempre que tengan las dimensiones adecuadas, se cumple las siguientes **propiedades del producto de matrices**:

1. $A(BC) = (AB)C$
2. Dada $A \in \mathcal{M}_{m \times n}(\mathbb{K})$, se verifica $I_m A = A$; $AI_n = A$.
3. $A(B+C) = AB + AC$
4. $(A+B)C = AC + BC$
5. $\alpha(AB) = (\alpha A)B$

Sin embargo, el producto de matrices **NO** es conmutativo. Por ejemplo,

Ejemplo 5.14

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

pero

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

Cuando dos matrices A y B cumplen $AB = BA$, diremos que A y B *conmutan*.

Por otra parte, también es posible que el producto de dos matrices sea nulo sin que lo sea ninguno de los factores:

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

Combinaciones de filas y columnas Supongamos que tenemos un producto de dos matrices: $C = AB$. Es natural preguntarse cómo interpretar en A y B las combinaciones de filas y columnas que hagamos en C. Se cumple que una combinación de filas (resp. columnas) en C se obtiene haciendo esa misma combinación en A y después multiplicando por B (resp. en B y después multiplicando por A). Lo mismo vale para las operaciones elementales, incluido el intercambio de filas o de columnas. Esta sencilla observación será muy útil, incluso en el caso aparentemente trivial en que una de las matrices sea la identidad.

En efecto, obsérvese que la fila i-ésima del producto es el resultado de multiplicar la fila i-ésima de A por las columnas de B.

Proposición 5.4 *Sean $A \in \mathcal{M}_{m\times p}$ y $B \in \mathcal{M}_{p\times n}$, se cumple que:*

$$rg(AB) \leq min\{rg(A), rg(B)\}$$

5.3.4. Inversión de Matrices

Supongamos que A es una matriz cuadrada, y consideramos la matriz ampliada $(A|I_n)$. A partir de operaciones elementales obtendremos una matriz escalonada reducida. Esta matriz será de la forma $(I_n|C)$ cuando A sea invertible. Entonces las columnas de C serán las soluciones de los sistemas anteriores, con lo que $AC = I_n$, luego $C = A^{-1}$.

Ejemplo 5.15 *Comprobemos que la matriz*

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

es invertible y calculemos su inversa, según acabamos de explicar. En primer lugar restamos la primera fila a la tercera y su cuádruplo a la segunda, para obtener

$$(A|I_3) = \left(\begin{array}{ccc|ccc} 1 & 2 & -2 & 1 & 0 & 0 \\ 4 & 2 & -2 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & -2 & 1 & 0 & 0 \\ 0 & -6 & 6 & -4 & 1 & 0 \\ 0 & -2 & 1 & -1 & 0 & 1 \end{array}\right) \rightsquigarrow$$

Restamos a la segunda fila el triple de la tercera e intercambiamos después las dos últimas filas:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & -2 & 1 & 0 & 0 \\ 0 & 0 & 3 & -1 & 1 & -3 \\ 0 & -2 & 1 & -1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & -2 & 1 & 0 & 0 \\ 0 & -2 & 1 & -1 & 0 & 1 \\ 0 & 0 & 3 & -1 & 1 & -3 \end{array}\right)$$

Sumamos a la primera fila $\frac{2}{3}$ de la tercera y restamos a la segunda $\frac{1}{3}$ de la tercera para obtener

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & \frac{1}{3} & \frac{2}{3} & -2 \\ 0 & -2 & 0 & -\frac{2}{3} & -\frac{1}{3} & 2 \\ 0 & 0 & 3 & -1 & 1 & -3 \end{array}\right)$$

Por último, sumamos la segunda fila a la primera, dividimos la segunda por -2 y la tercera por 3 para llegar a

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 1 & 0 & \frac{1}{3} & \frac{1}{6} & -1 \\ 0 & 0 & 1 & -\frac{1}{3} & \frac{1}{3} & -1 \end{array}\right)$$

En consecuencia A es invertible, y su inversa es

$$A^{-1} = \left(\begin{array}{ccc} -\frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{6} & -1 \\ -\frac{1}{3} & \frac{1}{3} & -1 \end{array}\right)$$

5.3.5. Potenciación de Matrices

Las potencias de una matriz cuadrada $A \in \mathcal{M}_n(\mathbb{K})$ se definen como:

$$A^0 = I_n \quad A^1 = A \quad A^{k+1} = A^k \cdot A \quad k \geq 1$$

Por la propiedad asociativa, se cumple que $A^p \cdot A^q = A^{p+q}$.

Sin embargo, como el producto no es conmutativo,$(AB)^p \neq A^p B^p$.

Definición 5.16 (Fórmula de Newton:) *Si A y B conmutan, entonces*

$$(A+B)^p = \sum_{k=0}^{p} \binom{p}{k} A^k B^{n-k} \tag{5.1}$$

5.4. El determinante

El determinante es un concepto que asocia a cada matriz cuadrada un escalar. Esta matriz nos permitirá decidir cuándo una matriz cuadrada es invertible, y desarrollar un método alternativo para calcular el rango de una matriz no necesariamente cuadrada. Por tanto, dicha función deberá comportarse bien en relación con las nociones centrales al respecto: combinaciones e independencia. Por supuesto tendrá también un comportamiento especial para con las operaciones elementales.

Asociaremos la matriz A a un escalar, que denotaremos det(A) y que definiremos por recurrencia sobre el orden n de la matriz A.

El caso $n = 1$ es especial, pues A es un escalar, y se define su determinante como $det(A) = A$.

Si $n \geq 1$, y fijados dos índices i, j, se denota A_{ij} a la matriz de orden $n-1$ que resulta de eliminar en A la fila i y la columna j. Así, supuesto definida la noción de determinante para matrices de orden $n-1$, se define:

$$det(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} det(A_{i1}) \tag{5.2}$$

5.4.1. Propiedades de los determinantes

1. Sea B la matriz que se obtiene a partir de la matriz A intercambiando dos de sus filas o dos de su columnas. Entonces, $det(B) = -det(A)$.
2. Si dos de las filas dos de las columnas de la matriz A coinciden, entonces el determinante de A es nulo.
3. Sean A, B y C tres matrices de orden n cuyas entradas coinciden, salvo las de la fila i-ésima, que cumplen que las de A son la suma de las de B y C. Entonces, $det(A) = det(B) + det(C)$. La misma propiedad es cierta si reemplazamos la palabra fila por columna.
4. Sean A una matriz, λ un escalar y B la matriz que se obtiene a partir de A multiplicando la fila i-ésima de esta por λ. Entonces, $det(B) = \lambda det(A)$. La misma propiedad es cierta si sustituimos la palabra fila por columna. Aplicando esta propiedad a cada una de las filas (o columnas) de λA, se deduce que si A es de orden n, entonces $det(\lambda A) = \lambda^n det(A)$.
5. Sean A una matriz, λ un escalar y B la matriz que se obtiene a partir de A sumando a su i-ésima fila el resultado de multiplicar por λ su k-ésima fila, siendo $k \neq i$. Entonces, $det(B) = det(A)$. La misma propiedad es cierta si cambiamos la palabra fila por columna.
6. Si una de las filas de la matriz A es combinación lineal de las restantes, entonces su determinante es nulo. Una vez más, también esta propiedad es cierta al

sustituir la palabra fila por columna.

7. Una matriz y su traspuesta tienen el mismo determinante $det(A) = det(A^t)$.
8. El determinante del producto de dos matrices cuadradas del mismo orden es el producto de sus determinantes:

$$det(AB) = det(A)det(B) \tag{5.3}$$

9. El valor de un determinante se puede hallar "desarrollándolo por cualquiera de sus columnas, esto es, para cada índice j,

$$det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} det(A_{ij})$$

Ejemplo 5.16 *Sea n un número entero positivo y la matriz*

$$A_n = \begin{pmatrix} n & 1 & 1\cdots & 1 & \\ n & 2 & 1\cdots & 1 & \\ n & 1 & 3\cdots & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ n & 1 & 1\cdots & n & \end{pmatrix}$$

de orden n. Para calcular su determinante, restamos a cada fila la primera. Se obtiene así la matriz

$$B_n = \begin{pmatrix} n & 1 & 1\cdots & 1 & \\ 0 & 1 & 0\cdots & 1 & \\ 0 & 0 & 2\cdots & 0 & \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0\cdots & n-1 & \end{pmatrix}$$

cuyo determinante, por ser triangular superior es $n! = det(B_n) = det(A_n)$

Es conveniente señalar algunas observaciones respecto de estas propiedades:

- En primer lugar, nótese que la definición de determinante es su desarrollo por la primera columna, y la última propiedades nos dice que podemos desarrollar por cualquiera de ellas.
- Además, combinando este hecho con que una matriz y su traspuesta tienen el mismo determinante, este se puede calcular desarrollando por filas, es decir, fijando un índice i, se tiene:

$$det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} det(A_{ij})$$

- Ya observamos que el determinante de una matriz triangular superior es el producto de los elementos de la diagonal principal. Se deduce, de la propiedad 7, que lo mismo ocurre para las matrices triangulares inferiores.
- Por último, de la octava propiedad se desprende que si A y B son matrices cuadradas del mismo orden, entonces $det(AB) = det(A)det(B) = det(B)det(A) = det(BA)$.

5.4.2. Invertibilidad de Matrices

Definición 5.17 (Matriz adjunta:) *Sea A una matriz cuadrada de orden n. Se llama matriz adjunta de A a la matriz adj(A), de orden n, cuyo coeficiente de la fila i-ésima y la columna j-ésima es* $(-1)^{i+j}det(A_{ij})$.

Ejemplo 5.17 *La adjunta de*

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

es adj(A)$= \begin{pmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix}$

La **propiedad fundamental de la matriz adjunta** a es:

$$A(adj(A^t)) = det(A)I_n = (adj(A^t))A$$

De aquí se desprende que si $det(A) \neq 0$, entonces es invertible, pues la matriz $B = \frac{1}{det(A)}(adj(A^t))$ cumple que $AB = BA = I_n$, y así $B = A^{-1}$.

Recíprocamente, si A es invertible y A^{-1} es su inversa, entonces $A \cdot A^{-1} = I_n$, y por tanto, por (5.3)

$$1 = det(I_n) = det(A \cdot A^{-1}) = det(A) \cdot det(A^{-1}) \tag{5.4}$$

luego $det(A) \neq 0$, Además de (5.4) se deduce que el determinante de la inversa de A es el inverso del determinante de A. Por tanto, se tiene que

Proposición 5.5 *Una matriz A de orden n es invertible si y sólo si* $det(A) \neq 0$. *En tal caso, su inversa es*

$$A^{-1} = \frac{1}{det(A)}(adj(A^t))$$

cuyo determinante vale

$$det(A^{-1}) = \frac{1}{det(A)}$$

Como consecuencia,

Proposición 5.6 *Si dos matrices cuadradas A y B cumplen $AB = I_n$, entonces también $BA = I_n$ y, por tanto, A y B son mutuamente inversas, esto es, $B = A^{-1}$ y $A = B^{-1}$*

5.4.3. Cáculo del determinante

Regla de Sarrus Partiendo de (5.2), si $n = 2$, se tiene la conocida fórmula

$$det(A) = a_{11}det(A_{11}) - a_{21}det(A_{21}) = a_{11}a_{22} - a_{21}a_{12}$$

.

Si $n = 3$, se tiene la conocida *fórmula de Sarrus*:

$$\begin{aligned} det(A) &= a_{11}det(A_{11}) - a_{21}det(A_{21}) + a_{31}det(A_{31}) = \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22}) = \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{21}a_{12}a_{33} + a_{21}a_{13}a_{32} + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22} \end{aligned}$$

Ejemplo 5.18 *El cálculo del determinante de una matriz es tanto más laborioso cuanto mayor sea su orden, pero en cualquier caso la definición es un procedimiento eficaz. Por ejemplo, para determinar la matriz de orden 4*

$$A = \begin{pmatrix} 2 & -1 & 5 & 11 \\ 1 & 2 & 1 & 3 \\ 1 & 0 & 1 & 1 \\ 0 & -3 & 3 & 7 \end{pmatrix}$$

se tiene que

$$det(A) = 2 \cdot det \begin{pmatrix} 2 & 1 & 3 \\ 0 & 1 & 1 \\ -3 & 3 & 7 \end{pmatrix} - det \begin{pmatrix} -1 & 5 & 11 \\ 0 & 1 & 1 \\ -3 & 3 & 7 \end{pmatrix} + det \begin{pmatrix} -1 & 5 & 11 \\ 2 & 1 & 3 \\ -3 & 3 & 7 \end{pmatrix}$$

y aplicando Sarrus:

$$\begin{aligned} det(A) &= 2\cdot[2\cdot1\cdot7-2\cdot1\cdot3-0\cdot1\cdot7+0\cdot3\cdot3+(-3)\cdot1\cdot1-(-3)\cdot3\cdot1]- \\ &\quad -[(-1)\cdot1\cdot7-(-1)\cdot1\cdot3-0\cdot5\cdot7+0\cdot3\cdot11+(-3)\cdot5\cdot1\cdot-(-3)\cdot1\cdot11]+ \\ &\quad +[(-1)\cdot1\cdot7-(-1)\cdot3\cdot3-2\cdot5\cdot7\cdot+2\cdot11\cdot3+(-3)5\cdot3-(-3)1\cdot11]= \\ &= 2\cdot[14-6-0+0-3+9]-[-7+3-0+0-15+33]+ \\ &\quad +[-7+9-70+66-45+33]= \\ &= 28-14-14= \\ &= 0 \end{aligned}$$

Regla de Laplace El determinante de una matriz $A \in \mathcal{M}_n$ puede calcularse mediante el desarrollo por adjuntos a lo largo de una columna o una fila de A.

Por columnas:

$$det(A) = a_{1j}adj_{1j}(A) + \cdots + a_{nj}adj_{nj}(A) = \sum_{k=1}^{n} a_{kj}adj_{kj}(A)$$

Por filas:

$$det(A) = a_{i1}adj_{i1}(A) + \cdots + a_{in}adj_{in}(A) = \sum_{k=1}^{n} a_{ik}adj_{ik}(A)$$

Teorema 5.2 *Sea la matriz $A \in \mathcal{M}_n$ si A es una matriz triangular, entonces $det(A)$ es el producto de los elementos de la diagonal principal.*

Universidad Rey Juan Carlos

Tema 6

Resolución de Sistemas Lineales

6.1. Ecuaciones y Sistemas Lineales

En esta sección probaremos el teorema de Roché, que es una condición necesaria y suficiente para la existencia de la solución de un sistema de ecuaciones lineales y que permite decidir, en caso de existir solución, si esta es única o no. Expondremos además, la regla de Cramer, que es un algoritmo para calcular todas las soluciones, en caso de que exista alguna.

6.1.1. Conceptos Básicos

Definición 6.1 (Ecuación Lineal:) *Una ecuación lineal (en n incógnitas o variables) es una expresión del tipo:*

$$a_1x_1 + \cdots + a_nx_n = b$$

donde los coeficientes a_i y el término independiente b son escalares y pertenecen a $\mathbb{K} = \mathbb{C}$ ó $\mathbb{R}$, y las variables x_i son las incógnitas o variables que consideramos ordenadas según sus índices.

Nótese que en una ecuación lineal no pueden aparecer términos como una incógnita al cuadrado, el producto de dos incógnitas o una función trigonométrica o logarítmica.

Ejemplo 6.1 *Las ecuaciones:*

$$2x+5y=0; \quad 3x-y+7z=13$$

son ecuaciones lineales, mientras que las siguientes no lo son:

$$2x^2+y=5; \quad xy+z=0; \quad sen(x)+y+z=14$$

Si $b=0$, la ecuación se llama **homogénea**.

Una **solución** de una ecuación es una asignación de valores a las incógnitas de forma que se verifique la igualdad.

Ejemplo 6.2 *Para la ecuación* $2x+3y=5$ *una solución es* $x=1, y=1$*; otra solución es* $x=0, y=\frac{5}{3}$

Las ecuaciones lineales se pueden clasificar en:

- **Triviales**: Si $a_1=a_2=\cdots=a_n=b=0$.
- **Incompatible**: Si $a_1=a_2=\cdots=a_n=0 \;\; b\neq 0$.
- **Compatible**: Ninguna de las dos. Éste sistema tiene solución, i.e:

$$\text{Si} \quad \exists i / 1 \leq i \leq n \quad \text{tal que} \quad a_i \neq 0 \text{ , tomando } c_j = \begin{cases} 0 \text{ si } j \neq i \\ \frac{b}{a_i} \text{ si } j = i \end{cases} \quad \text{entonces:}$$

$$c_1 \cdots c_n \text{ son soluciones}$$

Nótese que los tres conceptos anteriores son excluyentes. Además, si $a_i=0$, no se escribirá el sumando $a_i x_i$. Por tanto, la primera incógnita de una ecuación compatible es la primera variable que aparece en ella.

Definición 6.2 (Sistema Lineal:) *Un sistema lineal de m ecuaciones en n incógnitas es una lista finita de ecuaciones lineales de la forma:*

$$\begin{cases} a_{11}x_1+\cdots+a_{1j}x_j+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+\cdots+a_{2j}x_j+\cdots+a_{2n}x_n=b_2 \\ \cdots \\ a_{i1}x_1+\cdots+a_{ij}x_j+\cdots+a_{in}x_n=b_i \\ \cdots \\ a_{m1}x_1+\cdots+a_{mj}x_j+\cdots+a_{mn}x_n=b_m \end{cases}$$

Si todos los términos independientes $b_1, \cdots, b_m$ del sistema anterior son 0, decimos que el **sistema es homogéneo**.

Una *solución* del sistema anterior es una n-upla= $c = (c_1, \cdots, c_n) \in \mathbb{K}^n$ tal que:

$$\begin{cases} a_{11}c_1 + \cdots + a_{1j}c_j + \cdots + a_{1n}c_n = b_1 \\ a_{21}c_1 + \cdots + a_{2j}c_j + \cdots + a_{2n}c_n = b_2 \\ \cdots \\ a_{i1}c_1 + \cdots + a_{ij}c_j + \cdots + a_{in}c_n = b_i \\ \cdots \\ a_{m1}c_1 + \cdots + a_{mj}c_j + \cdots + a_{mn}c_n = b_m \end{cases}$$

Los sistemas lineales se pueden clasificar en:

- **Trivial**: Un sistema de ecuaciones es trivial cuando todas sus ecuaciones son triviales, o lo que es equivalente, cuando toda n-upla es solución.
- **Incompatible**: No tiene ninguna solución.
- **Sistemas compatible**: Tiene alguna solución (pero no todas).

Ejemplo 6.3 *Los sistemas*

$$\begin{cases} x+y=0 \\ x-y=0 \end{cases} \quad y \quad \begin{cases} x+2y=0 \\ x-3y=0 \end{cases}$$

son equivalentes, pues en ambos casos la única solución es x=y=0.

Ejemplo 6.4 *También son equivalentes los sistemas:*

$$\begin{cases} 2x_1+ & 6x_2+ & 4x_3+ & 2x_4 = & 14 \\ x_1+ & 4x_2+ & 3x_3+ & x_4 = & 9 \\ -2x_1- & 6x_2- & 7x_3+ & 19x_4 = & 4 \end{cases} \quad y \quad \begin{cases} x_1 & & & -6x_4 = & -5 \\ & x_2 & & +7x_4 = & 8 \\ & & x_3- & 7x_4 = & -6 \end{cases}$$

pero esto lo comprobaremos más adelante.

De este modo, diremos:

- *Discutimos o resolvemos un sistema* cuando estudiemos si es incompatible o compatible, en el último caso, encontremos todas sus soluciones.
- Dos sistemas son *equivalentes* cuando tengan las mismas soluciones.

Ejemplo 6.5

$$\begin{cases} x+y=1 \\ x+y=2 \end{cases}$$

- Las ecuaciones son *compatibles* si dado cualquier sistema siempre puedo suponer que todas sus ecuaciones son compatibles, pues las que no lo son, las elimino si lo que pretendo es discutirlo. En efecto, si hay alguna ecuación incompatible, el sistema no tiene solución, y si tiene ecuaciones triviales, suprimiéndolas obtenemos sistemas equivalentes. Nótese que un sistema puede ser incompatible y no tener ecuaciones incompatibles. Hay dos tipos de sistemas compatibles:
  - **Sistemas determinados:** Si sólo tienen una solución.
  - **Sistemas indeterminados:** Si tienen infinitas soluciones.

Definición 6.3 (Sistema Homogéneo Asociado:) *Dado un sistema:*

$$S = \begin{cases} a_{11}x_1 + \cdots + a_{1j}x_j + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2j}x_j + \cdots + a_{2n}x_n = b_2 \\ \cdots \\ a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n = b_i \\ \cdots \\ a_{m1}x_1 + \cdots + a_{mj}x_j + \cdots + a_{mn}x_n = b_m \end{cases}$$

se define como el sistema homogéneo asociado, al que resulta de poner a 0 todos los términos independientes. Es decir, al sistema:

$$S_h = \begin{cases} a_{11}x_1 + \cdots + a_{1j}x_j + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + \cdots + a_{2j}x_j + \cdots + a_{2n}x_n = 0 \\ \cdots \\ a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n = 0 \\ \cdots \\ a_{m1}x_1 + \cdots + a_{mj}x_j + \cdots + a_{mn}x_n = 0 \end{cases}$$

Nótese que:

- El sistema S_h siempre es compatible pues $\mathbf{0} = (0, \cdots, 0)$ es solución.
- S_h es determinado si su única solución es 0, o lo que es lo mismo, la solución trivial es la única que tiene.

Para estudiar la relación entre las soluciones de un sistema lineal y las de su homogéneo, es necesario aprender a *sumar y restar* uplas. Dadas dos uplas $c = (c_1, \cdots, c_n)$ y

$d = (d_1, \cdots, d_n)$, se define la suma y resta como sigue:

$$c + d = (c_1 + d_1, \cdots, c_n + d_n)$$
$$c - d = (c_1 - d_1, \cdots, c_n - d_n)$$

De este modo, si c y d son soluciones del sistema S, $c - d \in \mathbb{K}^n$ es solución de S_h. Por otro lado, si c es solución de S, cualquier otra solución d de S se obtiene sumando a c una solución del sistema homogéneo S_h.

6.1.2. Operaciones Elementales

Nuestra estrategia para resolver un sistema lineal consistirá en obtener sucesivamente sistemas equivalentes a él, cada vez más sencillos, hasta llegar a uno muy fácil de resolver. Para ello, utilizaremos la *combinación lineal de ecuaciones*

Definición 6.4 (Combinación lineal de ecuaciones:) *Se define como combinación lineal de ecuaciones a la ecuación resultante de la operación de multiplicar una ecuación f_i por un número α_i, una segunda f_j por otro α_j, y sumar los resultados para obtener $\alpha_i f_i + \alpha_j f_j$*

Con más precisión, si la ecuación $f_k = a_{k1}x_1 + \cdots a_{kn}x_n = b_k$, entonces las ecuaciones $\alpha_k f_k$ y $\alpha_i f_i + \alpha_j f_j$ son:

$$\begin{aligned} \alpha_k \cdot f_k &: (\alpha_k a_{k1})x_1 + \cdots + (\alpha_k a_{kn})x_n = \alpha_k b_k \\ \alpha_i \cdot f_i + \alpha_j \cdot f_j &: (\alpha_k a_{i1} + \alpha_j a_{j1})x_1 + \cdots + (\alpha_k a_{in} + \alpha_j a_{jn})x_n = \alpha_i b_i + \alpha_j b_j \end{aligned}$$

Esta nueva ecuación se llama combinación lineal de las dos ecuaciones f_i y f_j, y se dice que depende de ellas. Por supuesto, se pueden combinar más de dos ecuaciones, lo que representamos abreviadamente utilizando un sumatorio $\sum \alpha_i f_i$.

Obsérvese que si $c = (c_1, \cdots, c_n)$ es una solución de las ecuaciones que se combinan, también lo es de la combinación. Por tanto, al añadir combinaciones lineales no perdemos soluciones.

Por otra parte, si sabemos que una ecuación dada de nuestro sistema es combinación lineal de otras, cualquier solución de esas otras lo será automáticamente de la dada. Por tanto, esta ecuación será superflua y podremos eliminarla.

Así pues, nuestro propósito es doble. Por un lado, aprender a combinar ecuaciones para sustituir las iniciales por otras más simples. Por otro, descubrir qué ecuaciones dependen de las demás, para eliminarlas.

Ejemplo 6.6

$$\begin{cases} x_1+ & x_2+ & x_3 & =2 \\ x_1-1 & x_2+ & 2x_3 & =3 \\ 2x-1 & + & 3x_2 & =5 \end{cases} \rightsquigarrow \begin{cases} x_1+ & x_2+ & x_3 & =2 \\ x_1-1 & x_2+ & 2x_3 & =3 \end{cases} \tag{6.1}$$

son equivalentes, porque la tercera ecuación del primero es la suma de las dos primeras ecuaciones $f_3 = f_1 + f_2$.

Definición 6.5 (Operaciones elementales:)

- *Multiplicar una ecuación del sistema por un escalar no nulo:* $f_k \rightsquigarrow \lambda f_k$.
- *Sumar a una ecuación del sistema otra multiplicada por un escalar no nulo:* $f_k \rightsquigarrow f_k + \mu f_l$.
- *Intercambiar el orden de dos ecuaciones del sistema:* $f_k, f_l \rightsquigarrow f_l, f_k$.

Es claro que las dos primeras operaciones sirven para reemplazar una ecuación dada por una combinación de esa misma y otras cualesquiera. Las multiplicaciones por cero se excluyen para evitar trivialidades, y para que las operaciones sean reversibles. Para revertir, basta, respectivamente:

- Dividir (multiplicar por $1/\lambda$).
- Restar (multiplicar por -1 y sumar).

Esa reversibilidad garantiza que las soluciones siguen siendo las mismas. Con la tercera operación, sólo se cambia nuestra percepción del sistema. Por otra parte, las ecuaciones elementales conservan la independencia.

Definición 6.6 (Sistema escalonado reducido:) *Un sistema lineal se llama escalonado reducido cuando la primera incógnita de cada ecuación:*

- *Precede a la primera incógnita de la ecuación siguiente (el pivote está a la derecha).*
- *Tiene coeficiente 1 (el pivote es igual a 1).*
- *No aparece en ninguna de las demás ecuaciones (el resto de coeficientes de la columna son 0).*

Obsérvese además, que un sistema escalonado reducido no puede contener ni ecuaciones incompatibles ni triviales.

Ejemplo 6.7 *El sistema:*

$$\begin{cases} x_1 + 3x_2 + 2x_3 + x_4 = 7 \\ x_2 + x_3 = 2 \\ -3x_3 + 21x_4 = 18 \end{cases}$$

no es escalonado reducido, pues aunque la primera incógnita de cada ecuación precede a la de la siguiente, la de la segunda ecuación aparece en la primera. En este caso, las filas de 0s (las que sean combinación lineal del resto) se eliminan.

El siguiente sin embargo sí lo es:

$$\begin{cases} x_1 - 6x_4 = -5 \\ x_2 + 7x_4 = 8 \\ x_3 - 7x_4 = -6 \end{cases}$$

Proposición 6.1 *Dos sistemas escalonados reducidos compatibles equivalentes son el mismo/idénticos. En particular, si dos sistemas escalonados reducidos son equivalentes a un mismo sistema son el mismo.*

En conclusión, cada sistema compatible tiene asociado un único sistema escalonado reducido ya que dos sistemas escalonados reducidos del mismo sistema son equivalentes.

6.2. Discusión y Resolución de Sistemas Lineales

A continuación veremos métodos y teoremas para la resolución de sistemas de ecuaciones.

6.2.1. Método de escalonamiento de Gauss-Jordan

Por este método se transforma, aplicando sucesivamente operaciones elementales, cualquier sistema no trivial en un sistema incompatible o un sistema escalonado reducido. Por ejemplo:

Ejemplo 6.8 *Vamos a explicar el ejemplo (6.4). Aligeramos la explicación de cómo se*

aplica el algoritmo. Dividimos la primera ecuación por 2 ($f_1' = f_1 \cdot \frac{1}{2}$)*:*

$$\begin{cases} x_1 & + & 3x_2 & + & 2x_3 & + & x_4 & = & 7 \\ x_1 & + & 4x_2 & + & 3x_3 & + & x_4 & = & 9 \\ -2x_1 & - & 6x_2 & - & 7x_3 & + & 19x_4 & = & 4 \end{cases}$$

A continuación restamos la primera de la segunda, y sumamos el doble de la primera a la tercera ($f_2'' = f_2' - f_1'; f_3'' = f_3' + 2f_1'$)*:*

$$\begin{cases} x_1 & + & 3x_2 & + & 2x_3 & + & x_4 & = & 7 \\ & & x_2 & + & x_3 & & & = & 2 \\ & & & & -3x_3 & + & 21x_4 & = & 18 \end{cases}$$

Ahora restamos a la primera el triple de la segunda ecuación ($f_2''' = f_2'' - 3f_1''$)*:*

$$\begin{cases} x_1 & & - & x_3 & + & x_4 & = & 1 \\ & x_2 & + & x_3 & & & = & 2 \\ & & & -3x_3 & + & 21x_4 & = & 18 \end{cases}$$

A continuación dividimos la tercera ecuación por -3 ($f_3^{iv} = f_3''' \cdot \frac{1}{3}$)*:*

$$\begin{cases} x_1 & & - & x_3 & + & x_4 & = & 1 \\ & x_2 & + & x_3 & & & = & 2 \\ & & & x_3 & - & 7x_4 & = & -6 \end{cases}$$

Para terminar sumamos la tercera ecuación a la primera y la restamos de la segunda ($f_2^v = f_2^{iv} - f_3^{iv}; f_1^v = f_1^{iv} + f_3^{iv}$)*:*

$$\begin{cases} x_1 & & & & - & 6x_4 & = & -5 \\ & x_2 & & & + & 7x_4 & = & 2 \\ & & x_3 & & - & 7x_4 & = & -6 \end{cases}$$

Este sistema, equivalente al de partida, es ya escalonado reducido.

Formalmente, describimos este algoritmo en varios pasos:

1. Se hace $k = 1$.
2. Se examinan las ecuaciones a partir de la $k-$ésima inclusive: si hay alguna incompatible, el sistema es incompatible y hemos terminado; si hay triviales, se suprimen; si no hay otras ecuaciones, hemos terminado.

3. Si hay otras ecuaciones, se elige una ecuación entre ellas cuya primera incógnita x_{j_k} tengo índice j_k mínimo.
4. Se divide la ecuación elegida por el coeficiente de su primera incógnita x_{j_k}, y se intercambia la ecuación resultante con la k-ésima.
5. A cada ecuación distinta de la nueva k-ésima le restamos ésta multiplicada por el coeficiente que en aquélla tiene la incógnita x_{j_k}.
6. Se hace $k = k+1$ y se vuelve al paso 1.

Vamos a resumir qué se hace y para qué en los pasos anteriores. Nótese que primero sólo se hacen operaciones elementales, con lo que se obtiene siempre un sistema equivalente al de partida. En el paso 1, se decide si hay que hacer realmente algo. Si hay en efecto que hacerlo, en el paso 2 se selecciona la incógnita y la ecuación que se usarán para mejorar el posible escalonamiento que ya se tenga. En el paso 3 se hace 1 el coeficiente de esa incógnita en esa ecuación, y se coloca la ecuación en el lugar pertinente. En el paso 4 se elimina la susodicha incógnita de las demás ecuaciones; obsérvese que aquí pueden producirse ecuaciones triviales (a partir de la (k+1)-ésima). Y en el paso 5 se recomienza con otra incógnita posterior.

Es claro que haciendo esto un número finito de veces el sistema resulta incompatible o escalonado reducido.

Las incógnitas de un sistema escalonado reducido que son la primera variable de alguna ecuación se llaman *principales*. El resto se llaman *secundarias* o *parámetros*. Nótese que si una de las n incógnitas no aparece en el sistema, entonces es necesariamente secundaria.

Es importante observar que las ecuaciones de un sistema escalonado reducido son independientes, es decir, ninguna depende de las otras. En efecto, la primera incógnita de una ecuación dada no aparece en ninguna de las otras, luego no se puede obtener la ecuación dada combinando esas otras.

Para discutir y resolver un sistema basta discutir y resolver cualquier otro equivalente, y acabamos de explicar cómo obtener uno equivalente escalonado reducido. Veamos pues cómo se discuten y resuelven dichos sistemas.

6.2.2. Discusión y Resolución de Sistemas

Veamos cómo discutir y resolver el sistema del ejemplo anterior.

Ejemplo 6.9 *Comenzamos sustituyéndolo por el escalonado reducido equivalente:*

$$\begin{cases} x_1 & & & - & 6x_4 & = & -5 \\ & x_2 & & + & 7x_4 & = & 2 \\ & & x_3 & - & 7x_4 & = & -6 \end{cases}$$

En este sistema las incógnitas x_1, x_2, x_3 son principales, y x_4 es secundaria. Para resolverlo, hacemos $x_4 = \lambda$, y las ecuaciones paramétricas son:

$$\begin{cases} x_1 &= -5+ & 6\lambda \\ x_2 &= 8- & 7\lambda \\ x_3 &= -6+ & 7\lambda \\ x_4 &= & \lambda \end{cases}$$

En consecuencia, las soluciones son las uplas:

$$(-5+6\lambda, 8-7\lambda, -6+7\lambda, \lambda)$$

Si hacemos $\lambda = 0, 1$, obtenemos las dos soluciones (-5,8,-6,0), (1,1,1,1) y así vemos que el sistema es compatible indeterminado.

Formalmente, si las incógnitas principales de un sistema escalonado reducido son $x_1, \cdots, x_r$, y las incógnitas secundarias son $x_{r+1}, \cdots, x_n$, su expresión es del tipo:

$$\begin{cases} x_1 & & & + & \alpha_{1,r+1}x_{r+1} & + & \cdots & + & \alpha_{1n}x_n & = & \beta_1 \\ & x_2 & & + & \alpha_{2,r+1}x_{r+1} & + & \cdots & + & \alpha_{2n}x_n & = & \beta_2 \\ & & \ddots & & \dots & + & \cdots & + & \dots & = & \dots \\ & & x_r & + & \alpha_{r,r+1}x_{r+1} & + & \cdots & + & \alpha_{rn}x_n & = & \beta_r \end{cases}$$

con $\alpha_{ij}, \beta_k \in \mathbb{K}$. Para resolver este sistema basta despejar las incógnitas principales en aquellas ecuaciones del sistema que las tienen. De este modo, expresamos cada incógnita principal en función del término independiente y las incógnitas secundarias. Así, nos queda lo siguiente:

$$\begin{cases} x_1 &= -\alpha_{1,r+1}x_{r+1} - \cdots - \alpha_{1n}x_n + \beta_1 \\ & \dots \\ x_r &= -\alpha_{r,r+1}x_{r+1} - \cdots - \alpha_{rn}x_n + \beta_r \end{cases}$$

O bien, si escribimos $x_j = \lambda_j$ para $j = r+1, \cdots, n$, obtenemos las denominadas ecuaciones paramétricas:

$$\begin{cases} x_1 &= -\alpha_{1,r+1}\lambda_{r+1} & - \cdots - & \alpha_{1n}\lambda_n + \beta_1 \\ & \dots \\ x_r &= -\alpha_{r,r+1}\lambda_{r+1} & - \cdots - & \alpha_{rn}\lambda_n + \beta_r \\ x_{r+1} &= \lambda_{r+1} & & \\ & \dots & & \\ x_n &= & & \lambda_n \end{cases}$$

Ejemplo 6.10 *Discutimos a continuación el sistema:*

$$\begin{cases} x_1+x_2+x_3+x_4=7 \\ 2x_1-x_2-3x_3-x_4=3 \\ 8x_1-x_2-7x_3-x_4=10 \end{cases}$$

Restamos a las ecuaciones segunda y tercera el producto de la primera por 2 y 8, respectivamente, y multiplicamos por -1 las ecuaciones resultantes. Así el sistema dado es equivalente al siguiente:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 7 \\ 3x_2 + 5x_3 + 3x_4 = 11 \\ 9x_2 + 15x_3 + 9x_4 = 46 \end{cases}$$

Ahora restamos el triple de la segunda ecuación a la tercera, para obtener,

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 7 \\ 3x_2 + 5x_3 + 3x_4 = 11 \\ 0 = 13 \end{cases}$$

y la última ecuación nos dice que el sistema es incompatible.

6.2.3. Teorema de Rouché Fröbenious

Introducimos en esta sección los conceptos adecuados para sistematizar la discusión y resolución de los sistemas lineales. Observemos en primer lugar que el sistema:

$$(*) = \begin{cases} a_{11}x_1+\cdots+a_{1j}x_j+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+\cdots+a_{2j}x_j+\cdots+a_{2n}x_n=b_2 \\ \cdots \\ a_{i1}x_1+\cdots+a_{ij}x_j+\cdots+a_{in}x_n=b_i \\ \cdots \\ a_{m1}x_1+\cdots+a_{mj}x_j+\cdots+a_{mn}x_n=b_m \end{cases}$$

Se puede representar simbólicamente utilizando tablas de coeficientes y de variables:

$$(\bullet) = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x1 \\ \cdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b1 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{pmatrix}$$

Para hacer más fácil la escritura y la representación simbólica antes mencionada, es preferible utilizar *filas*, como $(x_1,\cdots,x_n)$ y $(b_1,\cdots,b_m)$, en lugar de las respectivas columnas que aparecen en $(\bullet)$. Pero entonces hacen falta notaciones que relacionen las filas y columnas anteriores. Así, si $x=(x_1,\cdots,x_n)$ es la fila de incógnitas y $b=(b_1,\cdots b_m)$ es la fila de términos independientes, denotaremos por x^t y b^t las columnas correspondientes. De esta manera $(\bullet)$ se abrevia:

$$A\cdot x^t=b^t$$

donde por supuesto, A representa la tabla de coeficientes del sistema. Esta forma de denotar los sistemas lineales mediante tablas de números e incógnitas se puede entender rigurosamente como un producto, definiendo éste del modo evidente que el sistema nos propone: se multiplican ordenadamente los elementos de cada fila de la tabla A por las incógnitas de la columna x^t, y se suman estos productos. Así resulta una columna con los primeros miembros de las ecuaciones del sistema. Ahora la igualdad de esta columna y la de los términos independientes b^t es el sistema dado.

Teorema 6.1 (Teorema de Rouché-Fröbenius:) *Un sistema lineal $Ax^t=b^t$ en n incógnitas es compatible si y sólo si $rg(A)=rg(\widetilde{A})$. En tal caso, el sistema es determinado si y sólo si $rg(A)=n$.*

Resumimos la información obtenida hasta el momento acerca del sistema de ecuaciones lineales como sigue:

1. Si $rg(A)=rg(\widetilde{A})=n$, el sistema es compatible y determinado.
2. Si $rg(A)=rg(\widetilde{A})<n$, el sistema es compatible e indeterminado.
3. Si $rg(A)\neq rg(\widetilde{A})$, el sistema es incompatible.

Ejemplo 6.11 *Vamos a discutir, en función de los valores de los números reales α y β, el siguiente sistema de ecuaciones lineales:*

$$\begin{cases} x & +y & +\alpha z & =1 \\ 2x & +\alpha y & & =\beta \\ \alpha x & +y & +z & =\beta^2 \end{cases}$$

La matriz que representa el sistema será

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

Por tanto, su determinante será

$$det(A) = \alpha(2-\alpha^2) + (\alpha - 2) = 3\alpha - \alpha^2 - 2 = -(1-\alpha)^2(\alpha+2)$$

Por ello,

- *Si* $\alpha \neq -2$ *y* $\alpha \neq 1$*, entonces* $rg(A) = rg(\widetilde{A}) = 3$*, por lo que el sistema es compatible y determinado.*
- *Para* $\alpha = -2$*,* $rg(\widetilde{A}) = 3 \neq rg(A) = 2$*, por lo que el sistema es incompatible.*
- *Si* $\alpha = 1$*,* $rg(A) = 2$*, mientras que el rango de* $\widetilde{A}$ *es 2 si* $\beta^2 = 1$ *ó 3 si* $\beta^2 \neq 1$*. Por tanto, el sistema es compatible e indeterminado si* $\alpha = 1$ *y* $\beta \pm 1$*, e incompatible si* $\alpha = 1$ *y* $\beta \not\pm 1$

Especial importancia tienen los sistemas homogéneos, que son aquellos cuya matriz de términos independientes es nula. Es obvio que en tal caso los rangos de A y $\widetilde{A}$ coinciden, luego dichos sistemas son compatibles. De hecho $x_1 = 0, \cdots x_n = 0$ es una solución, denominada trivial. De lo anterior se desprende que:

- Un sistema homogéneo con n incógnitas tiene alguna solución distinta de la trivial si y sólo si $rg(A) \neq n$.
- En particular, todo sistema homogéneo con menos ecuaciones que incógnitas admite alguna solución distinta de la trivial.

6.2.4. Regla de Cramer

La regla de Cramer es uno de los métodos de resolución de sistemas compatibles más populares.

Sea la matriz $A \in \mathcal{M}_n$ y sea $b \in \mathbb{R}^n$, se define $A_i(b)$ como la matriz obtenida al sustituir en A la columna i-ésima por el vector b

$$A_i(b) = \begin{pmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{pmatrix}$$

entonces, sea $A \in \mathcal{M}_n$ una matriz invertible, para cualquier $b \in \mathbb{R}^n$ del sistema $Ax = b$ la única solución viene dada por

$$x_i = \frac{det(A_i(b))}{det(A)} \quad \forall i = 1 \cdots, n$$

Por tanto, para conseguir todas las soluciones de un sistema compatible hay que obtener un sistema equivalente al dado y aplicar la regla de Cramer al sistema resultante de tomar las incógnitas $x_{r+1}, \cdots, x_n$ como parámetros.

Ejemplo 6.12 *Para cada terna de números reales α, β y γ se considera el sistema de ecuaciones:*

$$\begin{cases} -x & & +z & -t & = 4 \\ & y & -z & +\alpha t & = 1 \\ x & -y & & +t & = \beta \\ \alpha x & +y & -z & & = \gamma \end{cases}$$

Estudiaremos para qué valores de α, β y γ el sistema es compatible y, en esos casos, obtendremos todas sus soluciones.

El determinante de la matriz de coeficientes de A vale α^2.

CASO 1: Si α no es nulo, $rg(A) = rg(\widetilde{A}) = 4$ y el sistema es compatible y determinado. Podemos aplicar directamente la regla de Crammer y tenemos así:

$$x = \tfrac{1}{\alpha^2} det \begin{pmatrix} 4 & 0 & 1 & -1 \\ 1 & 1 & -1 & \alpha \\ \beta & -1 & 0 & 1 \\ \gamma & 1 & -1 & 0 \end{pmatrix} \quad y = \tfrac{1}{\alpha^2} det \begin{pmatrix} -1 & 4 & 1 & -1 \\ 0 & 1 & -1 & \alpha \\ 1 & \beta & 0 & 1 \\ \alpha & \gamma & -1 & 0 \end{pmatrix}$$

$$z = \tfrac{1}{\alpha^2} det \begin{pmatrix} -1 & 0 & 4 & -1 \\ 0 & 1 & 1 & \alpha \\ 1 & -1 & \beta & 1 \\ \alpha & 1 & \gamma & 0 \end{pmatrix} \quad t = \tfrac{1}{\alpha^2} det \begin{pmatrix} -1 & 0 & 1 & 4 \\ 0 & 1 & -1 & 1 \\ 1 & -1 & 0 & \beta \\ \alpha & 1 & -1 & \gamma \end{pmatrix}$$

Calculando los determinantes anteriores resulta que si $\alpha \neq 0$, la única solución es

$$x = \frac{4+\beta+\gamma}{\alpha} \quad y = \frac{9+2\beta+\gamma-\alpha\beta}{\alpha} \quad z = \frac{9+4\alpha+2\beta+\gamma}{\alpha} \quad t = \frac{5+\beta}{\alpha}$$

CASO 2: Suponemos en lo que sigue que $\alpha = 0$, con lo que el sistema se convierte en

$$\begin{cases} -x & & +z & -t & = 4 \\ & y & -z & & = 1 \\ x & -y & & +t & = \beta \\ & +y & -z & & = \gamma \end{cases}$$

Sumando las ecuaciones primera y terera resulta que $z - y = 4 + \beta$; luego, a la vista de las ecuaciones segunda y cuarta, para que sistema sea compatible es necesario que $\gamma = 1 = -(4+\beta)$. Por tanto,

- *Si $\alpha = 0$, y $\gamma = 1$ p $\beta \neq -5$, el sistema es incompatible.*
- *Solo resta analizar qué sicede si $\alpha = 0, \gamma = 1, \beta = -5$:*

  Las ecuaciones segunda y cuarta coinciden, mientras que la tercera es el resultado de cambiar de signo la suma de la primera y la segunda. En consecuencia el sitema es equivalente a

$$\begin{cases} -x & & +z & -t & = 4 \\ & y & -z & & = 1 \end{cases}$$

  En esta situación lo razonable es escribir

$$x = z - t - 4; \quad y = z + 1$$

  lo que expresa todas las soluciones en función de los parámetros z y t.

Universidad Rey Juan Carlos

Tema 7

Espacios Vectoriales

7.1. Espacios y Subespacios vectoriales

Ya comprobamos que los elementos de $\mathbb{R}^n$ se pueden sumar y multiplicar por escalares, siguiendo unas regla que parecen naturales. A continuación, veremos cómo en muchos otros conjuntos que no son familiares, tales como los polinomios, las matrices de tipo fijado o las funciones reales de variable real, también se puede sumar y multiplicar por escalares, lo que da lugar al concepto de espacio vectorial.

7.1.1. Definiciones y ejemplos

Definición 7.1 (Espacio Vectorial:) *Un espacio vectorial sobre $\mathbb{K}$, o equivalentemente un $\mathbb{K}$-espacio vectorial, es un conjunto (no vacío) V dotado de dos operaciones,*

una denominada suma $+ : V \times V \to V$*, y otra denominada producto por escalares* $\cdot : \mathbb{K} \times V \to V$ *que cumplen,*

1. *Para cada* $x, y \in V$*, su suma* $x+y \in V$*. (Se dice por ello que la suma de vectores es una operación interna en V).*
2. *Para cada* $x, y, z \in V$*, se tiene* $x+(y+z) = (x+y)+z$*. (Asociativa).*
3. *Para cada* $x, y \in V$*, se tiene* $x+y = y+x$*. (Conmutativa).*
4. *Existe un elemento en V, llamado vector nulo, y que se denota* 0_V*, tal que para cada* $x \in V$ *se tiene* $x + O_V = x$*.*
5. *Para cada* $x \in V$ *existe un vector en E, que se denota -x, y que denomina opuesto de x, tal que* $x + (-x) = 0_V$*.*
6. *Para cada* $a \in \mathbb{R}$ *y cada* $x \in V$ *el producto de a por x, que se denoata ax ó* $a \cdot x$*, es un elemento de V, esto es* $ax \in V$*.*
7. *Para cada* $a, b \in \mathbb{R}$ *y cada* $x \in V$ *se tiene que* $(a+b)x = ax + bx$*. (Distributiva).*
8. *Para cada* $a \in \mathbb{R}$ *y cada* $x, y \in V$ *se tiene que* $a(x+y) = ax + ay$*. (Distributiva).*
9. *Para cada* $a, b \in \mathbb{R}$ *y cada* $x \in V$ *se tiene que* $(ab)x = a(bx)$*. (Asociativa).*
10. *Para cada* $x \in V$ *se tiene* $1x = x$*.*

En el caso en el que $\mathbb{K} = \mathbb{R}$, diremos que V es un espacio vectorial real, mientras que si $\mathbb{K} = \mathbb{C}$ diremos que V es un espacio vectorial complejo.

Decimos que la suma de vectores es una operación interna porque el resultado de sumar dos elementos de V es otro elemento de V, mientras que el producto de escalares por vectores es una operación externa.

Por cumplirse las cinco primeras propiedades, que solo involucran a la suma, se dice que el conjunto V, dotado de la operación suma, o más concisamente el par $(V, +)$, es un *grupo abeliano*.

Y por cumplirse las diez propiedades anteriores, se dice que la terna $(V, +, \cdot)$ es un espacio vectorial sobre $\mathbb{R}$. Abreviaremos diciendo que V es un espacio vectorial, asumiendo que los escalares son números reales, y que no existe confusión en cuanto a la suma y el producto que se están empleando.

Ejemplo 7.1 $\mathbb{R}$ *es un espacio vectorial con la suma y el producto de números reales.*

Ejemplo 7.2 $\mathbb{R}^n$ *es un espacio vectorial con la suma y el producto definidos como:*

$$\begin{aligned} (x_1, x_2, \cdots x_n) + (y_1, y_2, \cdots, y_n) &= (x_1 + y_1, x_2 + y_2, \cdots x_n + y_n) \\ k(x_1, x_2, \cdots x_n) &= (kx_1, kx_2, \cdots kx_n) \end{aligned}$$

Ejemplo 7.3 *El conjunto $M_{m,n}(\mathbb{R})$ formado por las matrices con m filas y n columnas cuyas entradas son números reales es un espacio vectorial. En particular, lo es el conjunto $M_n(\mathbb{R})$*

Ejemplo 7.4 *Existe un espacio vectorial con un único vector, que ha de ser neutro para la suma y por tanto lo llamamos 0. La suma y el producto por escalares vienen definidos por;*

$$\begin{aligned} 0+0 &= 0 \\ k0 &= 0, \quad \forall k \in K \end{aligned}$$

Este espacio vectorial recibe el nombre de espacio vectorial cero o espacio trivial y es denotado habitualmente por $\{0\}$ o simplemente 0.

Proposición 7.1 *Sea V un espacio vectorial sobre $\mathbb{K}$. Sean $u,v \in V$ y $\lambda,\mu \in \mathbb{K}$. Se cumple:*

1. *El elemento neutro 0 es único.*
2. *$\lambda u = 0$ si y sólo si $\lambda = 0$ o $u = 0$.*
3. *$-u = (-1)\cdot u$, por lo que el opuesto es único.*
4. *Si $\lambda u = \mu u$ y $u \neq 0$, entonces $\lambda = \mu$*
5. *Si $\lambda u = \lambda v$ y $\lambda \neq 0$, entonces $u = v$.*

Una suma y un producto por escalares son las operaciones mínimas necesarias para tener una noción de combinación, como la que hemos utilizado repetidamente para ecuaciones lineales y para filas de matrices. En un espacio vectorial se llama **combinación lineal** a cualquier expresión de la forma

$$\lambda_1 u_1 + \cdots + \lambda_r u_r \quad \text{con} \quad u_i \in E, \lambda_i \in \mathbb{K}$$

7.2. Dependencia e Independencia lineal

En esta sección identificamos y estudiamos los subconjuntos que generan un espacio vectorial V o un subpacio de H de la manera más eficiente posible.

Definición 7.2 *Se dice que un conjunto de vectores $\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$ de V es **lineamente independiente** si la ecuación vectorial*

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \cdots + c_n\mathbf{v}_n = 0 \tag{7.1}$$

tiene solución trivial, $c_1 = 0,\cdots,c_n = 0$.

Definición 7.3 *Se dice que un conjunto* $\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$ *de V es* ***lineamente dependiente*** *si (7.1) tiene una solución no trivial, esto es, si existen pesos* $c_1,\cdots c_n$ *no todos cero, tales que se cumple (7.1). En tal caso, se dice que (7.1) es una* ***relación de dependencia lineal*** *entre* $\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$.

Proposición 7.2 *Sea V un espacio vectorial sobre* $\mathbb{K}$*,entonces:*

- *Si* $0\in\{v_1,\cdots,v_n\}$*, entonces* $\{v_1,\cdots,v_n\}$ *es linealmente dependediente.*
- $\{v_1\}$ *es linealmente independiente si , y sólo si* $v_1\neq 0$.
- *Si* $\{v_1,\cdots,v_n\}$ *es linealmente dependiente, entonces* $\{v_1,\cdots,v_n,v_{n+1},\cdots v_{n+r}\}$ *es linealmente dependiente.*
- *Si* $\{v_1,\cdots,v_n,v_{n+1},\cdots v_{n+r}\}$ *es linealmente independiente, entonces* $\{v_1,\cdots,v_n\}$ *es linealmente independiente.*
- *Un conjunto de vectores* $\{v_1,\cdots,v_n\}$ *es linealmente dependiente si, y sólo si, uno de los vectores es combinación lineal de los restantes.*

Ejemplo 7.5 *Estudiemos si el siguiente conjunto de vectores de* $\mathbb{R}^3$ *es linealmente dependiente o independiente:*

$$\{(1,0,1),(1,1,0),(1,1,1),(1,2,1)\}$$

Para ello planteamos el sistema

$$(0,0,0)=a(1,0,1)+b(1,1,0)+c(1,1,1)+d(1,2,1)$$

que tienes por coeficientes la matrix

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

en las que las columnas son los vectores del conjunto. Puesto que el sistema es homogéneo, la solución trivial será la única solución si el rango es igual al número de incógnitas (que son el número de vectores). Como el rango es 3 y hay 4 vectores, existen más soluciones además de la trivial, y los vectores son linealmente dependientes.

7.3. Sistemas Generadores de un Espacio Vectorial

Definición 7.4 *Un conjunto de vectores S será un* ***sistema generador*** *del espacio vectorial V si todo vector de V es combinación lineal de los vectores de S. Así, definiremos el conjunto Gen*$\{v_1,\cdots,v_n\}$ *al conjunto de todos los vectores que se pueden escribir como combinaciones lineales de,* $v_1,\cdots,v_n$.

Ejemplo 7.6 *$\{(1,1),(1,0),(1,-1)\}$ es un sistema de generadores para $\mathbb{R}^2$. Para comprobarlo plateamos el siguiente problema: dado un vector $(x,y) \in \mathbb{R}^2$ encontrar escalares a,b,c tales que*

$$a(1,1)+b(1,0)+c(1,-1)=(x,y)$$

Esto se convierte en estudiar si el sistema

$$\begin{cases} a & +b & +c & = x \\ a & & -c & = y \end{cases}$$

donde las incógnitas a,b,c, tienen solución para cualquier columna de términos independientes. Efectivamente, el rango de la matriz de coeficientes es 2 y el rango de la matriz ampliada no puede ser mayor de 2, luego el sistema es compatible y los vectores forman un sistema de generadores.

Lema 7.1 *Si $\{u_1,u_2,\cdots,u_n\}$ es un sistema de generadores del espacio vectorial V y u_i es combinación de los restantes vectores, entonces el conjunto de vectores que se obtiene eliminado u_i, $\{u_1,\cdots,u_{i-1},u_{i+1},\cdots,u_n\}$ es también un sistema de generadores de V.*

Ejemplo 7.7 *En el ejemplo (7.6) vimos que el conjunto $\{(1,1),(1,0),(1,-1)\}$ es un sistema de generadores para $\mathbb{R}^2$. Como el primer vector es combinación lineal de los otros, podemos quitarlo y obtener el también sistema de generadores $\{(1,0),(1,-1)\}$*

7.4. Bases de un espacio vectorial

Definición 7.5 *Dado un espacio vectorial V, un subconjunto $B \subseteq V$ es una **base** de V si*

- *B es linealmente independiente.*
- *B es sistema de generadores de V.*

Teorema 7.1 (Teorema de la Base:) *Si un espacio vectorial V tiene una base formada por un número finito de vectores, entonces todas las bases de V son finitas y tienen igual número de vectores.*

Definición 7.6 *De un espacio vectorial V que posee una base finita diremos que es un **espacio vectorial de dimensión finita** y llamaremos **dimensión** de V al número de vectores en cualquiera de sus bases, y lo denotaremos $dim_{\mathbb{K}}(V)$.*

Nótese que $dim_{\mathbb{K}}(V)$ es el mayor número posible de vectores independientes en V, y también el menor número posible de vectores en un sistema generador de V.

Ejemplo 7.8 *En el espacio vectorial $\mathbb{K}^n$ el conjunto formad por los n vectores*

$$(1,0,0,\cdots,0),(0,1,0,\cdots,0),\cdots(0,0,0,\cdots,1)$$

es una base que recibe el nombre de ***base canónica****.*

Teorema 7.2 *En un espacio vectorial no nulo, de cada sistema de generadores finito se puede extraer una base.*

Por tanto, para obtener una base basta con eliminar uno a uno los vectores de un sistema generador que sean combinación lineal del resto. Un método alternativo viene dado por el siguiente teorema:

Teorema 7.3 *Sea V un espacio vectorial sobre $\mathbb{K}$ de dimensión n y sea $\{v_1,\cdots,v_s\}$ un conjunto de vectores linealmente independientes. Entonces existen vectores $\{v_{s+1},\cdots,v_n\}$ tales que $\{v_1,\cdots,v_s,v_{s+1},\cdots v_n\}$ es una base.*

Ahora tenemos el segundo método para calcular una base: a partir de un conjunto de vectores linealmente independientes, añardir nuevos vectores de manera que se siga manteniendo la independencia.

Proposición 7.3 *Sea V un espacio vectorial de dimensión n, entonces dado un conjunto de exactamente n vectores $S=\{v_1,\cdots v_n\}$ son equivalentes:*

- *S es linealmente independiente.*
- *S es sistema de generadores de V.*
- *S es una base de V.*

7.4.1. Coordenadas de un vector respecto de una base

A continuación, desarrollamos la herramienta que nos permitirá trabajar en cualquier espacio vectoiral de dimensión finita, digamos n, como el correspondiente espacio $\mathbb{K}^n$.

Proposición 7.4 *Sea V un espacio vectoria sobre $\mathbb{K}$. Si $B=\{u_1,\cdots u_n\}$ es una base de V, entonces todo vecto $x\in V$ se escribe de* ***forma única*** *como combinación lineal de los vectores de B.*

Después de este resultado, tenemos que dado un vector y una base, el vector puede representarse por los escalares que aparecen en la anterior combinación lineal . Sea $B = \{u_1, \cdots u_n\}$ una base de V, si $x = x_1u_1 + x_2u_2 + \cdots + x_nu_n$ es la expresión única del vector $x \in V$ como combinación lineal de los vectores de la base B, diremos que $(x_1, x_2, \cdots, x_n)$ son las **coordenadas** de x respecto de la base B, y lo representaremos por

$$x = (x_1, x_2, \cdots, x_n)_B$$

Ejemplo 7.9 *Consideremos en* $\mathbb{R}^3$ *la base canónica* $B = \{(1,0,0),(0,1,0),(0,0,1)\}$. *El vector (2,3,1) tiene coordenadas respecto de la base canónica* $x = (2,3,1)_B$, *mientras que para la base* $B' = \{(1,1,0),(0,1,1),(0,0,1)\}$, *sus coordenadas son* $x = (2,1,-2)_{B'}$.

Definición 7.7 (Coordenadas y operaciones con vectores:) *Sea V un espacio vectorial de dimensión n, B una base de V, e x e y vectores de V de coordenadas* $x = (x_1, \cdots, x_n)_B$ *e* $y = (y_1, \cdots, y_n)_B$ *, entonces*

- $x + y = (x_1 + y_1, \cdots, x_n + y_n)_B$
- $kx = (kx_1, \cdots, kx_n)_B$

7.4.2. Coordenadas y dependencia lineal

Las coordenadas respecto de una base son también muy útiles cuando se trata de determinar la dependencia o independencia lineal de un conjunto de vectores.

Proposición 7.5 *Sea V un espacio vectorial y sea B una base de V. Un conjunto de r vectores* $\{u_1, \cdots u_r\}$ *en V es linealmente independiente si, y sólo si, la matriz cuyas columnas (respectivamente filas) son sus coordenadas respecto de B tiene rango r.*

Ejemplo 7.10 *Consideremos en* $\mathbb{R}^4$ *los vectores* $u_1 = (1,1,2,2), u_2 = (0,1,1,1)$ *y* $u_3 = (2,0,2,2)$. *Tomando la base canónica, la matriz cuyas filas son las coordenadas de estos vectores es*

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

que tiene forma de Hermite por filas

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

y que por tanto tiene rango 2. Así pues, los vectores u_1, u_2, u_3 son linealmente dependientes.

7.4.3. Cambio de base

Supongamos ahora que un espacio vectorial de dimensión n tenemos dos bases diferentes:

$$B = \{e_1, e_2, \cdots, e_n\}$$
$$B' = \{e'_1, e'_2, \cdots, e'_n\}$$

y queremos establecer la relación entre coordenadas de un **mismo** vector en las dos bases. Llamemos $(x_1, x_2, \cdots, x_n)$ a las coordenadas de $x \in V$ respecto e la base B y $(x'_1, x'_2, \cdots, x'_n)$ a las coordenadas de $x \in V$ respecto e la base B'. La relación que exista entre estas coordenadas dependerá de la que exista entre ambas bases, por tanto, debemos tener ciertos datos que las relacionen. Por ejemplo, supongamos conocidas las coordenadas de los vectores de B' en la base B:

$$\begin{aligned} e'_1 &= a_{11}e_1 + a_{21}e_2 + \cdots a_{n1}e_n \\ e'_2 &= a_{12}e_1 + a_{22}e_2 + \cdots a_{n2}e_n \\ &\cdots \\ e'_n &= a_{1n}e_1 + a_{2n}e_2 + \cdots a_{nn}e_n \end{aligned}$$

Ahora tenemos dos expresiones para el vector x:

$$x_1e_1 + x_2e_2 + \cdots x_ne_n = x = x'_1e'_1 + x'_2e'_2 + \cdots x'_ne'_n$$

y sustituyendo las ecuaciones del sistema:

$$\begin{aligned} x_1e_1 + x_2e_2 + \cdots x_ne_n &= x'_1(a_{11}e_1 + a_{21}e_2 + \cdots a_{n1}e_n) + \\ &+ x'_2(a_{12}e_1 + a_{22}e_2 + \cdots a_{n2}e_n) + \\ &\cdots \\ &+ x'_n(a_{1n}e_1 + a_{2n}e_2 + \cdots a_{nn}e_n) \end{aligned}$$

reordenando

$$\begin{aligned} x_1e_1 + x_2e_2 + \cdots x_ne_n &= (a_{11}x'_1 + a_{12}x'_2 + \cdots a_{1n}x'_n)e_1 + \\ &+ (a_{21}x'_1 + a_{22}x'_2 + \cdots a_{2n}x'_n)e_2 + \\ &\cdots \\ &+ (a_{n1}x'_1 + a_{n2}x'_2 + \cdots a_{nn}x'_n)e_n \end{aligned}$$

Como son dos expresiones de x en la misma base, los escalares en ambas combinaciones lineales deben ser iguales, con lo que quedan las **ecuaciones de cambio de base** de B' a B:

$$\begin{cases} x_1 = a_{11}x'_1 + a_{12}x'_2 + \cdots a_{1n}x'_n \\ x_2 = a_{21}x'_1 + a_{22}x'_2 + \cdots a_{2n}x'_n \\ \cdots \\ x_n = a_{n1}x'_1 + a_{n2}x'_2 + \cdots a_{nn}x'_n) \end{cases}$$

o bien en forma matricial:

$$X = PX'$$

donde X y X' son las matrices columna formadas por las coordenadas del vector x en las bases B y B' respectivamente, y P es la matriz que tiene por columnas las coordenadas de los vectores de B' en función de B.

La matriz P se llama la matriz de cambio de base de B' a B y es una matriz regular. Es decir P tiene inversa. De este modo también se pueden obtener las coordenadas de B' a partir de B:

$$X' = P^{-1}X$$

Proposición 7.6 *Toda matriz regular es una matriz de cambio de base.*

Ejemplo 7.11 *Sea la base* $B = \{(-9,1),(-5,-1)\}$ *y la base* $C = \{(1,-4),(3,-5)\}$, *para hallar la matriz de cambio de base habría que resolver cada uno de los siguientes sistemas:*

$$\begin{cases} x_1 + 3x_2 = -9 \\ -4x_1 - 5x_2 = 1 \end{cases} \quad y \quad \begin{cases} x_1 + 3x_2 = -5 \\ -4x_1 - 5x_2 = -1 \end{cases}$$

Esto se obtiene calculando la matriz escalonada reducida de la matriz:

$$\left(\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 7 & -35 & -21 \end{array}\right) \rightsquigarrow$$

$$\rightsquigarrow \left(\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 1 & -5 & -3 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right)$$

Por tanto $P = \begin{pmatrix} 6 & 4 \\ -5 & -3 \end{pmatrix}$

7.5. Subespacios Vectoriales

Definición 7.8 (Subespacios vectoriales:) *Un subconjunto H del espacio vectorial B se llama subespacio vectorial de B si H es, tiene las siguientes propiedades:*

1. *El vector cero de V está en H.*
2. *H es cerrado bajo la suma de vectores. Esto es, para cada u y v en H, la suma $u+v$ está en H.*
3. *H es cerrado bajo la multiplicación por escalares. Esto es, para cada u en H y cada escalar c, el vector $c \cdot u$ está en H.*

Las propiedades 1,2 y 3 garantizan que un subespacio sea un espacio vectorial en sí mismo. Por tanto, subespacio vectorial es un espacio vectorial. Recíprocamente, todo espacio vectorial es un subespacio (de sí mismo o de espacios mayores).

Es claro que aquellas propiedades de los espacios vectoriales que son satisfechas por todos los vectores de V lo son en particular por los de cualquier subconjunto H de V. Un subconjunto H de V puede no ser subespacio de V si,

- Existen dos vectores $x, y \in H$ tales que $x+y \notin H$.
- El vector nulo $0_V \notin H$.
- Existen $x \in H$ y $\lambda \in \mathbb{K} \quad / \quad \lambda x \notin H$.

De hecho, el siguiente resultado recoge que no existen más motivos por lo que no sería subespacio:

Proposición 7.7 *Un subconjunto H del espacio vectorial V es un subespacio vectorial de V si y sólo si $O_V \in H$ y para cada $a, b \in \mathbb{R}$ y cada $x, y \in H$ se cumple que $ax+by \in H$.*

Observemos que, en particular, las combinaciones lineales de vectores de un subespacio H también pertenecen a H; esto es, si $\lambda_1, \cdots, \lambda_m \in \mathbb{R}$ y $u_1, \cdots, u_m \in H$, entonces $\lambda_1 u_1 + \cdots + \lambda_m u_m \in H$.

Ejemplo 7.12 *Dados v_1 y v_2 en un espacio vectorial V, sea $H = Gen\{v_1, v_1\}$. Demuestre que H es subespacio de V.*

El vector cero está en H, puesto que $\mathbf{0} = 0v_1 + 0v_2$. Para demostrar que H es cerrado bajo la suma de vectores, tome dos vectores arbitrarios de H, por ejemplo,

$$\mathbf{u} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2$$
$$\mathbf{w} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2$$

Por las propiedades 2,3 y 7 de los espacios vectoriales (Def. 7.1),

$$\begin{aligned} \mathbf{u}+\mathbf{w} &= (s_1\mathbf{v}_1+s_2\mathbf{v}_2)+(t_1\mathbf{v}_1+t_2\mathbf{v}_2) \\ &= (s_1+t_1)\mathbf{v_1}+(s_2+t_2)\mathbf{v_2} \end{aligned}$$

entonces $\mathbf{u}+\mathbf{w}$ *está en H. Además, si c es un escalar, entonces por las propiedades 8 y 9 (Def. 7.1).*

$$\begin{aligned} c\mathbf{u} &= c(s_1\mathbf{v}_1+s_2\mathbf{v}_2) \\ &= (cs_1)\mathbf{v_1}+(cs_2)\mathbf{v_2} \end{aligned}$$

lo que muestra que $c\mathbf{u}$ *está en H y que H es cerrado bajo la multiplicación por escalares. Entones H es un subespacio de V.*

Teorema 7.4 *Si* $\mathbf{v}_1,\cdots,\mathbf{v}_n$ *son vectores en un espacio vectorial V, entonces* $Gen\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$ *es un subespacio de V.*

Ejemplo 7.13 *Sea H el conjunto de todos los vectores de la forma* $(a-3b,b-a,a,b)$ *donde a y b son escalares arbitrarios. Esto es, sea* $H=\{(a-3b,b-a,a,b) : a,b\in\mathbb{R}\}$. *Demuestre que H es un subespacio de* $\mathbb{R}^4$.

Escribiendo los vectores de H como vectores columna. Entonces un vector arbitrario en H tiene la forma:

$$\begin{pmatrix} a-3b \\ b-a \\ a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} -3 \\ 1 \\ 0 \\ 1 \end{pmatrix} = a\mathbf{v_1}+b\mathbf{v_2}$$

Este cálculo muestra que $H=Gen\{\mathbf{v}_1,\mathbf{v}_2\}$, *donde* $\mathbf{v}_1,\mathbf{v}_2$ *son los vectores antes indicados. Entonces H es un subespacio de* $\mathbb{R}^4$ *por el teorema 7.4*

7.5.1. Subespacio generado por un conjunto de vectores

Definición 7.9 *A partir de un conjunto cualquiera de vectores de V, digamos S, podemos considerar un nuevo conjunto formado por todas las posibles combinaciones lineales de vectores de S:*

$$L(S)=\{a_1s_1+\cdots a_ns_n;\quad n\in\mathbb{N},a_i\in\mathbb{K},s_i\in S,i=1\cdots,n\}$$

Al subespacio $L(S)$ *se le llama* ***subespacio generado por S****.*

Ejemplo 7.14 *Consideremos los vectores de* $\mathbb{R}^3$ $u=(1,1,0), v=(0,0,1)$. *Entonces*

$$L(u,v)=\{a\cdot u+b\cdot v|a,b\in\mathbb{R}\}=\{(a,a,b)|a,b\in\mathbb{R}\}$$

Proposición 7.8 *$L(S)$ es el menor subespacio vectorial de V que contiene al conjunto S.*

También podemos hablar de **base de un subespacio**. Es decir, un sistema de generadores del subespacio que es linealmente independiente. Por tanto, tenemos el concepto de **dimensión de un subespacio**. Si $U\subseteq V$, entonces aplicando el Teorema de la Ampliación de la Base, el conjunto de vectores linealmente independientes que es cualquier base de U se puede ampliar a una base de V, con los que $dimU\leq dimV$. Dado un sistema de generadores de un subespacio U, podemos obtener una base de U eliminando vectores que sean combinación lineal de los demás.

Ejemplo 7.15 *Consideremos el subespacio de* $\mathbb{R}^4$ $U=L(1,3,4,1),(2,6,8,2),(2,5,7,2)]$. *Los tres vectores dados son sistemas de generadores de U pero no son base porque no son linealmente independientes, ya que el rango de la matriz*

$$\begin{pmatrix}1&3&4&1\\2&6&8&2\\2&5&7&2\end{pmatrix}$$

es 2, y por tanto sólo hay 2 vectores linealmente independientes. Podemos elegir $\{(1,3,4,1),(2,5,7,2)\}$ *como base.*

7.5.2. Ecuaciones cartesianas y paramétricas de un subespacio

En el siguiente ejemplo vemos cómo un sistemas de ecuaciones lineales determina un subespacio vectorial de un cierto $\mathbb{K}^n$.

Ejemplo 7.16 *Supongamos que tenemos un sistema homogéneo de m ecuaciones con n incógnitas AX=0. Cada solución del sistema es un elemento del conjunto* $\mathbb{K}^n$. *El conjunto formado por todas las soluciones es un subespacio vectorial de* $\mathbb{K}^n$.

En efecto, si $(s_1,\cdots,s_n)$ *y* $(t_1,\cdots,t_n)$ *son soluciones del sistema,* $(s_1+t_1,\cdots,s_n+t_n)$ *y* $(k\cdot s_1,\cdots,k\cdots_n)$ *son soluciones del sistema.*

Por tanto, un subespacio vectorial de un espacio vectorial V de dimensión finita puede ser interpretado como el conjunto de soluciones de un sistema de ecuaciones. Para

ello, debemos considerar una base de V, digamos $B = \{e_1, \cdots, e_n\}$ respecto de la cual consideraremos las coordenadas de cada vector.

Si U es un subespacio vectorial de dimensión $r \leq n$, entonces podemos tomar la base de U, formada por r vectores: $B_U = \{u_1, \cdots, u_r\}$. Supongamos que conocemos los vectores u_i en función de la base B: $u_i = (a_{1i}, \cdots, a_{ni})_B$ con $i = 1, \cdots, r$. Cualquier vector $x \in U$ se expresa como combinación lineal de los u_i, digamos $x = \lambda_1 u_1 + \cdots \lambda_r u_r$. Si $x = (x_1, \cdots, x_n)_B$ entonces tenemos:

$$(x_1, \cdots, x_n) \quad = \quad \lambda_1(a_{11}, \cdots, a_{n1}) + \cdots + \lambda_r(a_{1r}, \cdots, a_{nr}) +$$

e igualando obtenemos las **ecuaciones paramétricas** de U respecto de la base B:

$$\begin{cases} x_1 = & a_{11} & \lambda_1 & + & a_{12} & \lambda_2 & + & \cdots & a_{1r} & \lambda_r \\ x_2 = & a_{21} & \lambda_1 & + & a_{22} & \lambda_2 & + & \cdots & a_{2r} & \lambda_r \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \ddots & \vdots & \vdots \\ x_n = & a_{n1} & \lambda_1 & + & a_{n2} & \lambda_2 & + & \cdots & a_{nr} & \lambda_r \end{cases}$$

De las ecuaciones de un sistema homogéneo cuyo conjunto de soluciones sea el anterior diremos que son unas **ecuaciones cartesianas o implícitas** de U respecto de la base B. Estas ecuaciones nos dan las condiciones que tienen que cumplir las coordenadas de un vector para que pertenezca al subespacio en cuestión.

Ejemplo 7.17 *Consideremos el subespacio vectorial de $\mathbb{R}^3$ formado por las soluciones del sistema homogéneo*

$$x_1 + x_2 + x_3 = 0$$

Despejando una de las incógnitas tenemos

$$x_1 = -x_2 - x_3$$

Tomando x_2, x_3 como parámetros, digamos λ, μ, podemos escribir

$$\begin{cases} x_1 = & -1 \cdot \lambda & -1 \cdot \mu \\ x_2 = & 1 \cdot \lambda & +0 \cdot \mu \\ x_1 = & 0 \cdot \lambda & +1 \cdot \mu \end{cases}$$

Es decir, todo vector del espacio es combinación lineal de los vectores formados por los coeficientes de cada uno de los parámetros. Tenemos así que $\{(-1,1,0),(-1,0,1)\}$ es una base del subespacio.

Si llamamos $n = dimV$ y $r = dimU$ entonces en unas ecuaciones paramétricas de U aparecen r parámetros. Dado un sistema homogéneo con n incógnitas, para que la solución dependa de r parámetros es necesario que la matriz de coeficientes tenga rango $n-r$, con lo que debe tener, al menor, $n-r$ ecuaciones. Por otra parte, si tiene más de $n-r$ ecuaciones, las que exceden pueden hacerse cero mediantes transformaciones elementales del sistema. Así, se tiene que

$$\text{nº ecuaciones cartesianas} = dimV - dimU \tag{7.2}$$

Ejemplo 7.18 *Consideremos en $\mathbb{R}^3$ el subespacio U generado por los vectores (1,-1,0),(1,1,0), entonces unas ecuaciones paramétricas de U son:*

$$U \equiv \begin{cases} x_1 &= \lambda + &\mu \\ x_2 &= -\lambda + &\mu \\ x_3 &= 0 & \end{cases}$$

Como U tiene dimensión 2 en un espacio de dimensión 3, solo necesitamos 3-2=1 ecuación cartesiana para describir U, y evidentemente es $x_3 = 0$.

Ejemplo 7.19 *Sea U el subespacio con base $\{(1,0,1,1),(0,1,1,0)\}$. A partir de ella, obtenemos las ecuaciones parámetricas*

$$U \equiv \begin{cases} x_1 &= \lambda & \\ x_2 &= & \mu \\ x_2 &= \lambda + &\mu \\ x_4 &= \lambda & \end{cases}$$

Obtenemos las ecuaciones cartesianas eliminando parámetros. Como U es un subespacio de $\mathbb{R}^4$ de dimensión 2, necesitaremos dos ecuaciones cartesianas:

$$U \equiv \begin{cases} x_1 &= \lambda \\ x_2 &= \mu \\ x_2 &= \lambda + \mu \\ x_4 &= \lambda \end{cases} \rightsquigarrow U \equiv \begin{cases} x_2 &= \mu \\ x_3 - x_1 &= \mu \\ x_4 - x_1 &= 0 \end{cases} \rightsquigarrow$$

$$\rightsquigarrow U \equiv \begin{cases} x_3 - x_1 - x_2 &= 0 \\ x_4 - x_1 &= 0 \end{cases} \rightsquigarrow U \equiv \begin{cases} x_1 + _2 - x_3 &= 0 \\ x_4 - x1 &= 0 \end{cases}$$

Universidad Rey Juan Carlos

Tema 8

Diagonalización de Matrices

8.1. Transformaciones Lineales

Supongamos que la transformación lineal A en la base canónica transforma el punto X en el punto X'.

$$A \cdot X = X'$$

Si la transformación A fuese una matriz diagonal, la transformación que realizaría en X consistiría en multiplicar por un escalar en cada una de las direcciones marcadas por la base canónica.

Por ejemplo, si A es:

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

y X es:

$$X = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

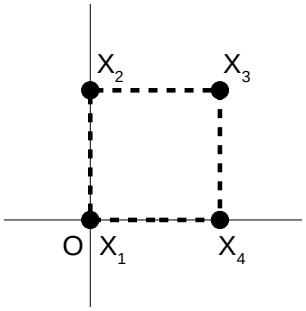
Entonces, X' sería:

$$X' = \begin{pmatrix} 2 \cdot 1 \\ 3 \cdot 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 9 \end{pmatrix}$$

Las transformaciones diagonales son muy fáciles de entender porque expresan muy claramente cuánto cambia en cada dirección un punto. Por ejemplo, la transformación A nos indica que en una dirección no cambia (se multiplica por 1) y en la otra dirección se multiplica por 3.

Desgraciadamente la mayoría de las aplicaciones no son diagonales. Por ejemplo, imaginemos que tenemos los 4 puntos de la Figura 1 en la base canónica 2D.

$$X_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, X_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, X_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ y } X_4 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$



Considérese la siguiente transformación A:

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

Esta transformación convierte los puntos X_1, X_2, X_3 y X_4 en los puntos X_1', X_2', X_3' y X_4' de la siguiente manera:

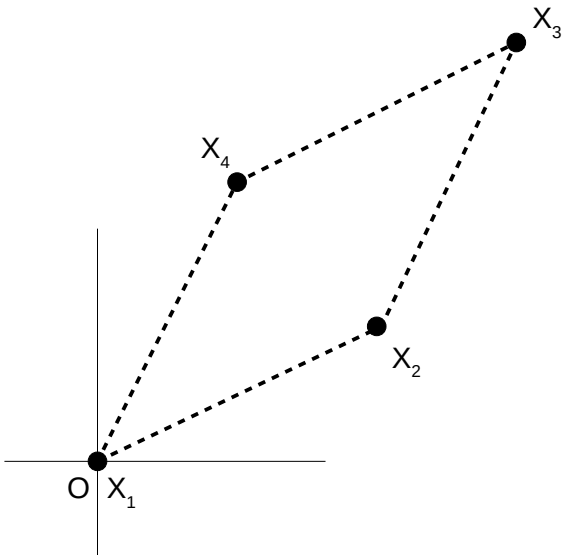
$$X_1' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$X_2' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$X_3' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

$$X_4' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

La siguiente figura ilustra cómo se han transformado los puntos de la Figura 1.



Debe notarse que observando la transformación A no es fácil descubrir cuál será su resultado sobre la Figura 1.

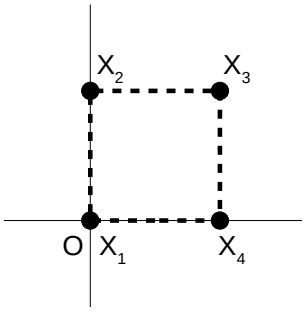
Además,

- Al aplicar la base P sobre un punto X obtenemos la representación de X en la base P que llamaremos X_P.
$$P \cdot X = X_P \tag{8.1}$$
- De la misma forma, al aplicar la base P sobre un punto X' obtenemos la representación de X' en la base P que llamaremos X'_P.
$$P \cdot X' = X'_P \tag{8.2}$$
- Además, en la base P la transformación D transforma X_P en X'_P.
$$D \cdot X_P = X'_P \tag{8.3}$$

Uniendo estas observaciones obtenemos un camino para obtener P desde A.

Ilustremos esto con un ejemplo numérico. Imaginemos que tenemos los 4 puntos de la Figura 1 en la base canónica de 2D.

$$X_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, X_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, X_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ y } x_4 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$



Considérese la siguiente transformación A:

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

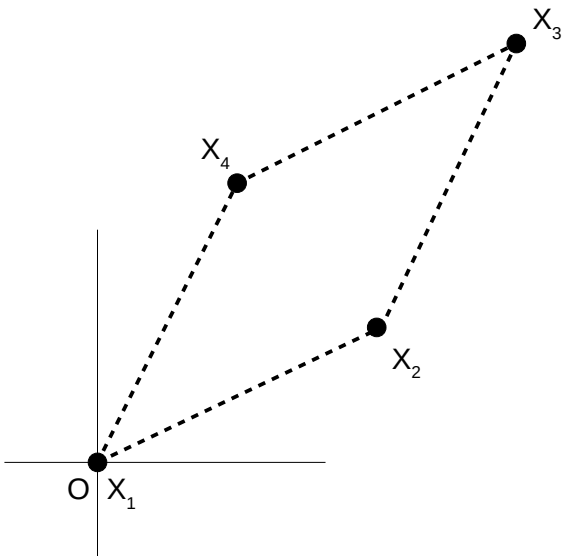
Esta transformación convierte los puntos X_1, X_2, X_3 y X_4 en los puntos X_1', X_2', X_3' y X_4' de la siguiente manera:

$$X_1' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

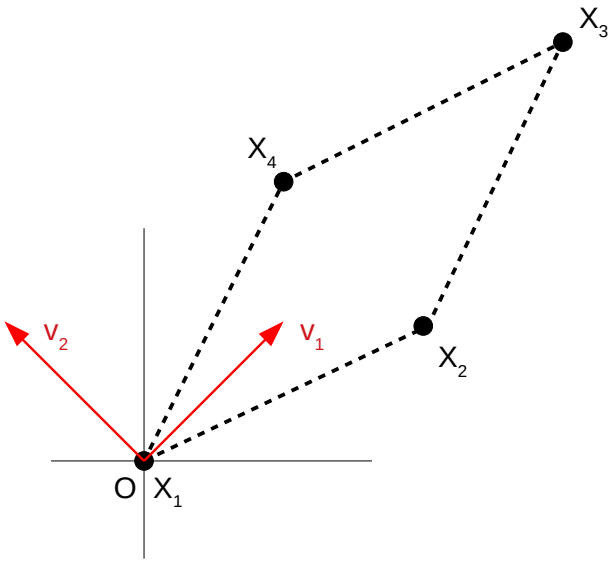
$$X_2' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$X_3' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

$$X_4' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$



Debe notarse que observando la transformación A no es fácil descubrir cúal es su resultado sobre la figura formada por los puntos x_1, x_2, x_3 y x_4.



Sin embargo, si se aplica la transformación P a los vectores v_1, v_2 que definen su sistema de referencia, estos vectores mantienen su dirección, Por eso, a estos vectores se les llama autovectores.

8.2. Autovalores y autovectores

En esta sección trabajaremos con transformaciones que vienen definidas mediante matrices

$$y = Ax$$

Aunque una transformación puede mover vectores en diversas direcciones, frecuentemente sucede que existen vectores especiales sobre los cuales la acción de A es muy sencilla.

Ejemplo 8.1 *Sea* $A=\begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}$, $u=\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ *y* $v=\begin{pmatrix} 2 \\ 1 \end{pmatrix}$

Como $Au=\begin{pmatrix} -5 \\ -1 \end{pmatrix}$

Como $Av=\begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2v$

Por tanto, A mueve el vector u mientras que a v solamente lo "estira" o dilata.

El objetivo de este tema será buscar vectores que sean transformados por A en múltiplos escalares de sí mismos.

Definición 8.1 *Un* ***vector propio o autovector*** *de una matriz* $A \in \mathcal{M}_n$ *es un vector* x *diferente de 0 tal que* $Ax = \lambda x$ *para algún escalar* λ.

Definición 8.2 *Un escalar* λ *se llama* ***valor propio o autovalor*** *de A si existe una solución no trivial* x *de* $Ax = \lambda x$. *Una* x *tal se llama* ***vector propio*** *correspondiente a* λ.

Ejemplo 8.2 *Sea* $A=\begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}$, $u=\begin{pmatrix} 6 \\ -5 \end{pmatrix}$ *y* $v=\begin{pmatrix} 3 \\ -2 \end{pmatrix}$.

¿Son u y v vectores propios de A?

$$Au = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 6 \\ -5 \end{pmatrix} = \begin{pmatrix} -24 \\ 20 \end{pmatrix} = -4 \begin{pmatrix} 6 \\ -5 \end{pmatrix} = -4u$$

$$Av = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} -9 \\ 11 \end{pmatrix} \neq \lambda \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

Por tanto, u es vector propio correspondiente al valor propio -4, pero v no.

Ejemplo 8.3 *Demuestre que 7 es un valor propio de A del ejemplo anterior, y encuentre los vectores propios correspondientes.*

Es escalar 7 es un valor propio de A si y sólo si la ecuación $Ax = 7x$ tiene una solución no trivial. Como

$$Ax = 7x \Leftrightarrow Ax - 7x = 0 \Leftrightarrow (A - 7I)x = 0$$

tenemos que resolver la ecuación homogénea:

$$\left(\begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix} - \begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -6 & 6 \\ 5 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Como rg(A-7I)=1, el sistema tiene soluciones no triviales. Por tanto, 7 es un valor propio de A. Para encontrar los vectores resolvemos el sistema

$$\begin{pmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

luego tenemos la solución

$$x_1 = \mu$$
$$x_2 = \mu$$

Por tanto, cada vector de esta forma con $\mu \neq 0$ es un vector propio correspondiente a $\lambda = 7$

De este ejemplo se deduce que λ es un valor propio si y sólo si la ecuación

$$(A - \lambda I)x = 0$$

tiene una solución no trivial.

Definición 8.3 *El conjunto de todas las soluciones de la ecuación $(A - \lambda I)x = 0$ es un subespacio denominado **espacio propio** de A correspondiente a λ. Es decir, el espacio propio consiste en los vectores 0 y todos los vectores propios correspondientes a λ.*

Ejemplo 8.4 *Sea* $A=\begin{pmatrix}4 & -1 & 6\\ 2 & 1 & 6\\ 2 & -1 & 8\end{pmatrix}$. *Un valor propio de A es 2. Encuentre una base para el espacio propio correspondiente.*

Como

$$A-2I=\begin{pmatrix}4 & -1 & 6\\ 2 & 1 & 6\\ 2 & -1 & 8\end{pmatrix}-\begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}=\begin{pmatrix}2 & -1 & 6\\ 2 & -1 & -6\\ 2 & -1 & 6\end{pmatrix}$$

y la matriz ampliada cumple que

$$\begin{pmatrix}2 & -1 & 6 & 0\\ 2 & -1 & -6 & 0\\ 2 & -1 & 6 & 0\end{pmatrix}\rightsquigarrow\begin{pmatrix}2 & -1 & 6 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$$

donde $rg(\widetilde{A})=1<3$ *existen infinitas soluciones y 2 es autovalor de A. Por otra parte, la solución general viene dada por*

$$\begin{cases}x_1=\frac{1}{2}\mu-3\gamma\\ x_2=\mu\\ x_3=\gamma\end{cases}$$

De aquí se obtiene que el espacio propio es un subespacio bidimensional de $\mathbb{R}^3$ *con base* $\{(\frac{1}{2},1,0)(-3,0,1)\}$.

Teorema 8.1 *Los valores propios de una matriz triangular son las entradas de su diagonal principal.*

Ejemplo 8.5 *Sea* $A=\begin{pmatrix}3 & 6 & -8\\ 0 & 0 & 6\\ 0 & 0 & 2\end{pmatrix}$ *y* $B=\begin{pmatrix}4 & 0 & 0\\ -2 & 1 & 0\\ 5 & 3 & 4\end{pmatrix}$. *Los valores propios de A son 3,0 y 2. Los valores propios de B son 4 y 1.*

Nótese que $Ax=0x$ tiene una solución no trivial si y sólo si A no es invertible. Esto nos lleva al siguiente resultado:

Teorema 8.2 (Teorema de la Matriz Invertible:) *Sea* $A\in\mathcal{M}_n$, *entonces A es invertible si y sólo si 0 no es valor propio de A.*

Teorema 8.3 *Si $v_1, \cdots, v_n$ son vectores propios que corresponden a distintos valores propios $\lambda_1, \cdots, \lambda_n$ de $A \in \mathcal{M}_n$, entonces el conjunto $\{v_1, \cdots, v_n\}$ es linealmente independiente.*

8.2.1. Polinomio Característico

Cierta información útil acerca de los valores propios de una matriz cuadrada A se encuentra codificada en una ecuación escalar llamada polinomio característico o ecuación característica de A.

Ejemplo 8.6 *Para encontrar los valores propios de* $A = \begin{pmatrix} 2 & 3 \\ 3 & -6 \end{pmatrix}$ *debemos encontrar todos los escalares λ tales que*

$$(A - \lambda I)x = 0$$

tenga solución no trivial. Por el teorema de la matriz invertible, este problema es equivalente a encontrar todas las λ tales que la matriz $(A - \lambda I)$ no sea invertible, donde

$$A - \lambda I = \begin{pmatrix} 2 & 3 \\ 3 & -6 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{pmatrix} = 0$$

Sabemos que una matriz no será invertible cuando su determinante sea 0, es decir cuando

$$\begin{aligned} det(A - \lambda I) &= det \begin{pmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{pmatrix} = (2-\lambda)\cdot(-6-\lambda) - 3\cdot 3 = \\ &= -12 + 6\lambda - 2\lambda + \lambda^2 - 9 = \\ &= \lambda^2 + 4\lambda - 21 = 0 \Leftrightarrow \\ &\Leftrightarrow (\lambda - 3)\cdot(\lambda + 7) = 0 \end{aligned}$$

Por tanto, los valores propios de A son 3 y -7.

Por tanto, podemos utilizar un determinante para decidir cuándo la matriz $A - \lambda I$ no es invertible.

Definición 8.4 *El **polinomio característico** de una matriz $A \in \mathcal{M}_n$ viene determinado por la expresión*

$$det(A - \lambda I) = 0 \tag{8.4}$$

Proposición 8.1 *Un escalar λ es u autovalor de $A \in \mathcal{M}_n$ si y sólo si λ satisface el polinomio característico*

$$det(A - \lambda I) = 0 \tag{8.5}$$

Ejemplo 8.7 *Encuentre la ecuación característica de*

$$A = \begin{pmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & 8 & 2 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

La ecuación característica viene dada por la expresión

$$det(A-\lambda I) = det \begin{pmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & 8 & 2 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{pmatrix} = (5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda) = 0$$

O equivalentemente

$$(5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda) = (5-\lambda)^2(3-\lambda)(1-\lambda) = \lambda^4 - 14\lambda^3 + 68v^2 - 130\lambda + 75$$

Definición 8.5 *Se define la **multiplicidad algebraica** de un valor propio λ a su multiplicidad como raíz de su polinomio característico.*

Ejemplo 8.8 *El polinomio característico de una matriz $A \in \mathcal{M}_6$ es $\lambda^6 - 4\lambda^5 - 12\lambda^4$.*

Factorizando el polinomio se tiene que

$$\lambda^6 - 4\lambda^5 - 12\lambda^4 = \lambda^4(\lambda^2 - 4\lambda - 12) = \lambda^4(\lambda - 6)(\lambda + 2)$$

Por tanto, los valores propios son 0 con multiplicidad 4, 6 con multiplicidad 1 y -2 con multiplicidad 1.

El siguiente teorema ilustra un uso del polinomio característico.

Definición 8.6 *Dos matrices cuadradas de A y B son **semejantes** si existe una matriz regular P de forma que $B = P^{-1}AP$.*

Nótese que convertir A en $P^{-1}AP$ constituye una transformación de semejanza.

Por tanto, sea A es una matriz de cambio de la base canónica a una base B, si D es un matriz semejante a A, $D = P^{-1}AP$ para cierta matriz regular P, entonces D será la matriz de cambio de base de la base canónica a una base B'.

Ejemplo 8.9 *Sea la matriz* $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$ *y* $P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ *la matriz de cambio de base de B' a B, entonces, la matriz de cambio de la base canónica a la base B' es:*

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

Teorema 8.4 *Si las matrices* $A, B \in \mathcal{M}_n$ *son semejantes, entonces tienen el mismo polinomio característico y por lo tanto los mismos autovalores con las mismas multiplicidades.*

8.3. Diagonalización de matrices

En muchos casos, la información de vector propio-valor propio contenida dentro de una matriz A se puede mostrar con una útil factorización de la forma $A = P^{-1}DP$. En esta sección, la factorización nos permite calcular rápidamente A^k para valores grandes de k.

Ejemplo 8.10 *Si* $D = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}$, *entonces* $D^2 = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 5^2 & 0 \\ 0 & 3^2 \end{pmatrix}$.

En general

$$D^k = \begin{pmatrix} 5^k & 0 \\ 0 & 3^k \end{pmatrix} \quad \forall k \geq 1$$

Si $A = PDP^{-1}$ para alguna P invertible y D diagonal, entonces también es fácil calcular A^k, como muestra el siguiente ejemplo.

Ejemplo 8.11 *Sea* $A = \begin{pmatrix} 7 & 2 \\ -4 & 1 \end{pmatrix}$. *Encuentre una fórmula para* A^k, *dado que* $A =$

PDP^{-1}, donde

$$P = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \quad y \quad D = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}$$

Como

$$P^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$$

Entonces, por la propiedad asociativa de la multiplicación de matrices,

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDDP^{-1} = PD^2P^{-1}$$

Análogamente,

$$A^3 = (PDP^{-1})A^2 = (PDP^{-1})(PD^2P^{-1}) = PDDDP^{-1} = PD^3P^{-1}$$

En general, para $k \geq 1$,

$$\begin{aligned} A^k = PD^kP^{-1} &= \\ &= \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 5^k & 0 \\ 0 & 3^k \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} = \\ &= \begin{pmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{pmatrix} \end{aligned}$$

Definición 8.7 *Decimos que una matriz cuadrada A es diagonalizable si A es semejante a una matriz diagonal, esto es, si $A = PDP^{-1}$ para alguna matriz P invertible y alguna matriz D diagonal.*

Por tanto, el problema que nos planteamos es, dada una matriz cuadrada A, encontrar una matriz D semejante a A de forma que D sea diagonal.

Teorema 8.5 *Una matriz $A \in \mathcal{M}n(\mathbb{R})$ es diagonalizable si y sólo si A tiene n vectores propios linealmente independientes.*

De hecho, $A = PDP^{-1}$, siendo D una matriz diagonal, si y sólo si las columnas de P son n vectores propios de A linealmente independientes. En este caso, las entradas diagonales de D son los valores propios de A que corresponden, respectivamente, a los vectores propios de P.

En otras palabras, A es diagonalizable si y sólo si hay suficientes vectores propios para formar una base de $\mathbb{R}^n (\Leftrightarrow \sum_i dim(\lambda_i) = n)$.

Definición 8.8 *La base formada por los vectores propios de una matriz $A \in \mathcal{M}^n(\mathbb{R})$. se denomina* ***base de vectores propios****.*

El siguiente resultado muestra cómo encontrar los vectores propios y, por tanto, cómo **diagonalizar una matriz**.

Proposición 8.2 (Algoritmo de Diagonalización de Matrices:)

1. ***Encontrar los valores propios de A****: A partir de la ecuación característica se obtienen los valores propios y su multiplicidad.*
2. ***Encontrar los vectores propios de A linealmente independientes****: Calcular el espacio propio de cada λ autovalor de A a partir de la ecuación $(A-\lambda I)x = 0$.*
3. ***Construir P*** *la matriz de coordenadas de la base de vectores propios. Si $\{v_1, \cdots, v_n\}$ son los autovectores de A, entoncesP $= (v_1 v_2 \cdots v_n)$.*
4. ***Construir D****: La matriz $D = P^{-1}AP$ será una matriz diagonal donde los elementos de la diagonal son los autovectores ordenados de mayor a menor.*

Ejemplo 8.12 *Vamos a diagonalizar* $A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}$. *Para ello:*

1. *Obtenemos los autovalores:*
   A partir de la ecuación característica se tiene que:

$$\begin{aligned} 0 &= det(A-\lambda I) = det\left(\begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right) = \\ &= det\begin{pmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{pmatrix} = det\begin{pmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 0 & -2-\lambda & -2-\lambda \end{pmatrix} = \\ &= (1-\lambda)det\begin{pmatrix} -5-\lambda & -3 \\ -2-\lambda & -2-\lambda \end{pmatrix} - (-3)det\begin{pmatrix} 3 & 3 \\ -2-\lambda & -2-\lambda \end{pmatrix} = \\ &= (1-\lambda)[(-5-\lambda)(-2-\lambda) - (-3)(-2-\lambda)] + 3\cdot 0 = \\ &= (1-\lambda)(10+5\lambda+2\lambda+\lambda^2-6-3\lambda) = \\ &= (1-\lambda)(\lambda^2+4\lambda+4) = (1-\lambda)(\lambda+2)^2 \end{aligned}$$

   Por tanto, los autovalores son $\lambda = 1$ con multiplicidad 1 y $\lambda = -2$ con multiplicidad 2.

2. *Obtenemos los autovectores:*

   Dado que $A \in \mathcal{M}_3$, necesitamos 3 vectores. Éste es el paso crítico. Si falla, entonces el teorema anterior dice que A no puede diagonalizarse. luego $v_1 = (1,-1,1)$ es autovector de A asociado a $\lambda = 1$.

   Análogamente para $\lambda = -2$,resolvemos $(A+2I)x = 0$

$$(A+2I)x = 0 \Leftrightarrow \left(\begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix} + 2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) x = 0 \Leftrightarrow$$

$$\Leftrightarrow \begin{pmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{pmatrix} x = 0 \Leftrightarrow \begin{cases} 3x_1 & 3x_2 & +3x_3 & = 0 \\ -3x_1 & -3x_2 & -3x_3 & = 0 \\ 3x_1 & +3x_2 & +3x_3 & = 0 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \{x_1 + x_2 + x_3 = 0 \Leftrightarrow \{x_1 = -x_2 - x_3 \Leftrightarrow \begin{cases} x_1 = -\mu - \beta \\ x_2 = \mu \\ x_3 = \beta \end{cases}$$

   Luego los vectores $v_2 = (-1,1,0), v_3 = (-1,0,1)$ son autovectores de A asociados a $\lambda = -2$.

3. *Obtenemos la matriz de coordenadas de la base propia:*

   La matriz de coordenadas será:

$$P = (v_1 v_2 v_3) = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

4. *Obtenemos la matriz diagonalizada:*

   D será la matriz diagonal cuyos elementos no cero son los autovalores ordenados de mayor a menor, es decir

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

Sin embargo, no todas las matrices son diagonalizables. El siguiente resultado, da una condición suficiente para que una matriz sea diagonalizable:

Teorema 8.6 *Una matriz $A \in \mathcal{M}_n(\mathbb{R})$ con n valores propios distintos es diagonalizable.*

Sin embargo, para ser diagonalizable no es necesario que una matriz cuadrada de orden n tenga n autovalores disintintos, como hemos visto en el ejemplo anterior. Cuando A es diagonalizable pero tiene menos de n valores distintos, aún es posible construir P de alguna forma que haga a P automáticamente invertible, como muestra el siguiente teorema.

Teorema 8.7 *Sea $A \in \mathcal{M}_n(\mathbb{R})$ cuyos valores propios distintos son $\lambda_1, \cdots, \lambda_p$*

1. *Para $1 \leq k \leq p$, la dimensión del espacio propio para λ_k es menor o igual a la multiplicidad del valor propio λ_k.*
2. *La matriz A es diagonalizable si y sólo si la suma de las dimensiones de los distintos espacios propios es igual a n y esto sucede si y sólo si la dimensión del espacio propio para cada λ_k es igual a la multiplicidad de λ_k.*
3. *Si A es diagonalizable y $\mathcal{B}_k$ es una base para el espacio correspondiente a λ_k para cada k, entonces la colección total de vectores en los conjuntos $\mathcal{B}_1, \cdots \mathcal{B}_p$ forma una base de vectores propios para $\mathbb{R}^n$.*

Ejemplo 8.13 *Queremos diagonalizar la matriz* $A = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 1 & 4 & -3 & 0 \\ -1 & -2 & 0 & -3 \end{pmatrix}$.

Como A es una matriz triangular, $det(A - \lambda I) = (5-\lambda)^2(-3-\lambda)^2$, luego los autonvalores serán 5 y -3, cada uno de ellos con multiplicidad 2. Resolviendo $(A-5I)x = 0$, obtenemos los vectores:

$$\begin{pmatrix} -8 \\ 4 \\ 1 \\ 0 \end{pmatrix} y \begin{pmatrix} -16 \\ 4 \\ 0 \\ 1 \end{pmatrix}$$

y resolviendo $(A+3I)x = 0$, obtenemos los vectores:

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

Por el teorema anterior, como $\{v_1, v_2, v_3, v_4\}$ es linealmente independiente,luego la suma de los espacios propios es 4, y $P = [v_1 v_2 v_3 v_4]$ es invertible y A es diagonalizable,

con

$$P = \begin{pmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \quad y \quad D = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}$$

8.3.1. Diagonalización en Matrices Simétricas

Las matrices simétricas surgen en las aplicaciones, de una u otra manera, con mayor otra frecuencia que cualquier otra clse importante de matrices. Por ello, nos centramos en las características de la diagonalización de este tipo de matrices.

Para ello definimos previamente los siguiente conceptos:

Definición 8.9 (Base ortogonal:) *Se dice que una base es ortogonal si los vectores que la forman son ortogonales dos a dos, es decir* $< e_i, e_j >= 0 \quad \forall i \neq j$.

Definición 8.10 (Base ortonormal:) *Se dice que B es una base ortonormal si es ortogonal y además todos los vectores que la forman tienen norma 1, es decir* $||e_i|| = \sqrt{x_1^2 + \cdots x_n^2} = 1$.

Definición 8.11 (Matrices ortogonal:) *Una matriz P es ortogonal si verifica*

$$P^T = P^{-1}$$

Si P es una matriz ortogonal estará formada por vectores ortonormales.

Proposición 8.3 *La matriz de cambio de base entre dos bases ortonormales es una matriz diagonal.*

Para plantear la diagonalización en matrices simétricas, estudiamos previamente el siguiente ejemplo:

Ejemplo 8.14 *Se quiere diagonalizar la matriz* $A = \begin{pmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{pmatrix}$.

La ecuación característica de A es

$$det(A-\lambda I)=0 \Leftrightarrow -(\lambda-8)(\lambda-6)(\lambda-3)=0$$

Luego los autovectores son 8, 6 y 3. Resolviendo los sistemas $(A-8I)x=0$, $(A-6I)x=0$, $(A-3I)x=0$ *obtenemos, respectivamente, los autovectores*

$$\lambda=8:\ v_1=\begin{pmatrix}-1\\1\\0\end{pmatrix}\quad \lambda=6:\ v_2=\begin{pmatrix}-1\\-1\\2\end{pmatrix}\quad \lambda=3:\ v_3=\begin{pmatrix}1\\1\\1\end{pmatrix}$$

Estos tres vectores forman una base para $\mathbb{R}^3$ *y los podemos usar como columnas para una matriz P que diagonalice A. Sin embargo, se puede comprobar que* $\{v_1,v_2,v_3\}$ *es un conjunto ortogonal, y P será más útil si sus columnas son ortonormales. Para ello, normalizamos los vectores* $\{v_1,v_2,v_3\}$ *obteniendo*

$$u_1=\begin{pmatrix}-1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}\quad u_2=\begin{pmatrix}-1/\sqrt{6}\\-1/\sqrt{6}\\2/\sqrt{6}\end{pmatrix}\quad u_3=\begin{pmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{pmatrix}$$

entonces tenemos que $P=[u_1u_2u_3]$ *será una matriz ortonormal.*

El siguiente teorema explica por qué los vecotres propios del ejemplo anterior son ortogonales:

Teorema 8.8 *Si A es simétrica, entonces cualesquier par de vectores propios de espacios propios diferentes son ortogonales.*

De este modo definimos:

Definición 8.12 *Se dice que A es* ***diagonalizable ortogonalmente*** *si hay una matriz ortogonal P y una matriz diagonal D tales que* $A=PDP^T=PDP^{-1}$.

Para diagonalizar ortogonalmente a una matriz $A\in\mathcal{M}_n$, debemos encontrar n vectores propios linealmente independientes y ortonormales. El siguiente resultado nos dice cuando eso es posible.

Teorema 8.9 *Una matriz* $A\in\mathcal{M}_n$ *es diagonalizable ortogonalmente si y sólo si A es una matriz simétrica.*

El siguiente ejemplo trata una matriz cuyos valores propios o son todos diferentes.

Ejemplo 8.15 *Queremos diagonalizar ortogonalmente* $A = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix}$

cuya ecuación característica es

$$0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2).$$

Luego los autovalores y autovectores correspondientes son:

$$\lambda = 7 : v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} v_2 = \begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix} \qquad \lambda = -2 : v_3 = \begin{pmatrix} -1 \\ -1/2 \\ 1 \end{pmatrix}$$

Aunque v_1, v_2 *son linealmente independientes, no son ortogonales pues* $\langle v_1, v_2 \rangle = -1/2 \neq 0$. *Sin embargo, si calculamos la proyección de* v_2 *sobre* v_1

$$z_2 = v_2 - \frac{\langle v_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{pmatrix} -1/1 \\ 1 \\ 0 \end{pmatrix} - \frac{-1/2}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/4 \\ 1 \\ 1/4 \end{pmatrix}$$

tenemos que $\{v_1, z_2\}$ *es un conjunto ortogonal es el espacio propio para* $\lambda = 7$. *Además, como* z_2 *es combinación lineal de* v_1 *y* v_2, z_2 *sigue estando en el espacio propio. Por otra parte, como el espacio propio es bidimensional,* $\{v_1, z_2\}$ *es una base ortogonal para el espacio propio.*

De este modo, normalizando $\{v_1, z_2\}$ *obtenemos la siguiente base ortonormal para el espacio propio* $\lambda = 7$*:*

$$u_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix} \quad u_2 = \begin{pmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{pmatrix}$$

Análogamente, una base ortonormal para el espacio con $\lambda = -2$ *será*

$$u_3 = \frac{1}{||2v_3||} 2v_3 = \frac{1}{3} \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix}$$

Por tanto, por el teorema 8.8, u_3 es ortogonal a los vectores propios u_1, u_2. Por lo tanto, $\{u_1, u_2, u_3\}$ es un conjunto ortonormal. Sea

$$P = [u_1 u_2 u_3] = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{pmatrix}$$

$$D = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Entonces P diagonaliza ortogonalmente a A, y $A = PDP^T$.

8.3.2. Teorema Espectral

Definición 8.13 *El conjunto de valores propios de una matriz A se conoce como **espectro** de A.*

Teorema 8.10 (Teorema Espectral:) *Una matriz simétrica $A \in \mathcal{M}_n$ tiene las siguientes propiedades:*

1. *A tiene n valores propios reales, contando multiplicidades.*
2. *La dimensión del espacio propio para cada valor propio λ es igual a la multiplicidad de λ como raíz de la ecuación característica.*
3. *Los espacios propios son mutuamente ortogonales, en el sentido de que vectores propios correspondientes a valores propios diferentes son ortogonales.*
4. *A es diagonalizable ortogonalemente.*

Descomposición Espectral

Suponga que $A = PDP^{-1}$, done las columnas de P son vectores propios ortonormales $u_1, \cdots, u_n$ de A y los valores propios correspondientes $\lambda_1, \cdots, \lambda_n$ están en la matriz diagonal D. Entonces, como $P^{-1} = P^T$

$$\begin{aligned} A = PDP^T &= (u_1 \cdots u_n) \begin{pmatrix} \lambda_1 & & 0 \\ & \vdots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} u_1^T \\ \cdots \\ u_n^T \end{pmatrix} = \\ &= (\lambda_1 u_1 \cdots \lambda_n u_n) \begin{pmatrix} u_1^T \\ \cdots \\ u_n^T \end{pmatrix} = \lambda_1 u_1 u_1^T + \cdots + \lambda_n u_n u_n^T \end{aligned}$$

Definición 8.14 *A la representación de una matriz*

$$A = \lambda_1 u_1 u_1^T + \cdots + \lambda_n u_n u_n^T$$

Con $A \in \mathcal{M}_n$ se conoce como ***descomposición espectral*** *de A.*

Ejemplo 8.16 *Vamos a construir la descomposición espectral de la matriz A que tiene la diagonalización ortogonal:*

$$A = \begin{pmatrix} 7 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 8 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$$

Sea $P = [u_1 u_2]$, entonces

$$A = 8u_1 u_1^T + 3u_2 u_2^T$$

Para verificar esta descomposición, calculamos

$$\begin{aligned} u_1 u_1^T &= \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} = \begin{pmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{pmatrix} \\ u_2 u_2^T &= \begin{pmatrix} -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} = \begin{pmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{pmatrix} \end{aligned}$$

y

$$8u_1 u_1^T + 3u_2 u_2^T = 8 \cdot \begin{pmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{pmatrix} + 3 \begin{pmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{pmatrix} = \begin{pmatrix} 7 & 2 \\ 2 & 4 \end{pmatrix} = A$$

Universidad Rey Juan Carlos

Tema 9

Espacios euclídeos y aproximación por mínimo cuadrados

9.1. Producto Escalar

En este tema veremos como, al introducir un producto escalar, se pueden llevar a los espacios vectoriales nociones geométricas de ortogonalidad, ángulo, longitud, distancias, áreas, etc.

La geometría euclídea se desarrolla en los siglos XIX y XX, tras la aparición del concepto de espacio vectorial. Recibe su nombre en honor a Euclides, matemático griego ($\sim$300 a.C.) quien estudió los conceptos básicos de la Geometría plana, aunque por supuesto no en un contexto vectorial.

Para generalizar conceptos geométricos, observaremos, entre otros, el comportamiento de los vectores del plano. Por ejemplo, en R^2 tenemos definido el producto escalar usual

$$(a1,a2)\cdot(b1,b2) = a1\cdot b1 + a2\cdot b2$$

que es una operación entre dos vectores, cuyo resultado es un escalar (de ahí el nombre "producto escalar").

El producto escalar permite, entre otras cosas, reconocer si dos vectores son ortogonales (forman un "ángulo recto"),i.e., dos vectores son ortogonales si su producto escalar es cero (por ejemplo, $(1,3)$ y $(-6,2)$).

También veremos como la noción de distancia no permitirá resolver sistemas de ecuaciones inconsistentes (que no tienen solución) utilizando la aproximación por mínimos cuadrados y su aplicación para crear modelos de regresión lineal.

9.1.1. Conceptos Básicos

Definición 9.1 (Producto Escalar:) *Denominamos* producto escalar*, o producto interno, en un espacio vectorial V sobre $\mathbb{K}$, a una función que asocia a cada par de vectores $u,v \in V$ un número escalar $\alpha \in \mathbb{K}$ que cumple las siguientes propiedades:*

1. *Conmutativa:* $u \cdot v = v \cdot u$ $\forall u,v \in V$
2. *Distributiva:* $u \cdot (v+w) = u \cdot v + u \cdot w$ $\forall u,v,w \in V$
3. *Reubicación el escalar:* $\alpha(u \cdot v) = (\alpha u) \cdot v = u \cdot (\alpha v)$ $\forall \alpha \in \mathbb{K}, u,v \in V$
4. *Definida positiva:* $v \cdot v \geq 0$*, y* $v \cdot v = 0$ *si solo si* $v = 0$*)* $\forall v \in V$

también se puede utilizar la notación $\langle u,v \rangle$.

Cualquier operación que se pueda expresar como $V \times V \to \mathbb{K}$, y que cumpla las propiedades definidas en Def. 9.1 se denomina producto escalar. Por ejemplo:

Definición 9.2 (Producto escalar usual en R^n) *El producto escalar usual en R^n se define como:*

$$(a_1,a_2,\ldots,a_n) \cdot (b_1,b_2,\ldots,b_n) = a_1b_1 + a_2b_s + \ldots a_nb_n$$

y puede verse como el producto de una matriz fila por una matriz columna:

$$\begin{pmatrix} a_1 & a_2 & \ldots & a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{pmatrix} = a_1b_1 + a_2b_s + \ldots a_nb_n$$

Definición 9.3 *También podemos "inventar" una operación en R^3 que cumpla las propiedades de la Def. 9.1 y por tanto podemos considerarla un producto escalar:*

$$(a_1,a_2,a_3) \cdot (b_1,b_2,b_3) = a_1b_1 + 2a_2b_s + 3a_3b_3$$

Definición 9.4 (Producto escalar en $\mathcal{M}_2$) *Como el producto ordinario de matrices no es un producto escalar (su resultado no es un escalar, no es conmutativo, etc.), podemos definir el siguiente producto escalar:*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = aa' + bb' + cc' + dd'$$

9.2. Espacios prehilbertianos y euclídeo

Definición 9.5 (Espacio vectorial prehilbertianos:) *Un espacio vectorial provisto de un producto escalar es un espacio prehilbertiano. Por lo tanto, un espacio prehilbertiano es un tipo de espacio métrico, con la métrica inducida por una* norma *definida a partir del producto escalar:*

$$||v|| = \sqrt{v \cdot v} \tag{9.1}$$

Para que un espacio prehilbertiano sea un espacion de Hilbert (hilbertiano) tiene que ser un espacio completo, es decir, es condición necesaria que el cuerpo base $\mathbb{K}$ sea R o C. Por lo tanto, ningún espacio prehilbertiano sobre Q, puede ser un espacio de Hilbert.

Definición 9.6 (Espacio Euclídeo:) *Decimos que un espacio vectorial es euclídeo si es un espacio de Hilbert de dimensión finita.*

En espacios euclídeos la norma $||v||$, define la "longitud" del vector v.

9.2.1. Conceptos geométricos en un espacio euclídeo

Al igual que ocurre en el plano o el espacio con el producto escalar usual, en general, dado un espacio euclídeo si nos referimos a un cierto producto escalar obtenemos los siguientes conceptos geométricos:

Definición 9.7 (Módulo de un vector:) *El módulo de un vector $|v|$, que corresponde, intuitivamente, a la "longitud" del vector, se define utilizando la norma del espacio vectorial:*

$$|v| = ||v|| = \sqrt{v \cdot v}$$

Al estar definida a partir del producto escalar:

- *El único vector de modulo cero es el vector 0.*

- *El módulo de un vector es el mismo que el de su opuesto,* $|v| = |-v|$.
- *El módulo de* αv *es* $|\alpha v| = |\alpha||v|$ *(es decir, el módulo queda multiplicado por el valor absoluto del escalar).*
- *Se cumple la desigualdad triangular* $|u+v| \leq |u|+|v|$ *para cualquier* u, v.

Definición 9.8 (Ángulo entre dos vectores:) *A partir del producto escalar usual de* R^2 *tenemos que* $u \cdot v = |u||v|cos\theta$, *donde* $\theta \in [o, \pi]$ *es el ángulo que forman los vectores u y v. Por lo tanto, generalizamos la noción de ángulo a cualquier espacio euclídeo, definiendo el ángulo entre u y v como el único número* $\theta \in [o, \pi]$ *tal que:*

$$cos\theta = \frac{u \cdot v}{|u||v|}$$

Definición 9.9 (Vectores ortogonales:) *Dos vectores u, v son ortogonales si su producto escalar es cero:* $u \cdot v = 0$ *y de denota* $u \perp v$.

Diremos que un conjunto de vectores es un conjunto ortogonal si cada uno de ellos es ortogonal a todos los demás. (Exigimos además que ninguno de los vectores sea el 0!).

Notar que si dos vectores u, v *son ortogonales entonces también lo son sus múltiplos* αu *y* βv *(*$\forall \alpha, \beta$ *escalares).*

Teorema 9.1 (Teorema de Pitágoras) *Dos vectores* u, v *son ortogonales si y sólo si* $||u+v||^2 = ||u||^2 + ||v||^2$.

Proof 9.1 *Aplicando la definición de norma:*

$$||u+v||^2 = (u+v) \cdot (u+v) = u \cdot u + u \cdot v + v \cdot u + v \cdot v = ||u||^2 + ||v||^2 + 2u \cdot v$$

Definición 9.10 (Distancia entre dos vectores:) *La distancia entre u y v es la norma del vector diferencia entre ambos:*

$$dis(u, v) = ||u - v||$$

9.3. Aproximación por mínimos cuadrados

En temas anteriores hemos estudiado el problema de resolver sistemas lineales de la forma $Ax = b$. Si lo analizamos en términos de "distancia", cuando intentamos resolver tales sistemas, lo que buscamos es el conjunto de vectores x que hacen que la distancia entre Ax y b sea igual a 0:

$$||Ax - b|| = 0$$

equivalentemente para evitar raíces cuadradas, escribimos $||Ax-b||^2 = 0$. Si el conjunto de soluciones de $Ax = b$ es vacío, con frecuencia estamos interesados en encontrar un vector x_0 que haga $||Ax_0 - b||^2$ tan pequeño como sea posible (en una cierta norma). Tal vector lo denominaremos solución de mínimos cuadrados del sistema $Ax = b$. Este término se deriva del hecho de que $||Ax - b||$ es la raíz cuadrada de una suma de términos al cuadrado.

Definición 9.11 (Solución de mínimos cuadrados:) *Dados una matriz A de dimensión $m \times n$ y un vector $b \in R^m$, un vector $x_0 \in R^n$ se denomina* solución de mínimos cuadrados *de $Ax = b$ si y sólo si para todo $x \in R^n$:*

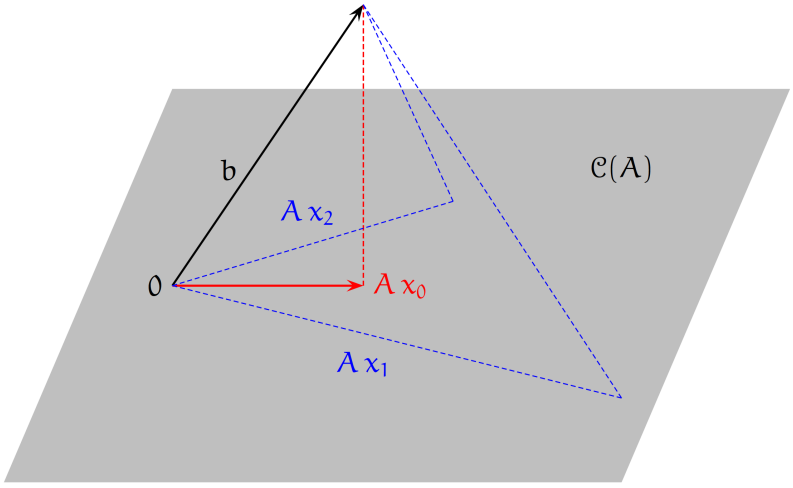
$$||Ax_0 - b||^2 \leq ||Ax - b||^2 \tag{9.2}$$

Obsérvese que, si x_0 resuelve el sistema $Ax = b$, entonces es una solución de mínimos cuadrados.

El método de mínimos cuadrados hace mínimo el error cuadrático.

9.3.1. Interpretación geométrica

La solución de mínimos cuadrados x_0 del sistema $Ax = b$ satisfará que el vector Ax_0 pertenece al espacio columna de A, $\mathcal{C}(A)$ y además será el vector que haga que la distancia de Ax_0 al vector b sea mínima.



Obviamente, la proyección ortogonal de b sobre el espacio columna de A proporcionará Ax_0. Supongamos que x_0 satisface $Ax_0 = P_{\mathcal{C}(A)}(b)$. Tal proyección cumple que

$b - P_{\mathcal{C}(A)}(b)$ es ortogonal a $\mathcal{C}(A)$ y así $b - Ax_0$ es ortogonal a cada columna A_j de A:

$$A_j^t(b - Ax_0) = 0, \quad j = 1, \ldots, m$$

Estas ecuaciones pueden ser escritas en una única expresión matricial de la forma:

$$A^t(b - Ax_0) = 0 \quad \rightarrow A^t Ax_0 = A^t b$$

Esto prueba el siguiente teorema:

Teorema 9.2 *El conjunto de soluciones de mínimos cuadrados de $Ax = b$ coincide con el conjunto de soluciones* no vacío *del sistema:*

$$A^t Ax = A^t b \tag{9.3}$$

9.3.2. Resolución del sistema de ecuaciones normales para x_0

El sistema escrito en forma matricial $A^t Ax = A^t b$ es denominado sistema de *ecuaciones normales para* x_0. Para encontrar al solución de mínimos cuadrados de $Ax = b$, el primer paso consiste en multiplicar por al izquierda ambos miembros por A^t.

- Si la matriz $A^t A$ es invertible, el sistema $A^t Ax = A^t b$ tiene solución única dada por:
$$x_0 = (A^t A)^{-1} A^t b \tag{9.4}$$
- Si la matriz $A^t A$ no es invertible, entonces el sistema $A^t Ax = A^t b$ tendrá infinitas soluciones (que encontraremos, por ejemplo, utilizando el método de Gauss).

Ejemplo 9.1 *Dado el sistema* $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$*, que es incompatible porque* $b = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ *no pertenece al subespacio S generado por* $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ *en* R^2.

Dado que $A^t A$ no es invertible obtendremos infinitas soluciones aplicando el método de Gauss:

Primero obtenemos la mejor aproximación de b en S. Para ello, utilizamos el vector (1,2) como base de S:

$$b' = proy_{(1,2)}(b) = \frac{(1,2)\begin{pmatrix} 3 \\ 5 \end{pmatrix}}{(1,2)\begin{pmatrix} 1 \\ 2 \end{pmatrix}}(1,2) = \frac{13}{5}(1,2) = \left(\frac{13}{5}, \frac{26}{5}\right)$$

y a continuación resolvemos el sistema (indeterminado) $\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}\frac{13}{5}\\ \frac{26}{5}\end{pmatrix}$ *cuyas infinitas soluciones son* $\begin{cases}x &= \frac{13}{5}- & 2\lambda\\ y &= & \lambda\end{cases}$

El error cuadrático es:

$$||b'-b||^2=\left|\left|\left(\frac{13}{5},\frac{26}{5}\right)-(3,5)\right|\right|^2=\frac{1}{5}=0,2$$

Ejemplo 9.2 *Dado el sistema* $\begin{pmatrix}2 & 3\\ 1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}0\\ 1\\ 1\end{pmatrix}$*, que es incompatible porque* $b=\begin{pmatrix}0\\ 1\\ 1\end{pmatrix}$ *no pertenece al subespacio S generado por* $\begin{pmatrix}2\\ 1\\ 0\end{pmatrix}$, $\begin{pmatrix}3\\ 0\\ 1\end{pmatrix}$ *en* R^3.

Dado que A^tA *es invertible podemos obtener la solución por mínimos cuadrados directamente,* $x_0=(A^tA)^{-1}A^tb$*:*

$$x_0=\left(\begin{pmatrix}2 & 1 & 0\\ 3 & 0 & 1\end{pmatrix}\begin{pmatrix}2 & 3\\ 1 & 0\\ 0 & 1\end{pmatrix}\right)^{-1}\begin{pmatrix}2 & 1 & 0\\ 3 & 0 & 1\end{pmatrix}\begin{pmatrix}0\\ 1\\ 1\end{pmatrix}=\begin{pmatrix}5 & 6\\ 6 & 10\end{pmatrix}^{-1}\begin{pmatrix}1\\ 1\end{pmatrix}=$$

$$=\begin{pmatrix}\frac{5}{7} & \frac{-3}{7}\\ \frac{-3}{7} & \frac{5}{14}\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix}=\begin{pmatrix}\frac{2}{7}\\ \frac{-1}{14}\end{pmatrix}$$

Por lo tanto la solución es $\begin{cases}x &= \frac{2}{7}\\ y &= \frac{-1}{14}\end{cases}$

El error cuadrático es:

$$||Ax_0-b||^2=\left|\left|\left(\frac{5}{14},\frac{2}{7},\frac{-1}{14}\right)-(0,1,1)\right|\right|^2=\frac{25}{14}\approx 1,79$$

9.3.3. Aplicación: análisis de regresión

A continuación aplicaremos el método de aproximación por mínimos cuadrados al ajuste de una nube de puntos (un conjunto de puntos en el plano).

Este ejemplo puede surgir al obtener resultados experimentales, estadísticos, etc, donde buscamos la recta que mejor se ajuste a dichos resultados.

Si planteamos que la recta pase por todos los puntos, obtendremos un sistema incompatible que tendremos que resolver aplicando el método de mínimos cuadrados.

Ejemplo 9.3 *Dados los puntos* $(1,2)$, $(2,3)$, $(3,5)$, *encontrar la recta que mejor se ajuste.*

Dada una recta con la ecuación $y = mx + b$, *si pasara por los tres puntos, debería cumplir que:*

$$\begin{cases} 2 &= m \cdot 1 + b \\ 3 &= m \cdot 2 + b \\ 5 &= m \cdot 3 + b \end{cases}$$

Si lo ponemos en forma de sistema $Ax = b$, *con* m, b *como incógnitas tenemos:*

$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$$

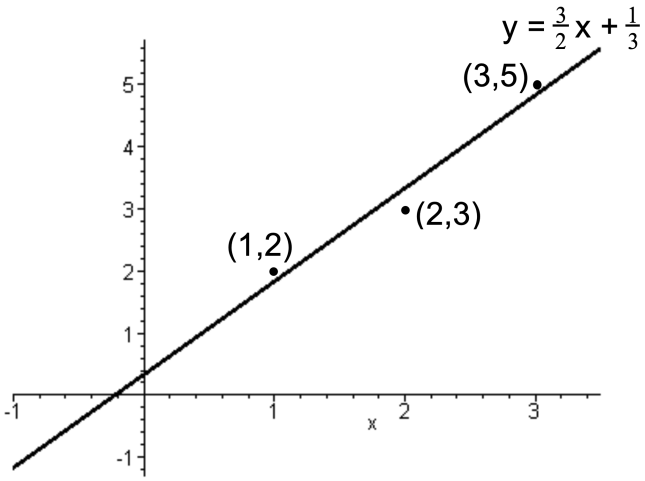
Como las columnas de A *son linealmente independientes, existe inversa para* A^tA *y por lo tanto hay solución única que se puede hallar directamente como* $x_0 = (A^tA)^{-1}A^tb$:

$$x_0 = \left(\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} =$$

$$= \begin{pmatrix} 14 & 6 \\ 6 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 23 \\ 10 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -1 \\ -1 & \frac{7}{3} \end{pmatrix} \begin{pmatrix} 23 \\ 10 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{1}{3} \end{pmatrix}$$

Es decir la recta que estamos buscando es $y = \frac{3}{2}x + \frac{1}{3}$.

Se puede comprobar en la figura siguiente que, si bien la recta no pasa por ninguno de los puntos dados, se aproxima a todos ellos.



El error cuadrático es:

$$||Ax_0 - b||^2 = \left|\left|\left(\frac{11}{6}, \frac{10}{3}, \frac{29}{6}\right) - (2,3,5)\right|\right|^2 = \frac{1}{6} \approx 0,167$$

También podríamos ajustar la nube de puntos mediante una parábola $y = ax^2 + bx + c$, o mediante polinomios de grado mayor; o también mediante una función de cualquier tipo como por ejemplo $y = acos(x) + blog(x)$, etc.

En todos los casos se debe "hacer pasar" la función por cada uno de los puntos, introduciendo la abscisa del punto en la x y la ordenada en la y, para obtener una ecuación por cada punto. Finalmente obtendremos un sistema incompatible en las incógnitas a, b, c... que resolveremos por mínimos cuadrados.

Como ya hemos comentado, puede ocurrir que el método de mínimos cuadrados se aplique a un sistema que ya era compatible. Es lo que ocurre si, por ejemplo, tratamos de ajustar por una recta una nube de puntos que ya estaban alineados.

En este caso no hay ningún impedimento para aplicar el método, y la solución "aproximada" es en realidad la solución "exacta", la misma que se obtendría resolviendo el sistema compatible por cualquiera de los métodos conocidos.

En este caso, el error cuadrático sería cero $||Ax_0 - b||^2 = 0$.

9.3.4. Primera aplicación en la historia

Mientras observaba una región en la constelación de Tauro el 1 de enero de 1801, Giuseppe Piazzi, astrónomo y director del observatorio de Palermo observó una pequeña "estrella" que nunca había visto antes. Mientras Piazzi y otros continuaban observando esta nueva "estrella" que en realidad era un asteroide. Se dieron cuenta de que se

había descubierto un nuevo "planeta", sin embargo, su nuevo "planeta" desapareció completamente en el otoño de 1801.

Los astrónomos conocidos de la época se unieron a la búsqueda de reubicar el "planeta perdido" pero todos los esfuerzos fueron en vano.

En septiembre de 1801 Carl F. Gauss decidió enfrentarse al desafío de encontrar este "planeta perdido". Gauss permitió la posibilidad de una órbita elíptica en lugar de obligarla a ser circular que era una suposición de los otros y procedió para desarrollar el método de mínimos cuadrados. En diciembre, la tarea se completó y Gauss informó a la comunidad científica no solo donde se encontraba el planeta perdido, sino que también predijo su posición en los tiempos futuros. Observaron y era exactamente donde Gauss había predicho que sería.

El asteroide se llamó Ceres y la contribución de Gauss fue reconocida por el nombramiento de otro asteroide menor Gaussia. Esta extraordinaria hazaña de localizar un diminuto y lejano cuerpo celeste a partir de datos aparentemente insuficientes asombró a la comunidad científica. Además, Gauss se negó a revelar sus métodos y hubo quienes incluso lo acosaron de brujería.

Estos acontecimientos condujeron directamente a la fama de Gauss en toda la Comunidad Europea y contribuyeron a establecer su reputación como un genio matemático y científico del más alto nivel.

Universidad Rey Juan Carlos

Tema 10

Códigos Lineales

10.1. Teoría de Códigos

Los códigos lineales son parte de la teoría de Códigos desarrollada a partir de los trabajos de Richar Hamming y Marcel Golay en 1950 en los que construyeron, códigos capaces de detectar y correguir un número especifico de errores producidos durante la transmisión de un determinado mensaje.

La transmisión de un mensaje se realiza mediante un sistema de comunicación que consta de cinco partes: el emisor, un codificador, un canal de comunicación, un descodificador y un receptor (ver Fig. 10.1).

La comunicación de un mensaje consiste en la transmisión de una secuencia de caracteres desde un emisor a un receptor a través de un canal de comunicación. En nuestro modelo, imperfecciones del canal, denominadas ruido, provoca que algunos caracteres sean incorrectamente recibidos por el receptor.

Si consideramos que el mensaje se codifica en un alfabeto digital $D^n = \{0,1\}^n$, al que llamaremos **Código**, esto implica que algún *bit* se recibe mal. Por lo tanto introducire-

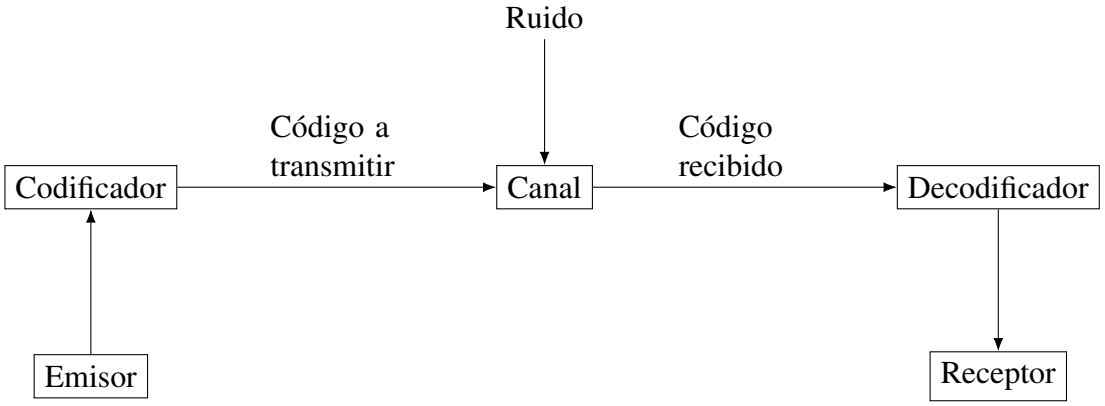


Figura 10.1: Modelo sistema de comunicación.

mos, de modo sistemático, redundancia en la información de modo que al trasmitir un determinado número de bits con una capacidad de corrección determinada, el número de bits añadidos sea mínimo.

Ejemplo 10.1 (Contando cartas) *Consideremos que Edward Thorp quiere contar cartas en un casino y enviarle el resultado a Andy Bloch utilizando un canal de transmisión cuya probabilidad de transmitir incorrectamente un bit es del uno por mil. Para contar las cartas necesitan un alfabeto de 4 símbolos correspondiente al palo de las cartas,* {pica, corazón, diamante, trébol} *y cada mensaje será una sucesión de cartas.*

Supongamos que codifican las cartas como $D^2_{cartas} = \{00,01,10,11\}$ *y las primeras cuatro cartas son dos picas y dos corazones el código a transmitir por Edward sería* 00000101*. Si por efecto del ruido el código recibido es* 00010111*, entonces Andy interpretaría que Edward ha visto una pica, dos corazones y un trébol. Con este código, la probabilidad de recibir una palabra (de dos bits) correctamente es del* $99,8001\,\%$*, por lo tanto si contamos 100 cartas la probabilidad de recibir el mensaje correctamente es del* $81,865\,\%$*.*

Supongamos que codifican las cartas como $D^3_{cartas} = \{000,011,101,110\}$*, el código a transmitir por Edward sería* 000000011011*. Si Andy recibe* 000010011111*, detectaría que la segunda carta,* 010*, y la cuarta carta,* 111*, son incorrectas y podría pedirle a Andy que vuelva a retransmitir el mensaje. Ahora la probabilidad de recibir una palabra (de tres bits) correctamente es del* $99,7003\,\%$*, por lo tanto para 100 cartas la probabilidad de recibir el mensaje correctamente baja al* $74,071\,\%$*, pero en un* $25,907\,\%$ *de los casos se detectará que el mensaje es incorrecto y la probabilidad de interpretar mal la recepción del mensaje pasa del* $18,135\,\%$ *a menos del* $0,025\,\%$*. Sin embargo, Andy no puede pedirle a Edward que repita el envío cuando hay un error.*

Por lo tanto deciden codificar las cartas como $D^5_{cartas} = \{00000,01101,10110,11011\}$*, y el código a transmitir por Edward sería* 0000000000110101101*. Ahora, si Andy recibiese* 0000001000011011101*, no solo detectaría que la segunda carta,* 01000*, y la cuarta carta,* 11101*, son incorrectas,*

sino que sería capaz de substituirlas por 00000 *y* 01101 *respectivamente. Con este código, la probabilidad de que el mensaje con 100 cartas sea correctamente interpretado aumenta hasta el* $99,9\,\%$.

10.1.1. Detección y corrección de errores

Para detectar y/o corregir errores necesitamos definir la noción de distancia entre palabras.

Definición 10.1 (Distancia Hamming) *Sean* $a = (a_1, \ldots, a_n)$ *y* $b = (b_1, \ldots, b_n)$ *dos palabras de* D^n*, la distancia Hamming entre a y b, denotada por* $d(a,b)$*, es el número de componentes no iguales:*

$$d(a,b) = |\{j : a \leq j \leq n, a_j \neq b_j\}|$$

donde $|S|$ *denota el cardinal de un conjunto S (número de elementos que contiene).*

A partir de la distancia Harmming definimos el tipo de error producido durante la transmisión del código.

Definición 10.2 (Error tipo λ) *Sean c y r dos palabras de* D^n*, tal que c es el código a transmitir por un canal con ruido y r es el código recibido. Diremos que se ha producido un error de tipo* λ *durante la transmisión, si* $d(c,r) = \lambda$.

Ejemplo 10.2 (Tipo de error) *En el ejemplo 10.1 los errores de transmisión de las cartas que se recibieron de manera incorrecta fueron de tipo 1, p.ej.,* $d(01101, 11101) = 1$.

Ya hemos visto dos parámetros importantes a la hora de definir un código: *n* la longitud de las palabras, y *M* el número de palabras. Sin embargo nos queda por definir el más importante a la hora de poder determinar el número de errores que un código puede detectar o corregir.

Definición 10.3 (Distancia mínima de un Código) *Sea C un Código, la distancia mínima de C, denotada por* $d(C)$*, es la menor de las distancias entre dos palabras cualesquiera de C:*

$$d(C) = min\{d(c,c') : c, c' \in C, c \neq c'\}$$

Ejemplo 10.3 (Distancia mínima) *En el ejemplo 10.1 la distancia mínima de D^2_{cartas} es 1, de D^3_{cartas} es 2 y de D^5_{cartas} es 3.*

Definición 10.4 (Código λ-detector) *Un Código C detecta λ errores si verifica que para cada palabra $c \in C$ y cada palabra r obtenida a partir de c cambiando entre 1 y λ bits, la palabra r no es una palabra del C.*

En otras palabras, si utilizamos un código λ-detector y al transmitir una palabra se produce un error de tipo λ o menos, tenemos la seguridad de que la palabra recibida no será una palabra del código y del descodificador la detectará como errónea (pudiendo solicitar una retransmisión del mensaje).

Ejemplo 10.4 (λ-detector) *Un código es λ-detector si, y sólo si, su distancia mínima es mayor que λ, i.e., $\lambda < d(C)$. En el ejemplo 10.1 el código D^2_{cartas} no detecta errores, el código D^3_{cartas} detecta errores tipo 1, y el código D^5_{cartas} detecta errores tipo 1 y 2.*

Definición 10.5 (Código λ-corrector) *Un Código C corrige λ errores si verifica que toda palabra recibida con a lo sumo λ errores es descodificada como la palabra transmitida. Utilizando al decodificación por distancia mínima significa que cada palabra $c \in C$ y cada palabra r obtenida modificando a lo sumo λ bits de c, la distancia Hamming entre c y r es estrictamente menor que al distancia entre r y cualquier otra palabra de C distinta de la palabra c.*

Ejemplo 10.5 (λ-corrector) *Un código es λ-corrector si, y sólo si, su distancia mínima es mayor que 2λ, i.e., $2\lambda < d(C)$. En el ejemplo 10.1 los códigos D^2_{cartas} y D^3_{cartas} no corrigen errores y el código D^5_{cartas} corrige 1 error (es un código 1-corrector).*

10.2. Códigos lineales

A continuación estudiaremos los códigos lineales que gracias a su estructura de espacio vectorial son más fáciles de implementar. En general un código lineal se puede definir sobre un alfabeto Z_p con p un número primo, sin embargo en este tema solo estudiaremos códigos lineales construidos a partir de un alfabeto binario $Z_2 = \{0,1\}$ que con las siguientes operaciones de suma '$+$' y multiplicación '$\cdot$':

$$0+0=0,\ \ 0+1=1,\ \ 1+0=1,\ \ 1+1=0$$

$$0\cdot 0=0,\ \ 0\cdot 1=0,\ \ 1\cdot 0=0,\ \ 1\cdot 1=1$$

es un Cuerpo, que denotamos como $F_2 = (Z_2, +, \cdot)$.

Proposición 10.1 (F_2^n es espacio vectorial) *Sea F_2 un Cuerpo, el conjunto F_2^n de palabras de longitud n en F_2 en un espacio vectorial en el que la suma de vectores '+' y la multiplicación escalar '·' son:*

$$\begin{aligned}(x_1\ x_2\ \cdots\ x_n)+(y_1\ y_2\cdots\ y_n) &= (x_1+y_1\ x_2+y_2\ \ldots\ x_n+y_n)\\ 0\cdot(x_1\ x_2\ \cdots\ x_n) &= (0\ 0\ \ldots\ 0)\\ 1\cdot(x_1\ x_2\ \cdots\ x_n) &= (x_1\ x_2\ \ldots\ x_n)\end{aligned}$$

A partir de F_2^n podemos definir un código lineal binario y sus dos parámetros principales (n,k).

Definición 10.6 (Código Lineal) *Cualquier subespacio vectorial C del espacio vectorial F_2^n es un (n,l) código lineal (binario) C, donde:*

- *n es la longitud de las palabras (dimensión del espacio vectorial F_2^n).*
- *k es la dimensión del subespacio vectorial C.*

Según la Definición 10.3 para calcular la distancia mínima de un código C debemos calcular la distancia entre todas las palabras del código, es decir, $(|C|^2-|C|)/2$ operaciones.

Definición 10.7 (Peso) *El peso $w(x)$ de una palabra $x \in F_2^n$ es el número de 1's en x. Equivale a la distancia de la palabra x con el vector cero del Código, i.e., $w(x) = d(x,\vec{0})$.*

El peso mínimo de un código lineal $C \subseteq F_2^n$, denotado por $w(C)$, es el menor de los pesos de las palabras código distintas de la palabra cero:

$$w(C) = min\{w(x) : x \in C, x \neq \vec{0}\}$$

Una de las propiedades más interesantes de los códigos lineales es que el peso mínimo de un código lineal coincide con la distancia mínima, i.e., $d(C) = w(C)$ y por lo tanto para obtener el parámetro d solo tenemos que realizar $|C|-1$ operaciones.

Teorema 10.1 *En un código lineal, la distancia mínima y el peso mínimo coinciden.*

Proof 10.1 *Dadas dos palabras cualesquiera x, y, el vector $x-y$ tiene tantas componentes no nulas como componentes diferentes tengan las palabras x e y entre sí. Por tanto, $d(x,y) = w(x,y)$.*

Sean c_1 y c_2 dos palabras código tales que $d(c_1,c_2) = d(C)$. Entonces $d(C) = d(c_1,c_2) = w(c_1-c_2) \geq w(C)$, pues c_1-c_2 es una palabra código no nulo. Recíprocamente, sea c una palabra código tal que $w(c) = w(C)$, entonces $w(C) = w(c) = d(c,0) \geq d(C)$, pues la palabra $\vec{0}$ es una palabra código distinta del vector c. □

Ejemplo 10.6 (Parámetros (n,k)) *En el ejemplo 10.1 los tres códigos son (n,k) códigos lineales binarios: D^2_{cartas} es un $(2,2)$ código lineal, D^3_{cartas} es un $(3,2)$ código lineal, y D^5_{cartas} es un $(5,2)$ código lineal.*

10.2.1. Construcción de Códigos Lineales

En esta sección vamos a definir formalmente una función de codificación que nos permita dividir un flujo original de bits (correspondiente a un mensaje de tamaño indeterminado) en bloques de tamaño finito k y asignarles una palabra codificada de longitud n.

Definición 10.8 (Función de codificación) *Una función de codificación es una aplicación lineal inyectiva, $f_C : F_2^k \to F_2^n$, de modo que a cada uno de los 2^k posibles bloques de F_2^k le asigna una palabra código en F_2^n.*

Ejemplo 10.7 (Función de codificación) *El ejemplo 10.1 es un caso particular donde el código D_{cartas^2} codifica el mensaje (la secuencia de cartas) en un flujo de bits y posteriormente, cada bloque de 2 bits (que se corresponde con la longitud de los símbolos del alfabeto en binario) se codifica asignando a cada bloque de 2 bits una palabra codificada de longitud $n = 3$, en el caso de D^3_{cartas}, y $n = 5$ en el caso de D^5_{cartas}.*

En general, dado que F_2^k y F_2^n son espacios vectoriales, candidatos posibles para la función de codificación son las transformaciones lineales, definidas por una matriz E.

Definición 10.9 (Matriz generadora de un código lineal) *Sea E, una matriz de n filas y k columnas de $\mathcal{M}_{n\times k}(Z_2)$ y sea C el (n,k) código lineal:*

$$C = \{x \in F_2^n : x^t = Ey^t, \forall y \in F_2^k\}$$

decimos que E es la matriz generadora de C, cuyas columnas son las palabra código correspondientes a los k vectores de la base canónica de F_2^k: $\{(10\ldots0), (01\ldots0), \ldots (00\ldots1)\}$.

Ejemplo 10.8 (Matrices generadoras) *Las matrices generadoras de los códigos D^3_{cartas} y D^5_{cartas}, del ejemplo 10.1 son:*

$$G_{D^3_{cartas}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \qquad G_{D^5_{cartas}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Definición 10.10 (Código sistemático y/o separable) *Un (n,k) código lineal C es un* código sistemático *si las k primeras filas de su matriz generadora E forman la matriz identidad I de orden k. Es decir, $G = \left(\frac{I}{Q}\right)$ siendo Q una matriz de $n-k$ filas y k columnas. En ese caso, se dice que la matriz generadora del código está en forma estándar o que es una matriz generadora estándar.*

Un (n,k) código lineal C es un código separable *si entre las filas de su matriz generadora están presentes las filas de I_k, la matriz identidad de orden k.*

En un código sistemático, al codificar el bloque $x = (x_1\ x_2\ \ldots\ x_k)$ de F_2^k la palabra código resultante $C(x) = (x_1\ x_2\ \ldots\ x_k\ x_{k+1}\ \ldots\ x_n)$ tiene dos partes diferenciadas:

- Los dígitos de **información**: Las primeras k entradas de la palabra codificada forman el bloque x, la información que se desea enviar.
- Los dígitos de **control**: Las $n-k$ entradas restantes forman la redundancia añadida por el código.

Ejemplo 10.9 (Códigos sistemáticos) *Los códigos D^3_{cartas} y D^5_{cartas} del ejemplo 10.1 son códigos sistemáticos.*

10.2.2. Detección de errores en códigos lineales

Dada la estructura lineal de un (n,k) código lineal C generado por una matriz G podemos definir una matriz, denominada matriz de control de paridad, que nos permite detectar si una palabra de longitud n pertenece (o no) al conjunto de palabras código de C.

Definición 10.11 (Matriz control de paridad) *Sea C un (n,k) código lineal. Una matriz $H \in \mathcal{M}_{(n-k)\times n}(Z_2)$ y rango $(n-k)$ es una matriz control de paridad para C si $Hx^t = \vec{0}$, para toda palabra código x de C.*

A partir de las ecuaciones definidas por $x^t = Gy^t$, del subespacio vectorial C, la matriz H es la matriz de coeficientes de un sistema de $(n-k)$ ecuaciones lineales con n incógnitas, homogéneas e independientes, que caracterizan C. Nota: un mismo código puede tener distintas matrices control de paridad, y una misma matriz puede ser matriz de control de paridad de varios códigos lineales.

Ejemplo 10.10 (Matriz H)

- *Si consideramos las ecuaciones del $(3,2)$ código lineal D^3_{cartas} del ejemplo 10.1, a partir de su matriz generadora, obtenemos el sistema correspondiente y aplicando transformaciones llegamos a obtener H:*

$$\begin{cases} x_1 = & y_1 & & \\ x_2 = & & y_2 & \\ x_3 = & y_1 + & y_2 & \end{cases} \rightarrow x_3 = x_1 + x_2 \rightarrow x_1 + x_2 + x_3 = 0 \rightarrow H = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$

y para toda palabra x de D^3_{cartas} se cumple que $Hx^t = \vec{0}$:

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

$$= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = (0)$$

- *Si consideramos el $(5,2)$ código lineal D^5_{cartas} del ejemplo 10.1 obtenemos:*

$$\begin{cases} x_1 = & y_1 & & \\ x_2 = & & y_2 & \\ x_3 = & y_1 + & y_2 & \\ x_4 = & y_1 & & \\ x_5 = & & y_2 & \end{cases} \rightarrow \begin{cases} x_3 = & x_1 + & x_2 & \\ x_4 = & x_1 & & \rightarrow \\ x_5 = & & x_2 & \end{cases}$$

$$\rightarrow \begin{cases} x_1 & + x_2 & + x_3 & & = 0 \\ x_1 & & & + x_4 & = 0 \\ & x_2 & & & + x_5 = 0 \end{cases} \rightarrow H = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

y para toda palabra x de D^5_{cartas} se cumple que $Hx^t = \vec{0}$:

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

Lema 10.1 *Si el $(n.k)$ código lineal C es un código sistemático, generado por una matriz $G = \left(\frac{I_k}{Q}\right)$, donde I es la matriz identidad de orden k y Q es una matriz de orden $(n-k) \times k$ entonces la matriz de paridad de C se puede definir como:*

$$H = \left(Q \;\middle|\; I_{(n-k)} \right)$$

donde $I_{(n-k)}$ es la matriz identidad de orden $(n-k)$, de paridad H de C es

Una vez tenemos la matriz de paridad estamos en disposición de verificar si una palabras es (o no es) una palabra código. Por la propia definición de la matriz control de paridad, una palabra x de F_2^n es una palabra código si, y sólo si, $Hx^t = \vec{0}$. En consecuencia, el proceso de detección de errores en los códigos lineales se simplifica.

Proposición 10.2 (Algoritmo detección de errores) *Sea H la matriz de paridad de un código C. Dado x un bloque recibido mediante un sistema de comunicación con ruido:*

- *si $Hx^t = \vec{0}$, la transmisión se considera **correcta**.*
- *si $Hx^t \neq \vec{0}$ se han producido **errores** en la transmisión*

Se elimina de esta forma la necesidad comparar la palabra recibida con cada una de las palabras código.

Ejemplo 10.11 (Detección de errores) *En el ejemplo 10.1 con el (3,2) código lineal D^3_{cartas} somos capaces de detectar que* 010 *y* 111 *son palabras incorrectas:*

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

10.2.3. Corrección de errores en códigos lineales

A partir de la matriz de paridad también podemos corregir errores en palabras recibidas con un error (donde el valor de uno de sus bits se ha modificado durante la transmisión).

Proposición 10.3 (Algoritmo corrección de errores) *Sea H la matriz de paridad de un código C. Dado que hay un error en uno de los bits de una palabra x recibida con respecto a la palabra código c enviada. El error ocurre en el i^{th} bit de x, donde i es el índice de la columna i^{th} de H cuyo vector es igual a Hx.*

Por lo tanto $c = x + e$, donde $e = (e_1 \ \ldots \ e_n)$ y los coeficientes e_j son nulos si $i \neq j$ y vale 1 si $i = j$.

Ejemplo 10.12 (Corrección de errores) *En el ejemplo 10.1 con el (5,2) código lineal D^5_{cartas} somos capaces de detectar que* 01000 *y* 11101 *son palabras incorrectas y podemos corregirlas:*

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

De modo que:

- *Para* $x=(01000)$ *el segundo bit es incorrecto (el vector* $(1\,0\,1)^t$ *es la* 2^a *columna de H) por lo tanto la palabra código enviada es* $(01000)+(01000)=(00000)$.
- *Para* $x=(11101)$ *es el primer bit el que es incorrecto (el vector* $(1\,1\,0)^t$ *es la* 1^a *columna de H) y la palabra código enviada es* $(11101)+(10000)=(01101)$.

¿Que errores (no) podemos corregir?

Hemos comentado anteriormente que la distancia mínima del código define el número de errores que se pueden detectar y/o corregir. Por lo tanto dado un código lineal de longitud n, no todas las 2^n posibles combinaciones de palabras podrán ser corregidas.

Ejemplo 10.13 *En el ejercicio 10.1 para el (5,2) código lineal* D^5_{cartas}*, ¿cuantas palabras de no pueden ser corregidas?*

- *La longitud de las palabras es* $n=5$ *por lo tanto tenemos* $2^5=32$ *posibles palabras de las que* 2^2 *son palabras código* $\{00000,01101,10110,11011\}$.
- *Para cada palabra código podemos detectar errores en cualquiera de sus* 5 *bits por lo tanto* $4\times 5=20$.
- *Por lo tanto de las* 32 *palabras,* 4 *son palabras código y* 20 *las podemos corregir:*

palabra códgio	*(00000)*	*(01101)*	*(10110)*	*(11011)*
error 1^{er} *bit*	*(10000)*	*(11101)*	*(00110)*	*(01011)*
error 2^o *bit*	*(01000)*	*(00101)*	*(11110)*	*(10011)*
error 3^{er} *bit*	*(00100)*	*(01001)*	*(10010)*	*(11111)*
error 4^o *bit*	*(00010)*	*(01111)*	*(10100)*	*(11001)*
error 5^o *bit*	*(00001)*	*(01100)*	*(10111)*	*(11010)*

- *De las* 32 *palabras posibles solo* 8 *no pueden ser corregidas (aquellas con una distancia mayor a 1 con respecto a las palabras código):* $\{(00011),(00111),(01010),(01110),(10001),(10101),(11000),(11100)\}$.

10.3. Códigos Lineales con Nombre Propio

Códigos de Golay: son cuatro códigos lineales sistemáticos, dos binarios y dos ternarios (sobre un alfabeto de tres símbolos), que publicó en 1949 en un trabajo de una página de extensión:

- El (24,12) código lineal binario G_{24} con $d(G_{24}) = 8$.
- El (23,12) código lineal binario G_{23} con $d(G_{23}) = 7$.
- El (12,6) código lineal ternario G_{12} con $d(G_{24}) = 6$.
- El (11,6) código lineal ternario G_{11} con $d(G_{11}) = 5$.

Se usó para transmitir fotos a color de Jupiter y Saturno desde los satélites Voyager en 1979-80.

Códigos Hamming: definidos independientemente por Marcel Golay (1949) y Richard Hamming (1950), son una familia de códigos lineales correctores de errores simples y con un sencillo algoritmo de descodificación para el alfabeto binario. Un código de Hamming binario $H(m,2)$ es un $(2^m - 1, 2^m - m - 1)$ con distancia 3, p.ej., para $m = 3$:

$$G_{H(3,2)} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad H_{H(3,2)} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Códigos Reed-Muller: introducidos en 1954, deben su nombre a sus descubridores, David E. Muller e Irving S. Reed. Es una colección de códigos binarios utilizado, por ejemplo, en la transmisión de fotografías del planeta Marte.

Códigos de Hadamard: son un caso particular de los códigos Reed-Muller. Un $(2^m, m)$ código lineal de Hadamard, denotado como H_m, con distancia mínima $d(H_m) = 2^{m-1}$, permite corregir $2^{m-2} - 1$ errores. La matriz generadora G_{H_m} es una matriz $\mathcal{M}_{2^m \times m}$ y se pueden construir fila a fila: la i^{th} fila es el entero i en representación binaria, p.ej., para $m = 3$:

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

Código Reed-Solomon: es un código cíclico no binario (sobre un alfabeto de 3 a 16 símbolos). El código fue inventado por Irving S. Reed y Gustave Solomon (de ahí su nombre) en el año 1960. La versión original se comprobó que en la práctica era ineficiente por lo tanto actualmente para corregir errores de transmisión se usa la Transformada Discreta de Fourier. Este código se encuentra actualmente aplicado en DVB (Digital Video Broadcasting), TDT, sitemas de almacenamiento (CD y DVD) en los Códigos QR.

Parte III

Prácticas

Universidad Rey Juan Carlos

Práctica A

Matemáticas Discreta

A.1. Algoritmos

Ejercicio 1: Dado el siguiente pseudocódigo

Algorithm 7:

Input: x y N.

```
1  X := x
2  i := 0
3  j := 0
4  while i ≤ N do
5  |   X := X + 2 × i
6  |   j := i
7  |   while j ≤ (N/2) do
8  |   |   X := X + j
9  |   |   j := j + 1
10 |   end
11 |   i := i + 1
12 end
13 return X
```

donde x es la suma de la primera cifra del DNI de cada integrante del grupo módulo 10 y N es la suma de las últimas dos cifras del DNI de cada integrante del grupo módulo

100.

a) Tradúcelo a R.
b) ¿Cuánto vale X al final del algoritmo?
c) ¿Cuántas instrucciones se realizan en total?

A.2. Combinatoria

Ejercicio 2: La empresa de ciberseguridad WebProt tiene 8 sedes en países al rededor del mundo: Madrid, Tokyo, París, Nueva York, Londres, Hong Kong, Singapur y Toronto. La empresa se comunica a través de una red cíclica (ver Figura A.1) de ordenadores en las que conectan un ordenador de cada sede. Sin embargo, para su buen funcionamiento, los ordenadores de los países del mismo continente deben estar colocados de forma adyacente. ¿Cuántas colocaciones distintas admite la red?

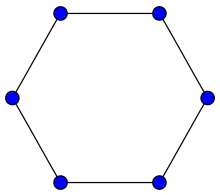


Figura A.1: Ejemplo de red cíclica

A.3. Teoría de Grafos

Ejercicio 3: En una actualización de la red, la empresa WebProt del ejercicio 2 quiere mejorar la conectividad de la red y establece una red con las siguientes conexiones:

- Madrid está conectado con París, Londres y Singapur.
- Tokyo está conectado con Hong Kong, Singapur y Toronto
- Nueva York está conectado con Toronto y con Londres

a) Calcula a mano la matriz de adyacencia y matriz de incidencia del grafo resultante de las conexiones de la empresa e introdúcelas en R.
b) Dibuja el grafo en R (usando el paquete *igraph*) a partir de la lista de aristas y la matriz de adyacencia.
c) Obtén con R los grados de cada vértice.

A.4. Aritmética modular

Ejercicio 4: Resuelve la siguiente congruencia lineal a través de un programa de R, mediante fuerza bruta (es decir, probando hasta encontrar la solución).

$$(D+3)\cdot x \equiv 13\cdot(D+7) \ \ mod\ (83)$$

siendo D la suma de la cuarta cifra del DNI de cada integrante del grupo módulo 10.

Universidad Rey Juan Carlos

Práctica B

Álgebra

B.1. Cálculo matricial

Ejercicio 1: Sea A la matriz que contiene en sus filas los DNIs de los integrantes del grupo (sin letra). Calcula a través de R las siguientes operaciones:

a) $C = 2 \cdot A \cdot A^t$

b) ¿Cuál es la suma de todos los elementos de C^2?

B.2. Espacios vectoriales y sistemas lineales

Ejercicio 2: Sea A la matriz y b el vector definidos en este enunciado , responde usando R las cuestiones enumeradas a continuación.

$$A = \begin{pmatrix} -1 & -2 & 3 \\ -1 & 3 & -1 \\ 2 & -5 & 5 \end{pmatrix}; b = \begin{pmatrix} 9 \\ -6 \\ 17 \end{pmatrix} \tag{B.1}$$

a) ¿Qué rango tiene la matriz ampliada $(A|b)$? ¿Son los vectores de sus filas independientes entre sí?
b) ¿Tiene solución el sistema lineal $A \cdot x = b$? ¿Cuál?

B.3. Diagonalización

Ejercicio 3: Sea la matriz A del ejercicio 2:

a) Calcula sus autovectores y autovalores con R.
b) ¿Es A diagonalizable? Si es así, comprueba con R que $A = P \cdot D \cdot P^{-1}$ siendo P y D las correspondientes matrices de autovectores y autovalores.

B.4. Aproximación por mínimos cuadrados

Ejercicio 4: Sean $DNI1$ y $DNI2$ el número de DNI de dos integrantes del grupo. Sean $DNI1_i$ y $DNI2_i$ las i-ésimas cifras de los correspondientes DNIs. Definimos la siguiente nube de 6 puntos:

$$(DNI1_1, DNI2_1), (DNI1_2, DNI2_2), ..., (DNI1_8, DNI2_8)$$

a) Encuentra a través de R la recta que mejor se ajusta por mínimos cuadrados.
b) Encuentra a través de R la parábola que mejor se ajusta por mínimos cuadrados.
c) Pinta con R los puntos, la recta ajustada y la parábola ajustada en el mismo gráfico.

B.5. Códigos lineales

Ejercicio 5: Al final del archivo `codigo_lineal.R` (ver Código B.1) está definida la variable `mensaje_recibido`. La información contenida en dicha variable corresponde a los bits de un mensaje cifrado en morse, codificado y enviado a través de un canal con ruido. Para ello se han usando las funciones definidas en el archivo: (i) El mensaje se ha traducido a código morse usando la función `string2morse(mensaje)`. El resultado es una sucesión de los cuatro símbolos del código morse:'/' representa la separación de palabras, '.' y '-' con el punto y raya, y ' ' separa los códigos que representan las letras del alfabeto latino (ver `https://morsedecoder.com/es/`); (ii) Después, con la función `morse2binary(mensaje)` se ha codificado en binario con un alfabeto de cuatro símbolos, $\{00, 01, 10, 11\}$ [1] y como el mensaje se transmite por un canal con

[1] Para facilitar la lectura los bits es†án ordenados en una matriz donde cada palabra es una columna.

ruido hemos decidido codificarlo con un (5,2) código lineal denominado `Codigo_52`:[2]

a) Calcula el numero de errores generados durante la transmisión del mensaje. Para ello se usará la función `detectar_errores(codigo,mensaje)`.

b) Descifra el mensaje recibido para obtener la frase en español, usando las funciones, `decodificar(mensaje)`, `binary2morse(mensaje)` y `morse2string(mensaje)`.

c) Corrige el mensaje recibido usando `corregir(codigo,mensaje)` y descifra el mensaje corregido.

d) Escribe la respuesta al mensaje.

Código B.1: Fichero codigo_lineal.R

```
### Código para la práctica 2 de la Asignatura de Matemáticas Discreta y
### Algebra del grado de ingeniería en Ciberseguridad del curso 2021-2022
### Autores: Marina Cuesta Santa Teresa y Joaquín Arias


options(max.print=1000000)  ## opción para imprimir sin limites

### Texto en Español a Código Morse ###

# Vector ordenado con todas las letras.
letras = c("␣","A","B","C","D","E","F","G","H","I","J","K","L","M","N",
           "O","P","Q","R","S","T","U","V","W","X","Y","Z","Ñ",
           "0","1","2","3","4","5","6","7","8","9",
           ".",",","?","'","!","/","(",")","&",":",
           ";","=","+","-","_","\"","$","@","¿","¡" )
# Vector ordenado con todos los códigos morse.
morses = c("/", ".-", "-...", "-.-.", "-..", ".", "..-.", "--.", "....", "..",
    ".---", "-.-", ".-..", "--", "-.",
           "---", ".--.", "--.-", ".-.", "...", "-", "..-", "...-", ".--",
               "-..-", "-.--", "--..", "--.--",
           "-----", ".----", "..---", "...--", "....-", ".....", "-....",
               "--...", "---..", "----.",
           ".-.-.-", "--..--", "..--..", ".----.", "-.-.--", "-..-.", "-.--.",
               "-.--.-", ".-...", "---...",
           "-.-.-.", "-...-", ".-.-.", "-....-", "..--.-", ".-..-.",
               "...-..-", ".--.-.", "..-.-", "--...-" )

# Recibe una letra "A" y devuelve su código morse ".-"
letra2morse<-function(letra) {
  return( morses[match(letra,letras)] )
}

# Recibe un código morse ".-" y devuelve su letra "A"
morse2letra<-function(morse) {
  return( letras[match(morse,morses)] )
```

[2]La función para codificar (resp. decodificar) palabras binarias en palabras código, es `codificar_mensaje(codigo,mensaje)` (resp. `decodificar_mensaje(codigo,mensaje)`).

```
}

# Recibe un mensaje y devuelve un vector con los códigos morse
string2morse<-function(string){
  STRING_=toupper(string) # convierte el mensaje a Mayusculas
  s_vector=matrix(strsplit(STRING_,"")[[1]],nrow=1,byrow=T)
  m_vector=sapply(s_vector,letra2morse,USE.NAMES = FALSE)
  m_string=paste(m_vector,collapse="␣")
  return(matrix(strsplit(m_string,"")[[1]],nrow=1,byrow=T))
}

# Recibe un vector con códigos morse y devuelve el mensaje
morse2string<-function(m_vector) {
  m_string=paste(m_vector,collapse="")
  morse_vector=matrix(strsplit(m_string,"␣")[[1]],nrow=1,byrow=T)
  vector=sapply(morse_vector,morse2letra,USE.NAMES = FALSE)
  return(paste(vector,collapse=""))
}

# Examples
# mensaje="Hola Mundo"
# string2morse(mensaje)
# morse2string(string2morse(mensaje))


### Codificar mensajes en morse usando un alfabeto binario ###

# Operadores auxiliares para manipular vectores binarios
mult_binaria <- function(A,B){ return((A%*%B)%%2) }
suma_binaria <- function(A,B){   return((A+B)%%2) }

# Recibe un código morse "." y devuelve su palabra binaria [01]
morse2bin<-function(simbolo){
  if(simbolo=="/"){ return(c(0,0)) }
  else if(simbolo=="."){ return(c(0,1)) }
  else if(simbolo=="-"){ return(c(1,0)) }
  else if(simbolo=="␣"){ return(c(1,1)) } }

# Recibe una palabra binaria [01] y devuelve su código morse "."
bin2morse<-function(palabra){
  if(all(palabra==c(0,0))){ return("/") }
  else if(all(palabra==c(0,1))){ return(".") }
  else if(all(palabra==c(1,0))){ return("-") }
  else if(all(palabra==c(1,1))){ return("␣") } }

# Recibe vector de códigos morse y devuelve matriz de palabras binarias
morse2binary<-function(mensaje_morse){
  return(sapply(mensaje_morse,morse2bin,USE.NAMES = FALSE))
}

# Recibe matriz de palabras binarias y devuelve vector de códigos morse
binary2morse<-function(mensaje_binario){
  return(matrix(apply(mensaje_binario,2,bin2morse),nrow=1,byrow = T))
}
```

```
85
86  # Ejemplos
87  # mensaje="Hola Mundo"
88  # string2morse(mensaje)
89  # morse2binary(string2morse(mensaje))
90  # binary2morse(morse2binary(string2morse(mensaje)))
91  # morse2string(binary2morse(morse2binary(string2morse(mensaje))))
92
93
94  ### Usar Códigos Lineales para enviar mensajes binarios ###
95
96  ## Definición de un (n,k) código lineal
97  # Matriz G - generadora de (5,2) código lineal con d=3
98  G_52=matrix(c(1,0,0,1,1,1,1,0,0,1),ncol=2,byrow = T)
99  # Matriz H - de paridad
100 H_52=matrix(c(1,1,1,0,0,1,0,0,1,0,0,1,0,0,1),nrow=3,byrow=T)
101 # Tabla para decodificar el (5,2) código lineal
102 Tabla_52<-function(codigo){
103   if (all(codigo==c(0,0,0,0,0))){ return(c(0,0)) }
104   else if (all(codigo==c(0,1,1,0,1))){ return(c(0,1)) }
105   else if (all(codigo==c(1,0,1,1,0))){ return(c(1,0)) }
106   else if (all(codigo==c(1,1,0,1,1))){ return(c(1,1)) }
107   else return(codigo[c(1,2)]) }
108 # Tabla para corregir códigos con 1 error
109 Tabla_Sindromes_52<-function(codigo, sindrome){
110   if     (all(sindrome==c(1,1,0))){ return(suma_binaria(codigo,c(1,0,0,0,0))) }
111   else if(all(sindrome==c(1,0,1))){ return(suma_binaria(codigo,c(0,1,0,0,0))) }
112   else if(all(sindrome==c(1,0,0))){ return(suma_binaria(codigo,c(0,0,1,0,0))) }
113   else if(all(sindrome==c(0,1,0))){ return(suma_binaria(codigo,c(0,0,0,1,0))) }
114   else if(all(sindrome==c(0,0,1))){ return(suma_binaria(codigo,c(0,0,0,0,1))) }
115   else { return(codigo) } }
116 # Definición del (5,2) código lineal Codigo_52
117 Codigo_52 <- list(G_52,H_52,Tabla_52,Tabla_Sindromes_52)
118
119
120 ## Funciones de codificación / decodificación para (n,k) código lineal
121
122 # Codificación
123 # Recibe matriz de palabras binarias y devuelve matriz con palabras código
124 codificar_mensaje<-function(codigo,mensaje){
125   G=codigo[[1]]
126   return(mult_binaria(G,mensaje))
127 }
128
129 # Decodificación
130 # Recibe matriz de palabras código y devuelve palabra binaria
131 decodificar_mensaje<-function(codigo,mensaje){
132   Tabla=codigo[[3]]
133   return(apply(mensaje,2,Tabla))
134 }
135
136 # Ejemplo
137 # mensaje="Hola Mundo"
138 # morse2binary(string2morse(mensaje))
```

```
# codificar_mensaje(Codigo_52,morse2binary(string2morse(mensaje)))
# decodificar_mensaje(Codigo_52, codificar_mensaje(Codigo_52,
    morse2binary(string2morse(mensaje))))


## Funciones para detectar / corregir errores

# Detectar errores
# Recibe matriz de palabras código (con errores) y devuelve el número de
    errores
detectar_errores<-function(codigo,mensaje){
  H=codigo[[2]]
  return(length(which(mult_binaria(H,mensaje)!=0)))
}

# Recibe matriz de palabras código (con errores) y devuelve matriz corregida
corregir_errores<-function(codigo,mensaje){
  H=codigo[[2]]
  Tabla_Sindromes=codigo[[4]]
  result = mensaje
  sindrome = mult_binaria(H, mensaje)
  for (c_index in 1:dim(mensaje)[2]){
    result[,c_index]=Tabla_Sindromes(mensaje[,c_index],sindrome[,c_index])
  }
  return(result)
}

# Ejemplo
# mensaje="Hola Mundo"
#
    detectar_errores(Codigo_52,codificar_mensaje(Codigo_52,morse2binary(string2morse(mensaje))))
#
    corregir_errores(Codigo_52,codificar_mensaje(Codigo_52,morse2binary(string2morse(mensaje))))


### Usar canal con ruido para enviar mensaje codificado ###

# Función que introduce interferencias en el canal de transmisión
ruido <- function(bit){ return(suma_binaria(bit, (sample.int(50,1)==1) ) ) }

# Canal de transmisión, recibe mensaje y devuelve mensaje con ruido
transmitir_mensaje <- function(mensaje_a_enviar){
  return(apply(mensaje_a_enviar,c(1,2) ,ruido))
}

# Ejemplo
# mensaje="Hola Mundo"
#
    mensaje_enviar=codificar_mensaje(Codigo_52,morse2binary(string2morse(mensaje)))
# mensaje_recibido=transmitir_mensaje(mensaje_enviar)
# detectar_errores(Codigo_52,mensaje_recibido)
#
# mensaje_enviar == mensaje_recibido
#
```

```
# mensaje_corregido=corregir_errores(Codigo_52,mensaje_recibido)
#
# mensaje_enviar == mensaje_corregido
#
# decodificar_mensaje(Codigo_52,mensaje_corregido)
# morse2string(binary2morse(decodificar_mensaje(Codigo_52,mensaje_corregido)))
# morse2string(binary2morse(decodificar_mensaje(Codigo_52,mensaje_recibido)))


## ------------------------------------------------------------------------ ##

## Función para generar el mensaje recibido
carga_bits_recibidos<-function(){
  data1=c(0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0,
      1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0,
      1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0,
      0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1,
      0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1,
      1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1,
      1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1,
      0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1,
      1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 1, 1, 0,
      0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0,
      0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0,
      1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0,
      1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1,
      1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1,
      1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1,
      0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1,
      0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1,
      1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0,
      1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0,
      0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1,
      1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1,
      1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1,
      1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0,
      1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1,
      0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0,
      1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0,
      1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1,
      0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0,
      0, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0,
      0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1,
      0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1,
      1, 0, 1, 1, 0, 1, 1)
  data2=c(0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0,
      1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0,
      1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0,
      1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0,
      1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1,
      0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1,
      0, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0,
      1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1,
      1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 0,
```

```
    1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1,
    1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1,
    1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1,
    0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0,
    1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1,
    0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0,
    1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1,
    0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1,
    1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0,
    1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0,
    1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0,
    0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0,
    1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0,
    1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0,
    1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1,
    1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1,
    1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1,
    0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1,
    0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1,
    1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1,
    1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 0,
    1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1,
    1, 0, 1, 0, 1, 1, 0)
data3=c(1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1,
    1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1,
    0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0,
    1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1,
    1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1,
    1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1,
    1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0,
    1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0,
    1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1,
    1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1,
    1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1,
    0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1,
    0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0,
    1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1,
    0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1,
    1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1,
    1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0,
    1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0,
    1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1,
    1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0,
    1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0,
    1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1,
    1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0,
    1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1,
    0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1,
    0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1,
    1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1,
    1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1,
    1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1,
    1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1,
    1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1,
```

```
        1, 0, 1, 1, 0, 1, 1)
204   data4=c(0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1,
        0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 0,
        1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1,
        0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1,
        0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1,
        0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1,
        1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1,
        0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1,
        1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1,
        1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1,
        1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0,
        1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1,
        1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0,
        1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1,
        1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0,
        1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1,
        1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1,
        0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1,
        1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1,
        0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0,
        1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1,
        1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1,
        0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1,
        1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0,
        1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 0,
        0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1,
        1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1)
205   data=c(data1, data2, data3, data4)
206   return(matrix(data, nrow = 5, byrow = F))
207 }
208 ## Variable con el mensaje recibido
209 mensaje_recibido = carga_bits_recibidos()
210
211
212 ### --- Práctica 2 --- Ejercicio 5 (Opcional) --- ----------------------- ###
213
214 # Apartado a)
215
216 # Apartado b)
217
218 # Apartado c)
219
220 # Apartado d)
```

Universidad Rey Juan Carlos

Práctica C

Soluciones

C.1. Soluciones Práctica Matemáticas Discreta

Solución Ejercicio 1:

a) Traducción a R:

```
algoritmo1 <- function(x,N) {
        X = x
        i = 0
        j = 0
        while (i <= N) {
                X = X+2*i
                j = i
                while(j <= (N/2)) {
                        X = X+j
                        j = j+1
                }
                i = i+1
        }
        return(X)
}
```

b) ¿Cuánto vale X al final del algoritmo?

```
print(paste("Para x=4 y N=67, X vale ",algoritmo(4,67)))
```

Al ejecutar `print` la salida por pantalla es:

"Para x=4 y N=67, X vale 17650".

c) ¿Cuántas instrucciones se realizan en total?

- Opción A: Calcular el número de instrucciones aplicando la fórmula.

- Opción B: Modificar el código del algoritmo para contar las instrucciones:

```
operaciones <- function(x,N) {
        op = 0
        X = x
        op = op + 1
        i = 0
        op = op + 1
        j = 0
        op = op + 1
        while (i <= N) {
                op = op + 1
                X = X+2*i
                op = op + 1
                j = i
                op = op + 1
                while(j <= (N/2)) {
                        op = op + 1
                        X = X+j
                        op = op + 1
                        j = j+1
                        op = op + 1
                }
                op = op + 1
                i = i+1
                op = op + 1
        }
        op = op + 1
        # return(X)
        op = op + 1
        return(op)
}

print(paste("El numero de operaciones para x=4 y N=67 es ",
    operaciones(4,67)))
```

La salida por pantalla es:

"El numero de operaciones para x=4 y N=67 es 2130".

Solución Ejercicio 2:

Número de combinaciones de continentes es una permutación cíclica: (3-1)!

Número de combinaciones dentro de Europa es una permutación: 3!

Número de combinaciones dentro de América es una permutación: 2!

Número de combinaciones dentro de Asia es una permutación: 3!

Cálculo del número de colocaciones posibles es:

$$comb_{total} = comb_{continentes} * comb_{europa} * comb_{america} * comb_{asia}$$

```
print(paste("El número de colocaciones posibles es ", factorial(3-1) *
    factorial(3) * factorial(2) * factorial(3)))
```

La salida por pantalla es:

"El número de colocaciones posibles es 144".

Solución Ejercicio 3:

a) Matriz de adyacencia:

$$\begin{array}{r} \\ Madrid \\ Tokyo \\ Pars \\ NuevaYork \\ Londres \\ HongKong \\ Singapur \\ Toronto \end{array} \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

y matriz de incidencia:

$$\begin{array}{r} \\ Madrid \\ Tokyo \\ Pars \\ NuevaYork \\ Londres \\ HongKong \\ Singapur \\ Toronto \end{array} \begin{array}{c} \begin{array}{cccccccc} a1 & a2 & a3 & a4 & a5 & a6 & a7 & a8 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix} \end{array}$$

b) Código en R:

```
# Llamo a la libreria
library(igraph)
# paises
paises=c("Madrid","Tokyo","Paris","Nueva York","Londres","Hong
    Kong","Singapur","Toronto")
```

```
## matriz de incidencia
matriz_incidencia=matrix(c(1,1,0,0,0,0,0,1,
                           0,0,1,1,1,0,0,0,
                           1,0,0,0,0,0,0,0,
                           0,0,0,0,0,1,1,0,
                           0,0,0,0,0,0,1,1,
                           0,0,0,1,0,0,0,0,
                           0,1,1,0,0,0,0,0,
                           0,0,0,0,1,1,0,0),byrow = T,ncol=8,nrow=8)
# añado nombres de filas
rownames(matriz_incidencia)=paises
# añado nombres de columnas
colnames(matriz_incidencia)=c("a1","a2","a3","a4","a5","a6","a7","a8")
## matriz de adyacencia
matriz_adyacencia=matrix(c(0,0,1,0,1,0,1,0,
                           0,0,0,0,0,1,1,1,
                           1,0,0,0,0,0,0,0,
                           0,0,0,0,1,0,0,1,
                           1,0,0,1,0,0,0,0,
                           0,1,0,0,0,0,0,0,
                           1,1,0,0,0,0,0,0,
                           0,1,0,1,0,0,0,0),byrow = T,ncol=8,nrow=8)
 # añado nombres de filas y columnas
colnames(matriz_adyacencia)=paises
rownames(matriz_adyacencia)=paises
matriz_adyacencia
## Grafo a partir de matriz de adyacencia
g1=graph_from_adjacency_matrix(matriz_adyacencia,mode="undirected")
plot(g1)
## lista de aristas
aristas=c("Madrid","Paris",
          "Madrid","Londres",
          "Madrid","Singapur",
          "Tokyo","Hong␣Kong",
          "Tokyo","Singapur",
          "Tokyo","Toronto",
          "Nueva␣York","Toronto",
          "Nueva␣York","Londres")
lista_aristas=matrix(aristas,ncol=2,byrow=T)
## Grafo a partir de lista de aristas
g2=graph_from_edgelist(lista_aristas, directed = F)
plot(g2)
```

c) Código en R:

```
deg <- degree(g2)
deg
```

Solución Ejercicio 4:

Código en R:

```
congruencia <- function(D) {
  valor_derecha=(13*(D+7))%%83
  x=0
  cond_parada=FALSE
  while (!cond_parada){
    valor_izquierda=((D+3)*x)%%83
    if (valor_izquierda==valor_derecha){
      cond_parada=TRUE
      return(x)
    }
    x=x+1
  }
}

print(paste("La solución (el valor de x) para D=3 es ", congruencia(3)))
```

La salida por pantalla es:

"La solución (el valor de x) para D=3 es 77".

C.2. Soluciones Práctica Álgebra

Solución Ejercicio 1:

a) Código en R:

```
A = matrix(c(5,2,...,8,9),nrow=3,byrow=T)
C = 2 * (A %*% t(A))
```

b) Código en R:

```
C_2 = C %*% C
suma = sum(C_2)
```

Solución Ejercicio 2:

a) Código en R:

```
A=matrix(c(1,-2,3,
           -1,3,-1,
           2, -5, 5), byrow=T,ncol=3)
b=c(9,-6,17)
# Junto la matriz A y el vector b para obtener la matriz A ampliada
A_ampliada=cbind(A,b)
```

```
# Rango de la matriz A ampliada
library(Matrix)
rango=rankMatrix(A_ampliada)
```

El rango de $(A|b)$ es 3.

Si, los vectores de sus filas son linealmente independientes entre si

b) ¿Tiene solución el sistema lineal asociado a la matriz A? ¿Cuál?

Si, tiene solución.

```
solve(A,b)
```

Obtenemos la siguiente solución:

```
[1]  1 -1  2
```

Por lo tanto $x_1 = 1$, $x_2 = -1$ y $x_3 = -2$.

Solución Ejercicio 3:

a) Código en R:

```
# Descomposicion de la matriz A
descomposicion = eigen(A)
# Autovalores
autovalores=descomposicion$values
# Matriz P de autovectores
P=descomposicion$vectors
```

b) Si, A es diagonalizable.

```
D = solve(P) %*% A %*% P
```

La matriz D es:

```
                [,1]          [,2]         [,3]
[1,]  7.721724e+00  8.881784e-16 1.054712e-15
[2,]  2.109424e-15  1.167336e+00 9.714451e-16
[3,] -4.843348e-15 -8.326673e-17 1.109405e-01
```

Solución Ejercicio 4:

a) Código en R:

```
# Dnis de ejemplo
DNI1=c(4,7,2,3,5,2,6,9)
DNI2=c(7,5,1,2,3,5,6,8)

## Aproximamos por minimos cuadrados
# Matriz y vector independiente del sistema de ecuaciones
A=cbind(DNI1,rep(1,8))
b=DNI2
# solucion aproximada
```

```
x_recta=(solve((t(A)%*%A)))%*%(t(A)%*%b)
# Obtenemos pendiente y ordenada en el origen de la recta
m=x_recta[1,1]
b=x_recta[2,1]
```

b) Código en R:

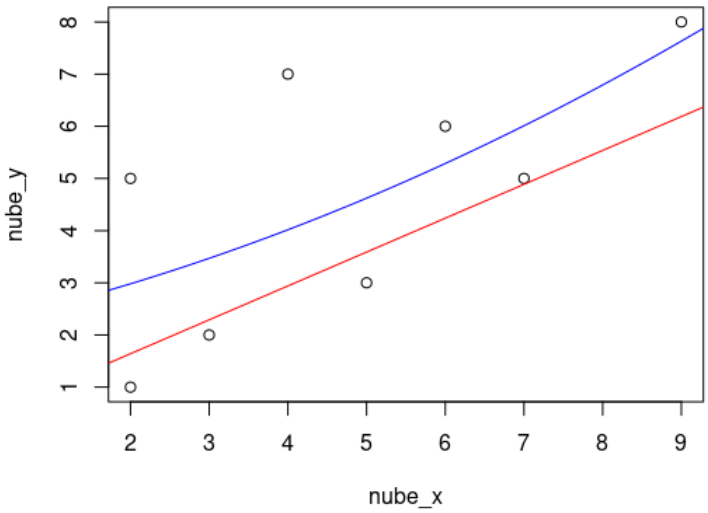
```
# Matriz y vector independiente del sistema de ecuaciones
A=cbind(DNI1**2,DNI1,rep(1,8))
b=DNI2
# solucion aproximada
x_parabola = (solve((t(A)%*%A)))%*%(t(A)%*%b)
# Obtenemos coeficientes a, b y c de la parabola
a=x_parabola[1,1]
b=x_parabola[2,1]
c=x_parabola[3,1]
```

c) Código en R:

```

# La nube de puntos que queremos pintar:
nube_x=DNI1
nube_y=DNI2
plot(nube_x,nube_y) #primero dibujamos la nube de puntos
# Dibujamos la recta:
x=seq(-100,100,0.1)
y_recta=m*x+b
lines(x,y_recta,col="red") # ahora agregamos la recta
# Dibujamos la parabola:
y_parabola=a*x**2+b*x+c
lines(x,y_parabola,col="blue") # ahora agregamos la parabola
```

El gráfico que se obtiene es el siguiente, siendo la línea roja la recta y la azul la parábola.



Solución Ejercicio 5:

a) Código en R:

```
errores = detectar_errores(codigo, mensaje_recibido)
```

b) Código en R:

```
mensaje = morse2string(biinary2morse(decodificar_mensaje(codigo,
    mensaje_recibido)))
```

c) Código en R:

```
mensaje_corregido = corregir(codigo,mensaje_recibido)
mensaje_original =
    morse2string(biinary2morse(decodificar_mensaje(codigo,
    mensaje_corregido)))
```

d) Sin resolver :-)