

MÉTODOS MATEMÁTICOS PARA LA INGENIERÍA MECÁNICA

NOTAS DE CLASE

Damián Ginestar Peiró
6 de septiembre de 2006

ESCUELA TÉCNICA SUPERIOR
DE INGENIERÍA DEL DISEÑO
DEPARTAMENTO DE MATEMÁTICA APLICADA
UNIVERSIDAD POLITÉCNICA DE VALENCIA

Índice general

1. Errores y aritmética finita	5
1.1. Introducción	5
1.2. Representación de los números en un ordenador	5
1.3. Propagación de errores	9
1.4. Ejercicios	10
2. Resolución de sistemas de ecuaciones	12
2.1. Introducción	12
2.2. Métodos directos para la resolución de sistemas	14
2.2.1. Factorización LU de una matriz	16
2.2.2. Pivotamiento	17
2.2.3. Errores y número de condición	20
2.2.4. Matrices especiales	21
2.3. Métodos iterativos	23
2.4. Ejercicios	25
3. Ecuaciones y sistemas de ecuaciones no lineales	28
3.1. Ecuaciones no lineales	28
3.1.1. Método de la bisección	28
3.1.2. Método del punto fijo	29
3.1.3. Método de Newton-Raphson	31
3.1.4. Método de la secante	32

3.2.	Resolución numérica de sistemas de ecuaciones	33
3.2.1.	Método iterativo del punto fijo	33
3.2.2.	Método de Newton-Raphson	34
3.2.3.	Método de Broyden	37
3.3.	Ejercicios	38
4.	Interpolación	41
4.1.	Introducción	41
4.1.1.	Varios problemas prácticos	42
4.2.	Interpolación lineal	44
4.3.	Polinomio interpolador de Lagrange	46
4.4.	Polinomio interpolador de Newton	50
4.5.	Caso particular de nodos igualmente espaciados	52
4.5.1.	Diferencias finitas	52
4.5.2.	Polinomio interpolador de Newton para puntos igualmente espaciados	53
4.6.	Interpolación inversa y extrapolación	54
4.6.1.	Interpolación inversa	54
4.6.2.	Extrapolación	54
4.7.	Splines	55
4.8.	Ejercicios	60
5.	Ajuste de curvas	63
5.1.	Recta de mínimos cuadrados	64
5.2.	Parábola de mínimos cuadrados	67
5.3.	Ajuste no lineal	68
5.4.	Ejercicios	70
6.	Introducción a las ecuaciones diferenciales en derivadas parciales	73
6.1.	Introducción	73

6.2.	Métodos de diferencias finitas	74
6.2.1.	Ecuación de Poisson 1D	75
6.2.2.	Ecuación de la difusión dependiente del tiempo	76
6.2.3.	La ecuación de ondas	77
6.2.4.	Ecuación de convección-difusión	78
6.2.5.	Ecuación de Poisson 2D	81
6.3.	Ejercicios	82

Capítulo 1

Errores y aritmética finita

1.1. Introducción

A la hora de realizar un determinado cálculo con un ordenador se pueden tener distintas fuentes de error:

- Errores de truncamiento.
- Error en los datos.
- Errores asociados a las aproximaciones numéricas utilizadas.

En este tema estudiaremos el origen de los errores de truncamiento derivados de utilizar una representación finita para los números y cómo se propagan los errores en los datos, que pueden deberse a los errores de truncamiento o a posibles errores experimentales, al realizar distintas operaciones aritméticas. Los errores derivados de utilizar una cierta aproximación numérica para realizar un cierto cálculo, los iremos estudiando a lo largo del curso.

1.2. Representación de los números en un ordenador

Por cuestiones de carácter técnico, los ordenadores utilizan una representación binaria de los números, donde los únicos dígitos posibles son el 0 o el 1, a diferencia de la representación decimal, o en base 10, que es la que utilizamos normalmente, donde se utilizan los dígitos del 0 al 9.

Dado un número p , este número en base b se escribe

$$(p)_b = B_n b^n + B_{n-1} b^{n-1} + \dots + B_0 + B_{-1} b^{-1} + B_{-2} b^{-2} + \dots .$$

Así, el número 32.5 en base 10 representa

$$(32,5)_{10} = 3 \cdot 10 + 2 + 5 \cdot 10^{-1}$$

y el número 10.01 en base 2 representa

$$(10,01)_2 = 1 \cdot 2 + 0 + 0 \cdot 2^{-1} + 1 \cdot 2^{-2} ,$$

o sea, es el número 2.25 en base 10.

Los ordenadores utilizan una representación de los números denominada *en coma flotante*. Esta representación utiliza un conjunto de dígitos comenzando por 0., denominado *mantisa* y un *exponente*. Por ejemplo, el número 32.5 en base 10, se puede expresar como

$$0,325 \cdot 10^2 .$$

En este caso la mantisa sería 325 y el exponente 2. De igual forma para el número 10.01 en base 2, tendríamos la representación

$$0,1001 \cdot 2^2 ,$$

con lo que la mantisa sería 1001 y el exponente 2.

Los ordenadores generalmente para representar un número real utilizan una posición de memoria (bit) para representar el signo, luego un cierto número de posiciones de memoria para representar el exponente y otro número de posiciones de memoria para representar la mantisa. La representación concreta cambia de unos ordenadores a otros. Por ejemplo, para un ordenador IBM cada posición de memoria está formada por 32 bits. Un número real p , en simple precisión se almacena de la forma siguiente: Una vez escrito p como

$$p = \pm p' \cdot 16^{(p''-64)} ,$$

con $p'' \geq 0$ y $\frac{1}{16} \leq p' < 1$. Se almacena primero el signo (0 si es + y 1 si es -), en los 7 bits siguientes se almacena el exponente p'' en base 2. En las 24 posiciones restantes se almacena la mantisa en base 2. Si el número está en doble precisión, se usan 64 bits para almacenarlo y para la mantisa se usan 56 bits.

Veamos un ejemplo, si queremos representar el número 1.1 en base 10. Lo primero que haremos es representar el número $(1,1)_{10}$ en base 2. La parte

entera del número es 1 con lo que queda de la misma forma. La parte decimal es 0.1 que podemos reescribir como

$$\begin{aligned} 0,1 &= 2^{-1}(0,2) = 2^{-4}(1,6) \\ &= 2^{-4}(1 + 0,6) = 2^{-4} + 2^{-5}(1,2) \\ &= 2^{-4} + 2^{-5} + 2^{-8}1,6 = \dots, \end{aligned}$$

así,

$$(0,1)_{10} = 0,0001100110011 \dots$$

Por otra parte,

$$\begin{aligned} 1,1 &= \frac{1,1}{16}16 = 1,0001112^{-4}16^{(65-64)} \\ &= 0,00010001116^{(65-64)}. \end{aligned}$$

El exponente es $65 = (64 + 1) = 2^6 + 1$. Por lo tanto, el exponente en base 2 es

$$(65)_2 = 1000001,$$

y, por tanto, la representación del número 1.1 en simple precisión es de la forma

0	1000001	000100011001100110011001
---	---------	--------------------------

Con este tipo de representación el número más alto que es posible obtener es

0	1111111	111111111111111111111111111111
---	---------	--------------------------------

que corresponde a

$$(2^{-1} + 2^{-2} + \dots + 2^{-24}) 16^{(127-64)} \approx 16^{23} \approx 7 \cdot 10^{75}.$$

Por otra parte, el número más pequeño que podemos representar es

$$\frac{1}{16} 16^{-64} = 16^{-65} \approx 6 \cdot 10^{-79},$$

cuya representación es

0	0000000	000100000000000000000000000000
---	---------	--------------------------------

El hecho que se disponga de un número finito para representar los números reales hace que se produzcan errores al hacer operaciones aritméticas con los mismos. Estos errores se denominan *errores de truncamiento* o *errores de redondeo* dependiendo de la técnica que se utilice para cortar la representación de los números.

Veamos algunos ejemplos que ilustran los errores que se pueden producir al utilizar aritmética finita. Supongamos que se tiene un ordenador ideal que puede representar los números en base 10, pero utiliza una mantisa de dos números. En este ordenador la representación del número $x = 15\pi/34 \approx 1,38$ es $x^* = 0,14 \cdot 10^1$, con lo que se comete un error de un 1%. Si ahora consideramos $y = \sqrt{2} \approx 1,428 \dots$, la representación de este número es $y^* = 0,14 \cdot 10^1$ y la resta de estos números será

$$x^* - y^* = 0,$$

con lo que cometemos un error

$$|x - y| = 0,0282168034,$$

que corresponde a un error del 100%.

En general, al restar dos números reales próximos entre sí, x e y , se obtiene una diferencia $x^* - y^*$, que aproxima mal la verdadera diferencia $x - y$.

Otra operación en la que se pueden cometer errores al utilizar aritmética finita es la división de un número x por un número pequeño δ . Supongamos que x se aproxima por $x^* = x + \varepsilon$. Al dividir por δ , $x^*/\delta = x/\delta + \varepsilon/\delta$. El nuevo error es ε/δ . Si δ es un número pequeño, el error cometido será una cantidad grande. Un razonamiento similar se puede hacer cuando se multiplica un número x por otro número muy grande.

En ocasiones el error de redondeo que se comete al realizar un cálculo se puede minimizar reformulando el problema. Por ejemplo, supongamos que queremos calcular las raíces de la ecuación

$$ax^2 + bx + c = 0, \quad a \neq 0.$$

Estas raíces vienen dadas por las expresiones

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \quad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

Consideremos la ecuación

$$x^2 + 62,10x + 1 = 0,$$

cuyas soluciones aproximadas son

$$x_1 = -0,01610723, \quad x_2 = -62,08390.$$

Supongamos que se calcula x_1 con una aritmética de 4 dígitos

$$\begin{aligned} \sqrt{b^2 - 4ac} &= \sqrt{(62,10)^2 - 4,000} = \sqrt{3856 - 4,000} = \\ &= \sqrt{3852} = 62,06, \end{aligned}$$

así,

$$\begin{aligned} x_1 &= \frac{-b + \sqrt{b^2 - 4ac}}{2a} = \frac{-62,10 + 62,06}{2,000} = \\ &= \frac{-0,04000}{2,000} = -0,02000, \end{aligned}$$

que es una aproximación pobre de $x_1 = -0,01611$. Por otro lado, el cálculo de x_2 involucra la adición de números casi iguales y no presenta problemas,

$$\begin{aligned} x_2 &= \frac{-b - \sqrt{b^2 - 4ac}}{2a} = \frac{-62,10 - 62,06}{2,000} = \\ &= \frac{-124,2}{2,000} = -62,10. \end{aligned}$$

Para obtener una mejor aproximación de x_1 tenemos en cuenta que

$$\begin{aligned} x_1 &= \frac{-b + \sqrt{b^2 - 4ac}}{2a} = \frac{b^2 - (b^2 - 4ac)}{2a(-b - \sqrt{b^2 - 4ac})} = \\ &= \frac{-2c}{b + \sqrt{b^2 - 4ac}} = \frac{-2,000}{62,10 + 62,06} = -0,0161. \end{aligned}$$

1.3. Propagación de errores

Supongamos que se quiere dar una cota al error que se comete al realizar un cálculo $y = f(x)$, si la variable x viene afectada de un error dado. Así, si $x = x^* + \Delta x$, y $f(x)$ es una función derivable, podemos escribir

$$y = f(x) = f(x^* + \Delta x) \approx f(x^*) + f'(x^*) \Delta x,$$

con lo que el error

$$\varepsilon f(x^*) = |y - f(x^*)| \approx |f'(x^*)| |\Delta x|.$$

Si se tiene un problema multidimensional $y = f(x_1, x_2, \dots, x_n)$, podemos escribir

$$\begin{aligned} |y - f(x_1^*, x_2^*, \dots, x_n^*)| &\approx \left| \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_1^*, x_2^*, \dots, x_n^*) \Delta x_i \right| \\ &\leq \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(x_1^*, x_2^*, \dots, x_n^*) \right| |\Delta x_i|. \end{aligned}$$

Así una expresión para una cota del error cometido es de la forma

$$\varepsilon f(x_1^*, x_2^*, \dots, x_n^*) = \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(x_1^*, x_2^*, \dots, x_n^*) \right| |\Delta x_i|,$$

que suele funcionar bien cuando n no es muy grande. En el caso que se tengan muchas variables esta expresión proporciona una cota muy conservadora del error cometido.

Otra posibilidad para definir un error es utilizar el *error cuadrático medio*

$$\varepsilon f(x_1^*, \dots, x_n^*) = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x_1^*, \dots, x_n^*) \Delta x_i \right)^2}.$$

1.4. Ejercicios

1. Escribir en base 2 los números:

a) 1.5, b) 1.4, c) 0.3 .

2. Escribir la representación en un ordenador IBM en simple precisión de $x = -232$.

3. Escribir en base 2 el número cuya representación en base 10 es 100.3. Utilizar este resultado para obtener la representación en simple precisión de este número en un ordenador IBM.

4. Dado el sistema de ecuaciones

$$\begin{aligned} 3x + ay &= 10, \\ 5x + by &= 20, \end{aligned}$$

donde $a = 2,100 \pm 5 \cdot 10^{-4}$ y $b = 3,300 \pm 5 \cdot 10^{-4}$. Obtener una cota del error asociado a x e y al resolver el sistema.

- Con qué exactitud se ha de medir el radio de una esfera y con cuántos decimales se ha de dar el número π , para que su volumen se conozca con un error realivo menor que el 0.01 %.
- Supongamos que se quiere hacer el cálculo

$$y = \ln \left(x - \sqrt{x^2 - 1} \right) ,$$

para $x = 30$, y que al calcular la raíz cuadrada tenemos un método que sólo nos da 6 decimales correctos. Estimad una cota del error que se comete al hacer el cálculo. Reescribir la función $y(x)$ para que el error cometido al hacer el cálculo sea menor.

- Consideremos una barra de longitud l y sección rectangular, de anchura a y altura b , empotrada por uno de sus extremos. Si en el extremo libre se aplica una fuerza F perpendicular a la barra, la flexión s experimentada viene dada por la expresión

$$s = \frac{4l^3}{Eab^3} F ,$$

donde E es el módulo de Young del material que forma la barra. De un ensayo, se sabe que una fuerza F de 140 kp aplicada a una barra de hierro de 125 cm de longitud y sección cuadrada de 2.4 cm de lado le produce una flexión de 1.71 mm. Calculad el módulo de Young del hierro y una cota del error asociado, si se supone que las medidas tomadas tienen la precisión asociada con la última cifra decimal.

- Calcular el error permitido en la inclinación de un cañón para asegurar que acierte en un objetivo rectangular de 40 m de longitud y 20 m de anchura, situado a la misma altura que el cañón, y cuyo centro está a una distancia de 3000 m de éste. La velocidad de salida del proyectil es de 600 m/s. Supóngase que todas las magnitudes son exactas y que las dimensiones del objetivo son pequeñas en relación a la distancia a la que se encuentra el cañón. Considérese también que se apunta al centro del objetivo y que está permitido tocar en cualquier lugar de éste.

Capítulo 2

Resolución de sistemas de ecuaciones

2.1. Introducción

Muchos problemas de ingeniería requieren para su solución plantear y resolver un sistema de ecuaciones algebraicas. Por ejemplo, supongamos que se tiene una estructura articulada como la de la figura

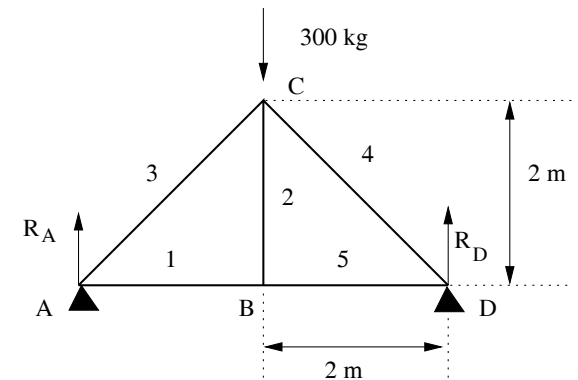


Figura 2.1.- Viga articulada.

La primera ecuación que se puede plantear es que la carga se compensa por las reacciones en los apoyos,

$$R_A + R_D = 300 ,$$

Planteando que la suma de las fuerzas sobre el nodo A son cero

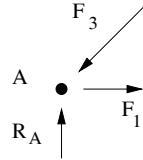


Figura 2.2.- Nodo A.

$$F_1 - F_3 \cos\left(\frac{\pi}{4}\right) = 0 ,$$

$$R_A - F_3 \sin\left(\frac{\pi}{4}\right) = 0 ,$$

para el nodo B se tiene

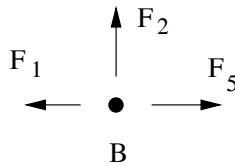


Figura 2.3.- Nodo B.

$$F_5 - F_1 = 0 ,$$

$$F_2 = 0 ,$$

y para el nodo C

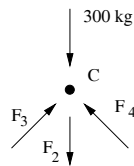


Figura 2.4.- Nodo C.

$$F_3 \cos\left(\frac{\pi}{4}\right) + F_4 \cos\left(\frac{\pi}{4}\right) - F_2 = 300 ,$$

$$F_3 \sin\left(\frac{\pi}{4}\right) - F_4 \sin\left(\frac{\pi}{4}\right) = 0 .$$

Con esto tenemos 7 ecuaciones con 7 incógnitas, que podemos escribir mediante la siguiente expresión matricial

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 1 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} R_A \\ R_D \\ F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \end{bmatrix} = \begin{bmatrix} 300 \\ 0 \\ 0 \\ 0 \\ 0 \\ 300 \\ 0 \end{bmatrix}$$

que tiene la forma genérica

$$Ax = b$$

donde A es la matriz de coeficientes del sistema y b el vector de términos independientes.

Formalmente la solución de este sistema se puede obtener como

$$x = A^{-1}b ,$$

pero en la mayoría de ocasiones no resulta conveniente calcular la matriz inversa, A^{-1} , para resolver el sistema y se recurre a técnicas alternativas. Nosotros en particular estudiaremos dos tipos de métodos para resolver los sistemas de ecuaciones lineales, métodos directos y métodos iterativos.

2.2. Métodos directos para la resolución de sistemas

Generalmente, los métodos directos se basan en transformar el sistema de ecuaciones inicial en un sistema equivalente (o sea, que tenga las mismas soluciones) que se pueda resolver más fácilmente. Para ello, se hace uso de la siguiente propiedad:

Dado un sistema de ecuaciones, si a una de las ecuaciones se le suma una combinación lineal de las ecuaciones restantes, el sistema de ecuaciones resultante es equivalente al sistema inicial.

que es la propiedad que utiliza el método de triangularización de Gauss. Recordemos cómo funciona este método con un ejemplo. Dado el sistema

$$\begin{bmatrix} 10 & -7 & 0 \\ -3 & 2 & 6 \\ 5 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ 6 \end{bmatrix},$$

si le restamos $-3/10$ veces la primera ecuación a la segunda obtenemos el sistema

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0,1 & 6 \\ 5 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 6,1 \\ 6 \end{bmatrix}.$$

Ahora restamos $5/10$ veces la primera ecuación a la tercera y obtenemos el sistema equivalente

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0,1 & 6 \\ 0 & 2,5 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 6,1 \\ 2,5 \end{bmatrix}.$$

Por último, si a la tercera ecuación la restamos $-2,5/0,1$ veces la segunda se llega al sistema

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0,1 & 6 \\ 0 & 0 & 155 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 6,1 \\ 155 \end{bmatrix}.$$

Con lo que hemos llegado a un sistema de ecuaciones equivalente al inicial cuya matriz de coeficientes es una matriz triangular superior. Este tipo de sistemas tiene una solución sencilla. Se empieza despejando la última ecuación, la solución obtenida se utiliza para despejar la penúltima ecuación y así sucesivamente.

Este método se llama el método de sustitución regresiva y se formaliza del modo siguiente. Dado el sistema

$$Ux = b, \quad U \in \mathbb{R}^{n \times n}, \quad b \in \mathbb{R}^{n \times 1},$$

donde $u_{ij} = 0$ si $i < j$ (triangular superior), un posible algoritmo que implementa el método es el siguiente

```
x=zeros(n,1);
for k=n:-1:1
    x(k)=b(k)/U(k,k);
    for i=1:k-1
        b(i) = b(i)-x(k)*U(i,k);
    end
end
```

2.2.1. Factorización LU de una matriz

Partimos del sistema de ecuaciones inicial

$$Ax = b,$$

con

$$A = \begin{bmatrix} 10 & -7 & 0 \\ -3 & 2 & 6 \\ 5 & -1 & 5 \end{bmatrix}$$

Se puede descomponer la matriz A como el producto

$$A = L_1 A_1,$$

con

$$L_1 = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{3}{10} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 10 & -7 & 0 \\ 0 & -0,1 & 6 \\ 5 & -1 & 5 \end{bmatrix}.$$

Por otra parte

$$A_1 = L_2 A_2,$$

con

$$L_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{5}{10} & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 10 & -7 & 0 \\ 0 & -0,1 & 6 \\ 0 & 2,5 & 5 \end{bmatrix}.$$

Por último,

$$A_2 = L_3 U,$$

con

$$L_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{2,5}{0,1} & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 10 & -7 & 0 \\ 0 & -0,1 & 6 \\ 0 & 0 & 155 \end{bmatrix}.$$

Con lo que tenemos que

$$A = L_1 L_2 L_3 U = LU,$$

donde U es la matriz triangular superior resultante al aplicar el método de Gauss y

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{3}{10} & 1 & 0 \\ \frac{5}{10} & -\frac{2,5}{0,1} & 1 \end{bmatrix}$$

es una matriz triangular inferior, cuyos elementos son los factores utilizados para triangularizar la matriz A . Este proceso es general y permite descomponer una matriz cuadrada A como producto de una matriz triangular inferior, L , por una matriz triangular superior, U .

2.2.2. Pivotamiento

Los elementos de la diagonal de la matriz U se llaman pivotes. El algoritmo de la sustitución regresiva realiza divisiones por los distintos pivotes, de este modo, el algoritmo no podrá llevarse a cabo si alguno de los pivotes es cero. Es lógico pensar que si alguno de los pivotes es muy pequeño se producirán errores grandes al calcular la solución.

Veamos un ejemplo. Supongamos que se quiere resolver el sistema

$$\begin{bmatrix} 10 & -7 & 0 \\ -3 & 2,099 & 6 \\ 5 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 3,901 \\ 6 \end{bmatrix},$$

cuya solución exacta es $x = (0, -1, 1)$. Ahora resolvemos el sistema mediante el método de Gauss con una aritmética de 5 dígitos significativos

$$\left[\begin{array}{ccc|c} 10 & -7 & 0 & 7 \\ -3 & 2,099 & 6 & 3,901 \\ 5 & -1 & 5 & 6 \end{array} \right] \approx \left[\begin{array}{ccc|c} 10 & -7 & 0 & 7 \\ 0 & -0,001 & 6 & 6,001 \\ 0 & 2,5 & 5 & 2,5 \end{array} \right] \approx \left[\begin{array}{ccc|c} 10 & -7 & 0 & 7 \\ 0 & -0,001 & 6 & 6,001 \\ 0 & 0 & 1,5005 \cdot 10^4 & 1,5004 \cdot 10^4 \end{array} \right]$$

Al hacer el cálculo para la sustitución regresiva, se tiene

$$x_3 = \frac{1,5004 \cdot 10^4}{1,5005 \cdot 10^4} = 0,99993,$$

que comparado con el resultado exacto, $x_3 = 1$, nos da un resultado aceptable. Para x_2 , se tiene

$$-0,001x_2 + (6)(0,99993) = 6,001,$$

o sea,

$$x_2 = -1,5,$$

que ya no es un resultado aceptable, comparado con el valor exacto $x_2 = -1$. Este problema se debe principalmente a la propagación del error debida a la división por el pivote 0,001. Esto se resuelve intercambiando la segunda ecuación por la tercera, ya que al realizar la triangularización ya no se obtiene un pivote tan pequeño. Esta estrategia de intercambiar filas se conoce como pivotación.

Si se permite el intercambio de filas, el algoritmo de Gauss nos lleva a una descomposición para la matriz A de la forma

$$PA = LU,$$

donde P es una matriz de permutación.

Una estrategia que se suele utilizar para evitar estos problemas del algoritmo de Gauss es la estrategia de *pivotación parcial*, que consiste en tomar como pivote en el paso k -ésimo del algoritmo de triangularización el elemento más grande en valor absoluto en la parte no reducida de la columna k -ésima. La fila que contiene este pivote se intercambia con la fila k -ésima para poner el pivote en la posición (k, k) de la matriz. Los mismos intercambios se han de llevar a cabo en el vector b .

A continuación se muestran dos funciones de Matlab, la primera `solvelu()` implementa el método de Gauss con pivotación parcial para triangularizar la matriz A . La segunda, `trisuper()` calcula la sustitución regresiva sobre la matriz triangularizada y el vector de términos independientes.

```
function [An,bn,v]=solvelu(A,b)
% [An,bn,v]=solvelu(A,b)
% Esta funcion realiza la trangularizacion
% de una matriz A y el termino independiente b
% para la resolucion de un sistema Ax=b.
%
% Entradas:
% A matriz del sistema n x n
% b termino independiente
% Salidas:
% An matriz triangularizada
% bn termino independiente
% v vector de punteros para la pivotacion
% parcial

[m,n]=size(A);
if (m~=n)
    error('La matriz A no es cuadrada');
end
An=A;
bn=b;
v=1:n; % permutaciones
```

```

for k=1:n-1
    M=k;
    for i=k+1:n
        if (abs(An(v(i),k))>abs(An(v(M),k)))
            M=i;
        end
    end
    if (k~=M)
        it=v(k);
        v(k)=v(M);
        v(M)=it;
    end
    if (abs(An(v(k),k))<=1.e-8)
        error('matriz singular');
    end
    for i=k+1:n
        alfa=An(v(i),k)/An(v(k),k);
        for j=k+1:n
            An(v(i),j)=An(v(i),j)-alfa*An(v(k),j);
        end
        bn(v(i))=bn(v(i))-alfa*bn(v(k));
    end
end

function x=trisuper(An,bn,v)
% x=trisuper(An,bn,v)
% Esta funcion resuelve un sistema triangular
% por el metodo de sustitucion regresiva.
% Se usa despues de la funcion solvelu
% Entradas:
% An matriz triangular superior como sale de solvelu
% bn termino independiente como sale de solvelu
% v vector de puneteros de la pivotacion parcial
% Salida:
% x solucion del sistema
[m,n]=size(An);
if (m~=n)
    error('La matriz A no es cuadrada');
end
for i=n:-1:1
    x(i)=bn(v(i));

```

```

for j=n:-1:i+1
    x(i)=x(i)-An(v(i),j)*x(j);
end
x(i)=x(i)/An(v(i),i);
end

```

2.2.3. Errores y número de condición

Cuando se obtiene la solución numérica de un sistema de ecuaciones, se obtiene un valor aproximado de la solución, x^* , mientras que la solución exacta, x , satisface

$$x = A^{-1}b .$$

Usualmente se tienen dos magnitudes que nos dan una idea del error cometido, el *error*, definido como

$$e = x - x^* ,$$

y el *residuo*

$$r = b - Ax^* .$$

Estas dos magnitudes no tienen porque ser pequeñas al mismo tiempo si se trabaja con aritmética finita.

Por otra parte, cuando se están resolviendo problemas prácticos los coeficientes de un sistema de ecuaciones y los términos independientes están afectados de un cierto error. Por ello, es interesante preguntarse cómo se pueden medir lo que cambia la solución, x , de un sistema si se hacen cambios en A y o en b .

Para evaluar esta dependencia se introduce el *número de condición* de una matriz, definido como

$$\kappa(A) = \frac{M}{m} ,$$

donde

$$M = \max \left(\frac{\|Ax\|}{\|x\|} \right) , \quad m = \min \left(\frac{\|Ax\|}{\|x\|} \right) \quad \forall x \neq 0 .$$

Consideramos

$$Ax = b ,$$

y el sistema

$$A(x + \delta x) = b + \delta b ,$$

se cumple

$$A\delta x = \delta b ,$$

además,

$$\|b\| \leq M \|x\| , \quad \|\delta b\| \geq m \|\delta x\| ,$$

y si $m \neq 0$, se tiene

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\delta b\|}{\|b\|} ,$$

así, el número de condición nos da un factor de magnificación del error. De este modo, el número de condición de la matriz del sistema es alto, cabe esperar que la solución numérica obtenida esté afectada de errores altos.

Si se introduce el concepto de norma matricial

$$\|A\| \equiv \max \left(\frac{\|Ax\|}{\|x\|} \right) , \quad \forall x \neq 0 ,$$

una definición del número de condición es

$$\kappa(A) = \|A\| \|A\|^{-1} .$$

El número de condición es un número difícil de calcular y se utilizan estimaciones para este número. Generalmente el número de condición de una matriz es muy alto si la matriz es casi singular.

2.2.4. Matrices especiales

Veremos ahora cómo es posible simplificar los métodos que hemos visto hasta ahora cuando la matriz de coeficientes tiene una estructura particular.

Matrices tridiagonales

Un caso especial de sistemas de ecuaciones que aparecen frecuentemente son aquellos cuya matriz de coeficientes es tridiagonal, o sea, un sistema de ecuaciones con la siguiente estructura

$$\begin{bmatrix} b_1 & c_1 & 0 & \cdots & & & \\ a_1 & b_2 & c_2 & \cdots & & & \\ & & \cdots & & & & \\ & & & a_{n-2} & b_{n-1} & c_{n-1} & \\ & & & 0 & a_{n-1} & b_n & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix} .$$

Para estos sistemas, el algoritmo de Gauss adopta una forma simple. La matriz se tiene almacenada en los vectores a , b , c , y el término independiente en el vector d .

Primero se copia en x el vector d

$x = d$;

luego se triangulariza la matriz y el término independiente mediante el siguiente bucle

```
for j=1:n-1
  mu=a(j)/b(j);
  b(j+1)=b(j+1)-mu*c(j);
  x(j+1)=x(j+1)-mu*x(j);
end
```

Posteriormente, se realiza la sustitución regresiva

```
x(n)=x(n)/b(n);
for j=n-1:-1:1
  x(j)=(x(j)-c(j)*x(j+1))/b(j);
end
```

Este algoritmo se conoce como el *algoritmo de Thomas*. No utiliza pivotación parcial y es mucho más rápido que el algoritmo de Gauss.

Matrices simétricas definidas positivas

Recordemos que una matrix A es simétrica si cumple que $A = A^T$. Además diremos que una matriz es definida positiva si cumple

$$x^T A x > 0 , \quad \forall x \neq 0 .$$

Si A es una matriz simétrica y definida positiva se puede encontrar una matriz triangular inferior, L , de forma que

$$LL^T = A .$$

Esta descomposición se denomina descomposición de Choleski de la matriz A .

Veamos cómo se puede calcular la matriz L . Para ello, consideremos un caso 3×3 .

$$\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

calculando el producto, se tienen las relaciones

$$\begin{aligned} a_{11} &= l_{11}^2, \\ a_{22} &= l_{21}^2 + l_{22}^2, \\ a_{33} &= l_{31}^2 + l_{32}^2 + l_{33}^2, \\ a_{12} &= l_{11}l_{21}, \\ a_{13} &= l_{11}l_{31}, \\ a_{23} &= l_{21}l_{31} + l_{22}l_{32} \end{aligned}$$

o sea,

$$\begin{aligned} l_{11} &= (a_{11})^{\frac{1}{2}}, \\ l_{21} &= \frac{a_{12}}{l_{11}}, \\ l_{22} &= (a_{22} - l_{21}^2)^{\frac{1}{2}}, \\ l_{31} &= \frac{a_{13}}{l_{11}}, \\ l_{32} &= \frac{a_{23} - l_{21}l_{31}}{l_{22}}, \\ l_{33} &= (a_{33} - l_{31}^2 - l_{32}^2)^{\frac{1}{2}}. \end{aligned}$$

Para una matriz $n \times n$, los elementos de L se pueden calcular mediante las expresiones

$$\begin{aligned} l_{ii} &= \left(a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 \right)^{\frac{1}{2}}, \\ l_{ji} &= \frac{1}{l_{ii}} \left(a_{ij} - \sum_{k=1}^{i-1} l_{ik}l_{jk} \right), \quad j = i+1, i+2, \dots, n. \end{aligned}$$

2.3. Métodos iterativos

Los métodos iterativos para la resolución de sistemas de ecuaciones lineales suelen utilizarse para problemas de gran dimensión, ya que usan menos

memoria y suelen ser más rápidos que los métodos directos. A continuación veremos alguno de los métodos iterativos más sencillos.

Se parte de un sistema de ecuaciones de la forma

$$Ax = b,$$

y realizamos la descomposición de la matriz de coeficientes

$$A = D - E - F,$$

donde D es la diagonal de A , $-E$ es la parte estrictamente triangular inferior de A y $-F$ es la parte estrictamente triangular superior. se supone que los elementos de D son todos no nulos.

El método de Jacobi se basa en iteraciones de la forma

$$Dx^{k+1} = (E + F)x^k + b,$$

o sea,

$$x^{k+1} = D^{-1}(E + F)x^k + D^{-1}b.$$

Otro método similar al método de Jacobi es el método de Gauss-Seidel. Este método en componentes se escribe de la siguiente forma

$$b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{k+1} - a_{ii}x_i^{k+1} - \sum_{j=i+1}^n a_{ij}x_j^k = 0,$$

para $i = 1, 2, \dots, n$.

En forma matricial el método de Gauss-Seidel se escribirá como

$$(D - E)x^{k+1} = Fx^k + b.$$

Para implementar este método hay que resolver un sistema triangular inferior.

Análogamente, se puede definir otro método de Gauss-Seidel de la forma,

$$(D - F)x^{k+1} = Ex^k + b,$$

donde se tendría que resolver un sistema triangular superior.

Por otra parte, podemos definir otra descomposición de la matriz A de la forma

$$\omega A = (D - \omega E) - (\omega F + (1 - \omega)D),$$

que da lugar al método iterativo conocido como el método SOR (successive over relaxation)

$$(D - \omega E)x^{k+1} = (\omega F + (1 - \omega)D)x^k + \omega b ,$$

donde ω es un parámetro que puede tomar distintos valores y que sirve para mejorar la convergencia del método.

Análogamente, se puede definir otro método SOR de la forma

$$(D - \omega F)x^{k+1} = (\omega E + (1 - \omega)D)x^k + \omega b .$$

Un método SOR simétrico, SSOR, viene definido por las ecuaciones

$$\begin{aligned} (D - \omega E)x^{k+1/2} &= (\omega F + (1 - \omega)D)x^k + \omega b , \\ (D - \omega F)x^{k+1} &= (\omega E + (1 - \omega)D)x^{k+1/2} + \omega b . \end{aligned}$$

2.4. Ejercicios

1. Resuelve el sistema

$$\begin{bmatrix} 0 & 2 & 4 \\ 1 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 3 \end{bmatrix} .$$

2. Dada la matriz

$$\begin{bmatrix} 3 & 0 & 2 \\ 0 & 4 & 3 \\ 0 & 0 & 2 \end{bmatrix} ,$$

obtén su inversa.

3. Dada la matriz

$$\begin{bmatrix} \alpha & 1 & 0 \\ \beta & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} ,$$

obtén todos los valores de α y β para los cuales la matriz es singular.

4. Resuelve los siguientes sistemas de ecuaciones factorizados:

a)

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & -1 \\ 0 & -2 & 1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} ,$$

b)

$$\begin{bmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} ,$$

5. Compara la solución que se obtiene mediante el método de Gauss para los sistemas de ecuaciones

$$\begin{aligned} x - y &= 1 \\ x - 1,01y &= 0 \end{aligned}$$

y

$$\begin{aligned} x - y &= 1 \\ x - 0,99y &= 0 \end{aligned}$$

6. Dado el sistema

$$\begin{bmatrix} 0,15 & 2,11 & 30,75 \\ 0,64 & 1,21 & 2,05 \\ 3,21 & 1,53 & 1,04 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -26,38 \\ 1,01 \\ 5,23 \end{bmatrix}$$

resuélvelo utilizando el método de Gauss sin pivotación parcial y con pivotación parcial.

7. Consideremos el sistema de ecuaciones

$$Ax = b ,$$

con

$$A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix} , \quad b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} ,$$

comprueba que el método de Jacobi diverge mientras que el método de Gauss-Seidel converge.

8. Determina las dos primeras iteraciones del método de Jacobi para los siguientes sistemas de ecuaciones, si tomamos $x_0 = (0, 0, 0)^T$.

a)

$$\begin{aligned} 10x_1 - x_2 &= 9 \\ -x_1 + 10x_2 - 2x_3 &= 7 \\ -2x_2 + 10x_3 &= 6 \end{aligned}$$

b)

$$\begin{aligned} 4x_1 + x_2 - x_3 + x_4 &= -2 \\ x_1 + 4x_2 - x_3 - x_4 &= -1 \\ -x_1 - x_2 + 5x_3 + x_4 &= 0 \\ x_1 - x_2 + x_3 + 3x_4 &= 1 \end{aligned}$$

9. Dado el sistema

$$\begin{aligned} 3x_1 + 2x_2 &= 1 \\ 4x_1 + 3x_2 &= 5 \end{aligned}$$

calcula dos iteraciones del método SSOR partiendo del vector inicial $(x_1^0, x_2^0) = (0, 0)$ y tomando $\omega = 1,25$.

10. Dado el sistema

$$\begin{aligned} 4x_1 + 3x_2 &= 24 \\ 3x_1 + 4x_2 - x_3 &= 30 \\ -x_2 + 4x_3 &= -24 \end{aligned}$$

compara 3 iteraciones del método de SOR y el método SSOR tomando $x_0 = (0, 0, 0)^T$, y $\omega = 1,25$.

Capítulo 3

Ecuaciones y sistemas de ecuaciones no lineales

En este tema trataremos de presentar algunos métodos numéricos para la resolución de ecuaciones no lineales donde, generalmente, no es posible 'despejar' las incógnitas.

3.1. Ecuaciones no lineales

3.1.1. Método de la bisección

Supongamos, por ejemplo, que se quiere calcular una solución de la ecuación

$$x^2 = 2,$$

sabiendo que la solución está en un intervalo $[a, b] = [1, 2]$. Probamos con un punto c que sea el punto medio del intervalo

$$c = \frac{a+b}{2},$$

calculamos $c^2 = (1,5)^2 = 2,25$. Como $c^2 > 2$, entonces tendremos que la raíz estará en un nuevo intervalo $[a', b'] = [a, c]$. Repitiendo esta estrategia se van obteniendo intervalos cada vez más pequeños que contienen la raíz buscada.

El siguiente código de Matlab resuelve el problema

```
M=2;
```

```

a=1;
b=2;
k=0;
tol=1.0e-5;
while (b-a) >tol
  x=(a+b)/2;
  if x^2>M
    b=x;
  else
    a=x;
  end
  k=k+1;
end
sol=(a+b)/2

```

Éste es un método lento pero seguro para obtener una raíz de una ecuación del tipo

$$f(x) = 0.$$

Si $f(x)$ es una función continua, sólo hace falta conocer un intervalo $[a, b]$ de forma que se satisfaga $f(a)f(b) < 0$.

3.1.2. Método del punto fijo

Definición 3.1 Diremos que un punto p es un punto fijo de una función $\varphi(x)$ si se satisface $\varphi(p) = p$.

Supongamos que se busca una raíz de una ecuación

$$f(x) = 0,$$

el método del punto fijo consiste en reescribir esta ecuación de la forma

$$x = \varphi(x),$$

y construir una sucesión de la forma

$$\begin{aligned}
x_1 &= \varphi(x_0) \\
x_2 &= \varphi(x_1) \\
&\vdots \\
x_{n+1} &= \varphi(x_n)
\end{aligned} \tag{3.1}$$

Si existe un m tal que $x_{m+1} \approx x_m$, se cumplirá que $x_m \approx \varphi(x_m)$ y, por tanto, se podrá tomar como valor aproximado de la raíz x_m .

Ejemplo 3.1 Suponemos que se quiere buscar una raíz de la función

$$f(x) = x^3 + 4x^2 - 10$$

en $[1, 2]$. Se pueden hacer diferentes elecciones de la función $\varphi(x)$, por ejemplo,

$$\begin{aligned}
a) \varphi_1(x) &= x - x^3 - 4x^2 + 10; & b) \varphi_2(x) &= \left(\frac{10}{x} - 4x\right)^{1/2}; \\
c) \varphi_3(x) &= \frac{1}{2}(10 - x^3)^{1/2}; & d) \varphi_4(x) &= (10/(4+x))^{1/2}; \\
e) \varphi_5(x) &= x - (x^3 + 4x^2 - 10)/(8x + 3x^2)
\end{aligned}$$

si tomamos como valor inicial $x_0 = 1,5$ y construimos las sucesiones correspondientes a las distintas elecciones, se obtienen los valores de la tabla 3.1.

Tabla 3.1.- Resultado para la iteración del punto fijo $x = \varphi(x)$.

n	$a)$	$b)$	$c)$	$d)$	$e)$
0	1.5	1.5	1.5	1.5	1.5
1	-0.875	0.8165	1.286953768	1.348399725	1.37333333
2	6.732	2.9969	1.402540804	1.367376372	1.365262015
3	-469.7	-	1.345458374	1.364957015	1.365230014
4	1.03×10^8	-	1.375170253	1.365264748	1.365230013
5	-	-	1.360094193	1.365225594	-
6	-	-	1.367846968	1.365230576	-
7	-	-	1.363887004	1.365229942	-
8	-	-	1.365916734	1.365230022	-
9	-	-	1.364878217	1.365230012	-
10	-	-	1.365410062	1.365230014	-

Para obtener una solución por este método, es necesario que la sucesión $\{x_n\}_{n=1}^{\infty}$ construida como en (3.1) sea convergente, y el límite de esta sucesión coincida con la raíz. Las condiciones necesarias para que esto ocurra se resumen en el siguiente teorema.

Teorema 3.1 (Teorema del punto fijo) Sea $\varphi : [a, b] \rightarrow [a, b]$ una función continua en $[a, b]$ y derivable en $]a, b[$, y además que cumple que $|\varphi'(x)| \leq k < 1, \forall x \in]a, b[$. Entonces existe un único $c \in]a, b[$ tal que $\varphi(c) = c$. Además para todo $x_0 \in]a, b[$, la sucesión $x_0, \{x_n\}_{n=0}^{\infty}$ obtenida de la forma $x_n = \varphi(x_{n-1})$ converge a c .

Este teorema ayuda a elegir una función $\varphi(x)$ adecuada, por ejemplo, para la opción *d*) del ejemplo anterior se tiene que

$$\begin{aligned}\varphi_4(x) &= (10/(4+x))^{1/2}; \\ \varphi'_4(x) &= \sqrt{10}(-1/2)(4+x)^{-3/2}; \\ |\varphi'_4(x)| &= \sqrt{10}/2|(4+x)^{-3/2}| < \sqrt{10}/(2 \cdot 5^{3/2}) = 0,141 < 1.\end{aligned}$$

Luego la sucesión converge en $[1, 2]$.

3.1.3. Método de Newton-Raphson

Este método se basa en utilizar el desarrollo de Taylor para aproximar una función derivable en las proximidades de un punto.

Escribimos $x_1 = x_0 + \Delta x$, y usando el desarrollo de Taylor

$$f(x_1) = f(x_0) + f'(x_0)\Delta x + o(\Delta x^2),$$

y suponiendo que $f(x_1) = 0$, queda

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

El método de Newton-Raphson se basa en esta ecuación y consiste en calcular los valores de una sucesión de la forma

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Otro modo de obtener este método consiste en suponer que $f: [a, b] \rightarrow \mathbb{R}$ es continua en $[a, b]$ y tal que $f''(x)$ no cambia de signo en $[a, b]$ con $f(a)f(b) < 0$. El proceso para encontrar un x tal que $f(x) = 0$ consiste en lo siguiente:

- 1) Fijamos $c = a$ ó b , tal que $f(c)f''(x) > 0, \forall x \in (a, b)$.
- 2) $x_0 = c$.
- 3) Hallamos la ecuación de la tangente que pasa por $(x_0, f(x_0))$ y el punto de corte de dicha tangente con el eje X . El proceso se repite hasta conseguir una sucesión de aproximaciones que converge a la raíz de $f(x) = 0$.

La ecuación de la recta tangente a la curva $y = f(x)$ en $(x_{n-1}, f(x_{n-1}))$ viene dada por

$$y = f(x_{n-1}) + f'(x_{n-1})(x - x_{n-1}).$$

La abscisa del punto de intersección de la recta tangente con el eje X , viene dada por

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})},$$

y mediante esta relación obtenemos una sucesión, $\{x_n\}_{n=1}^{\infty}$ de aproximaciones al valor de la raíz buscada.

El error que se comete en la iteración n -ésima será

$$|r - x_n| < \frac{M}{2m} |x_n - x_{n-1}|,$$

donde $0 < m \leq |f(x)|$ y $|f''(x)| \leq M, \forall x \in (a, b)$.

Ejemplo 3.2 Se busca la solución de $x = \cos(x)$. Se construye la función

$$f(x) = x - \cos(x),$$

utilizando la sucesión

$$x_{n+1} = x_n - \frac{x_n - \cos(x_n)}{1 + \sin(x_n)},$$

y el punto inicial $x_0 = 0,7$, se obtiene la siguiente tabla de valores

Tabla 3.2.- Resultado de la iteración de Newton.

n	x_n
1	0.7393981635
2	0.7395361337
3	0.7390851781
4	0.7390851332
5	0.7390851332

3.1.4. Método de la secante

Una de las desventajas del método de Newton para obtener las raíces de una ecuación de la forma

$$f(x) = 0,$$

es que es necesario conocer la derivada $f'(x)$. En ocasiones esta derivada es difícil de calcular o no se dispone de la misma, y se utiliza una aproximación de la forma

$$s_n = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}},$$

obteniendo el método de la secante, que se basa en iteraciones de la forma

$$x_{n+1} = x_n - \frac{f(x_n)}{s_n}.$$

El método de la secante hace uso de dos aproximaciones iniciales para la raíz.

3.2. Resolución numérica de sistemas de ecuaciones

Los métodos anteriores se pueden generalizar para el caso de sistemas de ecuaciones. Consideraremos sólo sistemas de dos ecuaciones con dos incógnitas, la generalización para un número mayor de ecuaciones es sencilla.

3.2.1. Método iterativo del punto fijo

Partimos de un sistema de la forma

$$\begin{aligned} f_1(x, y) &= 0, \\ f_2(x, y) &= 0, \end{aligned}$$

del que se pretende obtener la solución. Para utilizar el método del punto fijo, se reescribe el sistema de la forma

$$\begin{aligned} x &= g_1(x, y), \\ y &= g_2(x, y). \end{aligned}$$

Se construye la sucesión

$$(x_{n+1}, y_{n+1}) = (g_1(x_n, y_n), g_2(x_n, y_n))$$

que como ocurre en el caso escalar la eficiencia del método dependerá de la elección de las funciones g_1 y g_2 .

Ejemplo 3.3 Suponemos que se quiere buscar una solución del sistema

$$\begin{aligned} x^2 - 10x + y^2 + 8 &= 0, \\ xy^2 + x - 10y + 8 &= 0. \end{aligned}$$

Se eligen

$$\begin{aligned} x &= \frac{1}{10}(x^2 + y^2 + 8), \\ y &= \frac{1}{10}(xy^2 + x + 8). \end{aligned}$$

Se parte de $(x, y) = (0, 0)$ y, utilizando el método del punto fijo, se obtiene la sucesión mostrada en la tabla 3.3.

Tabla 3.3.- Resultado de la iteración del punto fijo para el sistema.

x_n	y_n
0	0
0.8000	0.8000
0.9414	0.9670
0.9821	0.9901
0.9945	0.9969
0.9983	0.9990
0.9995	0.9997
0.9998	0.9999
0.9999	1

en la siguiente iteración se obtiene que $(x_n, y_n) = (1, 1)$.

3.2.2. Método de Newton-Raphson

Este método se basa en utilizar el desarrollo de Taylor para aproximar una función derivable en las proximidades de un punto en este caso de funciones de dos variables. Partimos de un sistema de la forma

$$\begin{aligned} f_1(x, y) &= 0, \\ f_2(x, y) &= 0, \end{aligned}$$

del que se pretende obtener la solución. Se supone que

$$x = x_0 + \Delta x \quad \text{e} \quad y = y_0 + \Delta y$$

luego

$$\begin{aligned} f_1(x_0 + \Delta x, y_0 + \Delta y) &= 0, \\ f_2(x_0 + \Delta x, y_0 + \Delta y) &= 0. \end{aligned}$$

Utilizando el desarrollo de Taylor alrededor de (x_0, y_0) y quedandonos con el primer orden, tenemos

$$\begin{aligned} f_1(x_0, y_0) + \frac{\partial f_1(x_0, y_0)}{\partial x} \Delta x + \frac{\partial f_1(x_0, y_0)}{\partial y} \Delta y &\approx 0, \\ f_2(x_0, y_0) + \frac{\partial f_2(x_0, y_0)}{\partial x} \Delta x + \frac{\partial f_2(x_0, y_0)}{\partial y} \Delta y &\approx 0, \end{aligned}$$

si introducimos la notación $\vec{x} = (x, y)$, $\vec{F} = (f_1, f_2)$ y

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix}$$

queda

$$\vec{F}(\vec{x}_0) + J(\vec{x}_0) (\vec{x} - \vec{x}_0) \approx 0$$

entonces

$$\vec{x} = \vec{x}_0 - J^{-1}(\vec{x}_0) \vec{F}(\vec{x}_0)$$

y el método de Newton consiste en calcular la sucesión

$$\vec{x}_{n+1} = \vec{x}_n - J^{-1}(\vec{x}_n) \vec{F}(\vec{x}_n).$$

Dado que calcular explícitamente la matriz inversa del Jacobiano no es un proceso muy eficiente desde el punto de vista numérico, a la hora de implementar el método se hace en dos pasos:

1. se resuelve el sistema

$$J(\vec{x}_n) \Delta \vec{x}_{n+1} = -\vec{F}(\vec{x}_n),$$

2. se calcula

$$\vec{x}_{n+1} = \vec{x}_n + \Delta \vec{x}_{n+1}.$$

Ejemplo 3.4 Suponemos que se quiere buscar una solución del sistema

$$\begin{aligned} f_1 &= x^2 - 10x + y^2 + 8 = 0 \\ f_2 &= xy^2 + x - 10y + 8 = 0 \end{aligned}$$

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x - 10 & 2y \\ y^2 + 1 & 2xy - 10 \end{bmatrix}$$

y se parte de $(x, y) = (0, 0)$ se obtiene la sucesión:

Tabla 3.4.- Resultado de la iteración del método de Newton para el sistema

i	x	y
1	0	0
2	0.8	0.88
3	0.9918	0.9917
4	1	1
5	1	1

como se observa la convergencia con este método es, en general, más rápida que con el método del punto fijo.

Las siguientes funciones muestran una posible implementación en Matlab del método de Newton.

```
function y=f(x)
% funcion para utilizar con newtonsi.m
y(1)=x(1)^2-10*x(1)+x(2)^2+8;
y(2)=x(1)*x(2)^2+x(1)-10*x(2)+8;

function df=jac(x)
% matriz jacobiana para usar con newtonsi.m
df(1,1)=2*x(1)-10;
df(1,2)=2*x(2);
df(2,1)=x(2)^2+1;
df(2,2)=2*x(1)*x(2)-10;

function [xr,k]=newtonsi(x,tol,imax)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Metodo de Newton para sistemas de ecuaciones
% Uso: [xr,k]=newtonsi(x,tol,imax)
%
% Input:
% x = vector x1,x2,...,xn inicial,
% tol=tolerancia
%
% Se ha de disponer de las funciones:
%   f.m funcion y=f(x) donde se define el sistema
%   jac.m funcion df=jac(x) donde se define la matriz
%   derivada del sistema.
```

```

%
% Output: xr= raiz, k= numero de iteraciones.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
k=1;
epi=1;
x1=x;
while norm(epi)>tol
    x=x1;
    fxn=f(x);
    axn=jac(x);
    epi=axn\fxn';
    x1=x-epi';
    k=k+1;
    if k>imax
        disp('no converge')
        break
    end
end
xr=x1;

```

3.2.3. Método de Broyden

El método de Broyden generaliza el método de la secante para resolver ecuaciones no lineales par el caso de un sistema de ecuaciones.

Para el caso unidimensional de tomaba una aproximación de la derivada de $f(x)$

$$s_n = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}},$$

y se hacían iteraciones de la forma

$$x_{n+1} = x_n - s_n^{-1} f(x_n).$$

Para el caso de un sistema, se ha de obtener una aproximación de la matriz derivada, S_n , que cumpla

$$S_n (x_n - x_{n-1}) = f(x_n) - f(x_{n-1}).$$

Introducimos la siguiente notación

$$\begin{aligned} b_n &= x_{n+1} - x_n, \\ \Delta f_n &= f(x_{n+1}) - f(x_n), \\ S_{n+1} &= S_n + C_n. \end{aligned}$$

Con lo que tenemos que

$$S_{n+1} (x_{n+1} - x_n) = f(x_{n+1}) - f(x_n),$$

o sea,

$$(S_n + C_n) b_n = \Delta f_n,$$

y, por tanto,

$$C_n b_n = \Delta f_n - S_n b_n.$$

De este modo, dado un vector w_n^T tal que $w_n^T b_n \neq 0$, podemos elegir

$$C_n = \frac{1}{w_n^T b_n} (\Delta f_n - S_n b_n) w_n^T,$$

ya que

$$C_n b_n = \Delta f_n - S_n b_n.$$

Si elegimos $w_n^T = b_n$, se obtiene el primer método de Broyden, y si elegimos $w_n = S_n^T b_n$ se obtiene el segundo método de Broyden.

Así pues, el método de Broyden para sistemas parte de una aproximación inicial para la raíz, x_0 , y una aproximación para la matriz derivada del sistema, S_0 , y se basa en iteraciones de la forma

$$\begin{aligned} b_n &= -S_n^{-1} f(x_n), \\ x_{n+1} &= x_n + b_n, \\ \Delta f_n &= f(x_{n+1}) - f(x_n), \\ S_{n+1} &= S_n + \frac{1}{w_n^T b_n} (\Delta f_n - S_n b_n) w_n^T. \end{aligned}$$

3.3. Ejercicios

1. Supongamos que se tiene un objeto cayendo verticalmente a través del aire sujeto a una resistencia viscosa y a la fuerza de gravedad. Si

suponemos que la altura inicial a la que se encuentra es S_0 , la altura del objeto al cabo de t segundos es

$$S(t) = S_0 + \frac{mg}{k}t - \frac{m^2g}{k^2} \left(1 - e^{-\frac{kt}{m}}\right),$$

donde $g = -9,81\text{m/s}^2$ y la constante k representa el coeficiente de resistencia del aire. Si tomamos $S_0 = 300\text{m}$, $m = 0,25\text{kg}$, y $k = 0,1\text{kg s m}^{-1}$. Estima el tiempo que tardará el objeto en caer al suelo.

2. Una partícula parte del reposo y se desliza a lo largo de un plano inclinado cuyo ángulo de inclinación, θ , cambia con velocidad constante

$$\frac{d\theta}{dt} = \omega < 0.$$

Después de t segundos la posición de la partícula viene dada por

$$x(t) = \frac{g}{2\omega^2} \left(\frac{e^{\omega t} - e^{-\omega t}}{2} - \text{sen}(\omega t) \right).$$

Supongamos que la partícula ha recorrido 1.7m en un segundo. Obtén una estimación de la velocidad ω a la que θ cambia. Emplea el valor $g = -9,8\text{m/s}^2$.

3. Una catenaria es la curva que forma un cable colgante. Supongamos que el punto más bajo coincide con el origen de coordenadas. Para este caso, la función que define la catenaria es de la forma

$$y = C \cosh\left(\frac{x}{C}\right) - C.$$

Para determinar la catenaria que pasa por los puntos $(\pm a, b)$, se debe obtener C de la ecuación

$$b = C \cosh\left(\frac{a}{C}\right) - C.$$

Obtened la catenaria que pasa por $(\pm 10, 6)$.

4. Verifica que la ecuación

$$x^4 + 2x^2 - x - 3 = 0,$$

se puede escribir como

$$\begin{aligned} \text{a) } x &= (3 + x - 2x^2)^{\frac{1}{4}} & \text{c) } x &= \left(\frac{x+3}{x^2+2}\right)^{\frac{1}{2}}, \\ \text{b) } x &= \left(\frac{x+3-x^4}{2}\right)^{\frac{1}{2}}, \\ \text{d) } x &= \frac{3x^4+2x^2+3}{4x^3+4x-1}. \end{aligned}$$

Calcula 5 iteraciones del método del punto fijo para cada uno de los casos, tomando $x_0 = 1 + a$, donde a es un número aleatorio entre 0 y 1. ¿Qué ocurre si para el caso b) se toma $x_0 = 2 + a$? ¿Qué ocurre si para el caso c) se toma $x_0 = -3 - a$. ¿Qué ocurre si para el caso d) se toma $x_0 = 0,2367$?

5. Dado el sistema de ecuaciones

$$\begin{aligned} x_1 &= \sin(x_1 + x_2), \\ x_2 &= \cos(x_1 - x_2), \end{aligned}$$

tomando $(x_1^0, x_2^0) = (1, 1)$. Calcula dos iteraciones del método del punto fijo, y dos iteraciones del método de Newton.

6. Dado el sistema de ecuaciones

$$\begin{aligned} x_1 &= (x_1^2 + x_2^2), \\ x_2 &= (x_1 - x_2), \end{aligned}$$

tomando $(x_1^0, x_2^0) = (1, 1)$. Calcula dos iteraciones del primer método de Broyden.

Capítulo 4

Interpolación

4.1. Introducción

Frecuentemente, al realizar trabajos experimentales, la información respecto de cierta función desconocida o de difícil cálculo, $f(x)$, se tiene a partir de una tabla con valores numéricos de $f(x)$ para ciertos valores de x . Para una función representada de esta forma se necesita, a menudo, hallar el valor de $f(x)$ para un valor de x entre dos de los valores tabulados. Éste es el problema de la *interpolación directa*. De forma recíproca, podemos necesitar obtener el valor de x a partir de un valor de $f(x)$ situado entre dos valores tabulados. Éste es el problema de la *interpolación inversa*. El problema de determinar el valor de $f(x)$ en algún punto x que no esté situado entre dos valores de la tabla se denomina *problema de extrapolación*.

Trabajar sólo con la información de $f(x)$ que proporciona la tabla significa desconocer “casi todo” sobre $f(x)$, en particular si $f(x)$ es continua y derivable. En el desarrollo del tema se asumirá la hipótesis de continuidad y derivabilidad sobre la función $f(x)$. Una justificación para esto es que intentaremos resolver los problemas de interpolación y extrapolación aproximando la función “desconocida”, $f(x)$, mediante una función polinómica o una función construida a partir de polinomios (splines).

La interpolación, con las hipótesis descritas es un procedimiento usualmente satisfactorio, sobre todo si los valores para los que se quiere calcular $f(x)$ están cerca de los valores tabulados.

Veamos a continuación algunos problemas prácticos que se pueden plantear.

4.1.1. Varios problemas prácticos

- La fertilización nitrogenada se ha realizado durante muchos años de forma indiscriminada debido en parte al bajo coste de los fertilizantes químicos y a la creencia de que una mayor fertilización implica una mayor producción. El aumento del coste, la evidencia de un óptimo en la dosis de N para una producción máxima y la constatación de que la contaminación por nitrato aumenta con la dosis de abonado nitrogenado, deberían conducir a una práctica más racional del abonado nitrogenado.

Los datos proporcionados corresponden a una experiencia con maíz y muestran la relación entre fertilización (kg Nfert/ha), el contenido de nitrógeno en suelo y en la planta (kg N/ha), y la producción (tm/ha).

Tabla 4.1.-

Nfertilizante aplicado (kg N/ha)	N en planta (Kg/ha)	N en suelo (Kg/ha)	Producción (t/ha)
0	20	15	1.2
110	120	20	6.5
230	220	110	9.3
350	230	230	9.0
460	235	340	8.5
570	210	560	8.0

Por otra parte las dosis óptimas de N dependen del manejo del agua. En la experiencia anterior también se midió la lixiviación (lavado hacia capas más profundas del suelo) de nitrato en función de la dosis de fertilizante nitrogenado y del volumen de drenaje:

Tabla 4.2.-

Nfertilizante aplicado (kg N/ha)	Lixiviación de nitrato (kg N/ha)		
	drenaje alto	drenaje medio	drenaje alto
0	38	25	18
95	42	30	19
180	60	40	21
360	138	95	50

- En la siguiente tabla se muestra el cálculo de la extracción diaria de nitrógeno en el cultivo de cítricos a partir de los datos conocidos de las necesidades mensuales

Tabla 4.3.-

Meses	Nextraído kgN/ha kgN/ha	extraccione mensuales/totales	Fracción total de N extraído	Fracción del año
1	2.44	0.012	0.012	0.083
2	4.07	0.02	0.032	0.166
3	6.11	0.03	0.062	0.25
4	12.22	0.06	0.122	0.333
5	23.21	0.114	0.236	0.416
6	34.61	0.17	0.406	0.5
7	42.75	0.21	0.616	0.583
8	36.65	0.18	0.796	0.666
9	22.39	0.11	0.906	0.75
10	10.18	0.05	0.956	0.833
11	6.11	0.03	0.986	0.916
12	2.85	0.014	1.000	1.000

- Veamos ahora la curva de demanda de N por el cultivo como una función de la fracción del tiempo.

Debido a los problemas que supone la ingestión de aguas con un elevado contenido en nitratos por parte del ser humano (la OMS propone que en agua potable no se sobrepase los 50 ppm), y puesto que parece que en la contaminación de los acuíferos tiene una importancia crucial el exceso de la fertilización nitrogenada aportada por la agricultura intensiva, se realizó un ensayo de laboratorio con vistas a considerar la evolución de mineralización, proceso por el cual el N orgánico (proveniente de residuos vegetales, abonos ...) se transforma en N mineral (NO_3^- y NH_4^+), que se produce en el suelo de una parcela de Vinalesa con vistas a considerar estos aportes en el balance total de N y así al considerar este aporte disminuir la fertilización disminuyendo de esta forma la contaminación de los acuíferos. Los datos obtenidos para un ensayo a 25 grados C y a la humedad determinada fueron los siguientes:

Tabla 4.4.-

Tiempo (días)	Nmin (mg/kg)
7	9.466
14	8.211
27	15.590
41	17.615
55	20.215
83	21.734

Según la bibliografía la evolución de la mineralización se ha ajustado a funciones lineales ($N_{min} = N_0 + kt$), exponenciales de primer orden ($N_{min} = N_0(1 - \exp(-kt))$), potenciales ($N_{min} = ht^k$).

4.2. Interpolación lineal

Supongamos que se tiene una tabla de datos

Tabla 4.5.-

x_0	x_1	x_{n+1}
y_0	y_1	y_{n+1}

que gráficamente se pueden representar, por ejemplo, de la forma

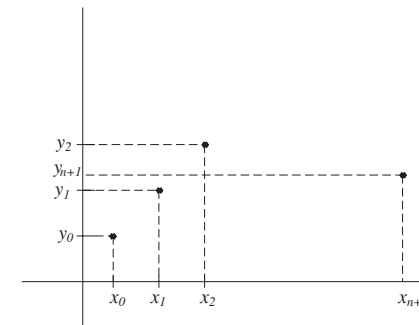


Fig 4.1.- Gráfica de los datos de una tabla.

Pretendemos conocer el valor que toma la variable y para un valor de x que no está recogido en la tabla. Por ejemplo, sea $x \in [x_i, x_{i+1}]$. Para aproximar el valor de y se supone que se encuentra en la recta que une los puntos (x_i, y_i) y (x_{i+1}, y_{i+1}) . Esta recta tiene por ecuación

$$\frac{x - x_i}{x_{i+1} - x_i} = \frac{y - y_i}{y_{i+1} - y_i},$$

o sea, el valor de y correspondiente al valor x es

$$y = y_i + \frac{y_{i+1} - y_i}{x_{i+1} - x_i}(x - x_i) = y_i \frac{x_{i+1} - x}{x_{i+1} - x_i} + y_{i+1} \frac{x - x_i}{x_{i+1} - x_i}. \quad (4.1)$$

Podemos escribir una función interpoladora, $I(x)$, que sea válida para todos los puntos comprendidos entre los puntos extremos de la tabla haciendo uso de la función característica,

$$\chi_{[x_i, x_{i+1}]} = \begin{cases} 1 & x \in [x_i, x_{i+1}] \\ 0 & x \notin [x_i, x_{i+1}] \end{cases}.$$

De este modo, para los puntos de la tabla 1.5 se puede escribir

$$\begin{aligned} I(x) &= \left(y_0 \frac{x_1 - x}{x_1 - x_0} + y_1 \frac{x - x_0}{x_1 - x_0} \right) \chi_{[x_0, x_1]} + \\ &+ \left(y_1 \frac{x_2 - x}{x_2 - x_1} + y_2 \frac{x - x_1}{x_2 - x_1} \right) \chi_{[x_1, x_2]} + \\ &\vdots \\ &+ \left(y_n \frac{x_{n+1} - x}{x_{n+1} - x_n} + y_{n+1} \frac{x - x_n}{x_{n+1} - x_n} \right) \chi_{[x_n, x_{n+1}]} . \end{aligned}$$

Introduciendo las funciones

$$\psi_i(x) = \frac{x - x_{i-1}}{x_i - x_{i-1}} \chi_{[x_{i-1}, x_i]} + \frac{x_{i+1} - x}{x_{i+1} - x_i} \chi_{[x_i, x_{i+1}]},$$

podemos escribir

$$I(x) = y_0 \frac{x_1 - x}{x_1 - x_0} \chi_{[x_0, x_1]} + \sum_{i=1}^n y_i \psi_i(x) + y_{n+1} \frac{x - x_n}{x_{n+1} - x_n} \chi_{[x_n, x_{n+1}]}, \quad (4.2)$$

que es el interpolante lineal asociado a la tabla 4.5.

4.3. Polinomio interpolador de Lagrange

Supongamos conocidos los valores de una función $f(x)$ en $n + 1$ puntos distintos x_0, x_1, \dots, x_n , que supondremos ordenados de menor a mayor. Esta información la representamos en la siguiente tabla:

Tabla 4.6.- Tabla de simple entrada para la función f .

x	x_0	x_1	\dots	x_n
$f(x)$	f_0	f_1	\dots	f_n

Nuestro problema es obtener, aproximadamente, el valor de $f(x)$ en un punto arbitrario $x \in [x_0, x_n]$. Para resolverlo, construiremos un polinomio $L_n(x)$ de grado menor o igual que n que cumpla que

$$L_n(x_i) = f_i, \quad i = 0, 1, \dots, n,$$

es decir, que en los puntos x_i toma los valores de la tabla. A $L_n(x)$ se le denomina *polinomio interpolador* de la función $f(x)$ en los puntos x_0, x_1, \dots, x_n que, a su vez, se llaman *nodos de interpolación* de la función f . Una posible forma de resolver el problema sería el plantear un polinomio de grado n

$$P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n,$$

con coeficientes a_i , $i = 0, 1, \dots, n$ indeterminados tal que $P_n(x_i) = f_i$, $i = 0, 1, \dots, n$. Esto significa que obtener el polinomio interpolador es equivalente a resolver el sistema de ecuaciones

$$\begin{aligned} a_0 + a_1x_1 + \dots + a_nx_1^n &= f_1, \\ a_0 + a_1x_2 + \dots + a_nx_2^n &= f_2, \\ &\vdots \\ a_0 + a_1x_n + \dots + a_nx_n^n &= f_n. \end{aligned}$$

Esta forma de afrontar el problema es, desde el punto de vista práctico, poco operativa.

El siguiente resultado proporciona una forma explícita del polinomio interpolador buscado.

Teorema 4.1 (Polinomio interpolador de Lagrange) *Dada la tabla 4.6, se considera el polinomio*

$$L_n(x) = \sum_{i=0}^n P_{n,i}(x) f_i$$

donde

$$P_{n,i}(x) = \frac{(x-x_0)(x-x_1)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_0)(x_i-x_1)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)},$$

con $i = 0, 1, \dots, n$. Este polinomio es un polinomio interpolador para la función $f(x)$.

Demostración.

En primer lugar, el polinomio $L_n(x)$ tiene grado menor o igual que n puesto que es combinación lineal de los polinomios $P_{n,i}(x)$, y éstos tienen grado menor o igual que n . Por otra parte, observemos que

$$P_{n,i}(x_k) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x_k - x_j}{x_i - x_j} = \begin{cases} 1 & \text{si } k = i \\ 0 & \text{si } k \neq i \end{cases},$$

de donde

$$L_n(x_i) = P_{n,i}(x_i) f_i = f_i, \quad i = 0, 1, \dots, n.$$

El polinomio obtenido en el teorema 4.1 se conoce con el nombre de polinomio interpolador de Lagrange. Sería más propio decir que está representado en la forma de Lagrange ya que, como probaremos en la siguiente proposición, el polinomio interpolador de una función $f(x)$ es único.

Proposición 4.1 Dada la tabla 4.6, el polinomio interpolador de la función $f(x)$ es único.

Demostración.

Supongamos que existe otro polinomio interpolador de $f(x)$, es decir, existe un polinomio $L_n^*(x)$ de grado menor o igual que n , tal que

$$L_n^*(x_i) = f_i, \quad i = 0, 1, \dots, n.$$

Consideramos el polinomio $Q_n(x) = L_n(x) - L_n^*(x)$ el cual tendrá, a lo sumo, grado n . Este polinomio verifica

$$Q_n(x_i) = L_n(x_i) - L_n^*(x_i) = 0, \quad i = 0, 1, \dots, n.$$

Por lo tanto $Q_n(x)$ es un polinomio de grado menor o igual que n con $n + 1$ raíces distintas, de donde $Q_n(x)$ es el polinomio idénticamente nulo. En consecuencia $L_n(x) = L_n^*(x)$.

Nota 4.1

- Hay otras formas de escribir el polinomio interpolador de una función $f(x)$ en $n + 1$ puntos. No obstante, hay que recordar que la única diferencia será de aspecto puesto que el polinomio interpolador es único.
- Si se sabe que la función $f(x)$ es un polinomio de grado menor o igual que n , $L_n(x) = f(x)$.
- Para el cálculo práctico del polinomio interpolador de Lagrange no hay necesidad alguna de que los nodos estén ordenados.
- En vista de que un polinomio interpolador depende linealmente de los valores f_i de la función f , el polinomio interpolador de la suma de dos funciones en $n + 1$ puntos x_0, x_1, \dots, x_n es igual a la suma de los polinomios interpoladores.

Ejemplo 4.1 Consideremos los datos dados por la tabla

Tabla 4.7.-

x	0,3	1	2	3
$f(x)$	1,35	2,718	7,389	20,086

que corresponde a algunos valores que toma la función $f(x) = e^x$. El polinomio interpolador, en la forma de Lagrange, viene dado por

$$\begin{aligned} L_3(x) &= \frac{(x-1)(x-2)(x-3)}{(0,3-1)(0,3-2)(0,3-3)} 1,35 + \\ &+ \frac{(x-0,3)(x-2)(x-3)}{(1-0,3)(1-2)(1-3)} 2,718 + \\ &+ \frac{(x-0,3)(x-1)(x-3)}{(2-0,3)(2-1)(2-3)} 7,389 + \\ &+ \frac{(x-0,3)(x-1)(x-2)}{(3-0,3)(3-1)(3-2)} 20,086. \end{aligned}$$

Para obtener, por ejemplo, una aproximación al valor de $f(0,44)$, calculamos $L_3(0,44) = 1,608$. El valor exacto de $f(x)$ es 1,553. Observamos las gráficas, superpuestas, de e^x y $L_3(x)$ en la figura 1.1.

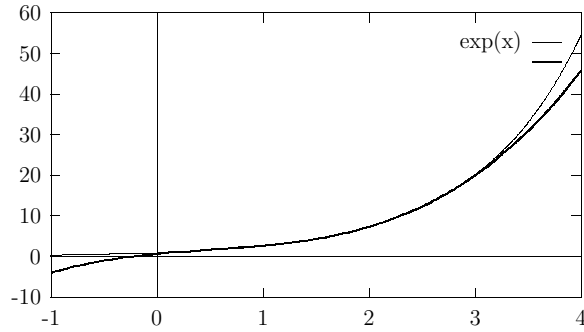


Fig. 4.2.- Gráficas de e^x y $L_3(x)$.

En el siguiente ejemplo se puede observar que el polinomio interpolador no siempre proporciona buenos resultados.

Ejemplo 4.2 Consideramos la tabla

x	0	3	5	9
$f(x)$	0	0,141	-0,959	0,412

correspondiente a algunos valores de la función $f(x) = \text{sen}(x)$. Si calculamos el polinomio interpolador en la forma de Lagrange, obtenemos

$$L_3(x) = \frac{(x-3)(x-5)(x-9)}{(0-3)(0-5)(0-9)}0 + \frac{(x-0)(x-5)(x-9)}{(3-0)(3-5)(3-9)}0,141 + \\ + \frac{(x-0)(x-3)(x-9)}{(5-0)(5-3)(5-9)}(-0,959) + \frac{(x-0)(x-3)(x-5)}{(9-0)(9-3)(9-5)}0,412.$$

De la observación de las gráficas de $\text{sen}(x)$ y de $L_3(x)$ (véase la figura 1.2) se deducen las grandes diferencias entre la función aproximadora ($L_3(x)$) y la función a aproximar ($\text{sen}(x)$).

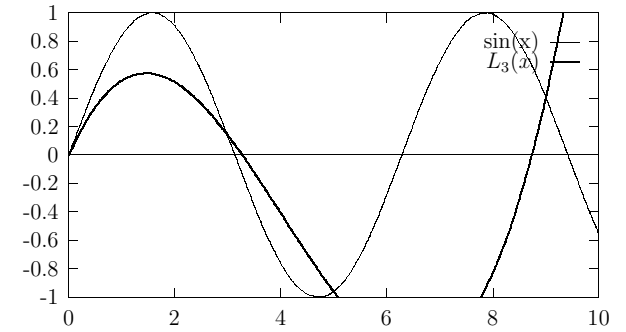


Fig. 4.3.- Gráficas de $\text{sen}(x)$ y $L_3(x)$.

4.4. Polinomio interpolador de Newton

A partir de los datos de la Tabla 4.5, vamos a obtener una nueva forma de expresar el polinomio interpolador para la función $f(x)$. Antes será interesante demostrar el siguiente resultado.

Proposición 4.2 Sea f una función de la cual se conocen sus valores en $n + 2$ valores reales distintos

$$f(x_i) = f_i, \quad i = 0, 1, \dots, n + 1.$$

Consideremos los polinomios interpoladores de f en los nodos x_0, x_1, \dots, x_n y $x_0, x_1, \dots, x_n, x_{n+1}$ a los cuales llamamos $L_n(x)$ y $L_{n+1}(x)$ respectivamente. Entonces, existe una constante $c_n \neq 0$ tal que

$$L_{n+1}(x) = L_n(x) + c_n(x - x_0)(x - x_1) \cdots (x - x_n).$$

Demostración. Observemos que el polinomio

$$L_n(x) + c_n(x - x_0)(x - x_1) \cdots (x - x_n),$$

coincide con $L_n(x)$, y por lo tanto con $f(x)$ en los puntos x_0, x_1, \dots, x_n , para cualquier valor de c_n . Además, este polinomio es de grado $n + 1$ a lo sumo.

Elegiremos c_n de forma que

$$f_{n+1} = L_n(x_{n+1}) + c_n(x_{n+1} - x_0)(x_{n+1} - x_1) \cdots (x_{n+1} - x_n),$$

de modo que

$$c_n = \frac{f_{n+1} - L_n(x_{n+1})}{(x_{n+1} - x_0)(x_{n+1} - x_1) \cdots (x_{n+1} - x_n)}.$$

Vamos a construir, de forma recursiva, un polinomio interpolador para la función f en los puntos x_0, x_1, \dots, x_n .

1. Definimos $N_0(x) = f_0$.
2. Teniendo en cuenta la demostración de la proposición 4.2 definimos

$$N_1(x) = N_0(x) + c_1(x - x_0), \text{ con } c_1 = \frac{f_1 - N_0(x_1)}{(x_1 - x_0)}.$$

3. En general, calculamos

$$N_k(x) = N_0(x) + \sum_{i=1}^k c_i(x - x_0)(x - x_1) \cdots (x - x_{i-1})$$

con

$$c_i = \frac{f_i - N_{i-1}(x_i)}{(x_i - x_0)(x_i - x_1) \cdots (x_i - x_{i-1})},$$

para $k = 2, 3, \dots, n$.

Llamaremos polinomio interpolador *en la forma de Newton* al polinomio

$$N_n(x) = N_0(x) + \sum_{i=1}^n c_i(x - x_0)(x - x_1) \cdots (x - x_{i-1}),$$

que se construye mediante el proceso recursivo descrito. Notar que

$$N_{k+1}(x) = N_k(x) + c_k(x - x_0)(x - x_1) \cdots (x - x_{k-1}).$$

Evidentemente $L_n(x) = N_n(x)$ aunque la forma de Newton permite aprovechar la información que se tiene al añadirse un nuevo nodo de interpolación.

4.5. Caso particular de nodos igualmente espaciados

4.5.1. Diferencias finitas

En muchos casos los nodos de interpolación pueden aparecer equidistantes entre sí o igualmente espaciados, es decir,

$$x_j = x_0 + jh, \quad j = 0, 1, \dots, n.$$

Sea $h \in \mathbb{R}$ constante y una función real $f(x)$. Llamaremos *diferencia de orden 1* de la función f , y se representa por $\Delta f(x)$, a la siguiente expresión

$$\Delta f(x) = f(x + h) - f(x)$$

Las diferencias de órdenes superiores se definen de forma recursiva:

$$\Delta^{k+1} f(x) = \Delta(\Delta^k f(x)), \quad k = 1, 2, \dots$$

Por convenio se considerará $\Delta^0 f(x) = f(x)$. Las diferencias sucesivas de una función $f(x)$ se pueden representar y calcular de forma sencilla mediante la siguiente tabla:

x	$f(x)$	$\Delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$	$\Delta^4 f(x)$
x_0	f_0	$\Delta f(x_0)$			
x_1	f_1	$\Delta f(x_1)$	$\Delta^2 f(x_0)$	$\Delta^3 f(x_0)$	
x_2	f_2	$\Delta f(x_2)$	$\Delta^2 f(x_1)$	$\Delta^3 f(x_1)$	$\Delta^4 f(x_0)$
x_3	f_3	$\Delta f(x_3)$	$\Delta^2 f(x_2)$	\vdots	\vdots
x_4	f_4	\vdots	\vdots	\vdots	\vdots
\dots	\dots	\dots	\dots	\dots	\dots

Ejemplo 4.3 Si calculamos la tabla de diferencias sucesivas correspondiente a la tabla de valores

x	0,5	1	1,5	2	2,5	3
$f(x)$	1,65	1,03	0,74	0,61	0,53	0,45

obtenemos

x_i	f_i	Δf	$\Delta^2 f$	$\Delta^3 f$	$\Delta^4 f$	$\Delta^5 f$
0,5	<u>1,65</u>					
1	1,03	<u>-0,62</u>				
1,5	0,74	-0,29	<u>0,33</u>			
2	0,61	-0,13	0,16	<u>-0,17</u>		
2,5	0,53	-0,08	0,05	-0,11	<u>0,06</u>	
3	0,45	-0,08	0,00	-0,05	0,06	<u>0,00</u>

4.5.2. Polinomio interpolador de Newton para puntos igualmente espaciados

Si en la forma de Newton del polinomio interpolador suponemos que los puntos están igualmente espaciados, es decir

$$x_j = x_0 + j h, \quad j = 0, 1, \dots, n.$$

se obtiene

$$c_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{f(x_0 + h) - f(x_0)}{(x_0 + h) - x_0} = \frac{\Delta f(x_0)}{h}$$

$$c_2 = \frac{f(x_2) - f(x_0) - c_1(x_2 - x_0)}{(x_2 - x_0)(x_2 - x_1)} =$$

$$= \frac{f(x_2) - f(x_0) - \frac{\Delta f(x_0)}{h}(x_2 - x_0)}{(x_2 - x_0)(x_2 - x_1)} =$$

$$= \frac{\frac{f(x_2) - f(x_0)}{x_2 - x_0} - \frac{\Delta f(x_0)}{h}}{x_2 - x_1} = \frac{\frac{f(x_2) - f(x_1) + f(x_1) - f(x_0)}{2h} - \frac{\Delta f(x_0)}{h}}{h} =$$

$$= \frac{\Delta f(x_1) + \Delta f(x_0) - 2\Delta f(x_0)}{2h^2} = \frac{\Delta f(x_1) - \Delta f(x_0)}{2h^2} = \frac{\Delta^2 f(x_0)}{2h^2}.$$

En general, se obtiene por inducción

$$c_k = \frac{\Delta^k f(x_0)}{h^k k!},$$

de donde la forma de Newton del polinomio interpolador quedaría del siguiente modo

$$N_n(x) = f(x_0) + \frac{\Delta f(x_0)}{h}(x - x_0) + \frac{\Delta^2 f(x_0)}{h^2 2!}(x - x_0)(x - x_1) + \dots$$

$$\dots + \frac{\Delta^n f(x_0)}{h^n n!}(x - x_0)(x - x_1) \dots (x - x_{n-1}).$$

Nótese que los coeficientes c_i , $i = 1, 2, \dots, n$ se obtienen, de forma sencilla, a partir de la tabla de diferencias sucesivas de la función.

Ejemplo 4.4 Dada la tabla de datos y los elementos subrayados de la tabla de diferencias ($\Delta^k f(x_0)$) del ejemplo 4.3, se construye muy fácilmente el polinomio interpolador en la forma de Newton

$$N_5(x) = 1,65 + \frac{(-0,62)}{0,5}(x - 0,5) + \frac{0,33}{0,5^2 2!}(x - 0,5)(x - 1) +$$

$$+ \frac{(-0,17)}{0,5^3 3!}(x - 0,5)(x - 1)(x - 1,5) +$$

$$+ \frac{0,06}{0,5^4 4!}(x - 0,5)(x - 1)(x - 1,5)(x - 2).$$

4.6. Interpolación inversa y extrapolación

4.6.1. Interpolación inversa

Dada una tabla de puntos, el problema de la interpolación inversa consiste, simplemente, en determinar para un valor dado de la función $f(x)$ la correspondiente x . Cualquiera de las formas de interpolación expuestas permite afrontar este problema, simplemente hay que considerar a la f como variable independiente y a x como variable dependiente. La interpolación inversa puede tener gran interés como método para contrastar los resultados obtenidos al resolver un problema de interpolación directa.

4.6.2. Extrapolación

La utilización de los polinomios interpoladores para calcular aproximaciones con valores no situados entre nodos de la tabla 1.5 da, en general, resultados bastante desastrosos. Solamente se obtienen buenos resultados en

la extrapolación para casos muy particulares de funciones. Por ejemplo, consideremos la tabla 1.5 donde f_i tiende a un límite f cuando i crece. Supongamos, además, que la sucesión $\{f - f_i\}$ es una progresión geométrica. Entonces, para tres valores consecutivos f_{k-1} , f_k , f_{k+1} , se tiene que

$$(f - f_{k-1})(f - f_{k+1}) = (f - f_k)^2,$$

y de aquí

$$f = f_{k+1} - \frac{(f_{k+1} - f_k)^2}{f_{k+1} - 2f_k + f_{k-1}}.$$

Esta fórmula se conoce como fórmula de extrapolación δ^2 de Aitken y más que como fórmula general de extrapolación, se utiliza para acelerar la convergencia en ciertos métodos iterativos.

4.7. Splines

Podemos pensar que, dada una nube de puntos, basta calcular el polinomio interpolador para obtener una aproximación aceptable de la función. Pero cuando el número de puntos de la tabla a ajustar es elevado, este proceso es computacionalmente, costoso y el resultado obtenido no suele ser bueno.

Otro método que se utiliza es la aproximación mediante Splines. Aquí se sigue la filosofía de que la aproximación obtenida será una función que pase por todos los puntos dados por la tabla y que esta función estará construida a partir de polinomios de grado a elegir, pero no muy alto en la práctica. Supongamos que se quieren aproximar, a partir de los datos de la tabla 1, los valores de una función $f(x)$, donde se supondrá que los puntos $a = x_0 \leq x_1 \leq \dots \leq x_n = b$ definen una partición del intervalo $[a, b]$.

Definición 4.1 Dado un intervalo $[a, b] \subset \mathbb{R}$ y una partición \mathcal{P} de este intervalo

$$a = x_0 \leq x_1 \leq \dots \leq x_n = b ,$$

llamaremos spline de orden k asociado a la partición, a la función $S(x)$ definida del siguiente modo

$$S(x) = S_i(x) \quad \forall x \in [x_{i-1}, x_i], \quad i = 1, \dots, n$$

donde $S_i(x)$ son polinomios de grado k , que satisfacen:

$$\begin{aligned} S_i(x_{i-1}) &= f_{i-1} \\ S_i(x_i) &= f_i \end{aligned} \quad i = 1, \dots, n \quad (4.3)$$

y además,

$$\begin{cases} S'_i(x_i) = S'_{i+1}(x_i) \\ \vdots \\ S_i^{k-1}(x_i) = S_{i+1}^{k-1}(x_i) \end{cases} \quad i = 1, \dots, n-1. \quad (4.4)$$

A k se le llama grado del spline.

Los splines más utilizados en la práctica, son los splines de grado 3 o cúbicos, o sea, un conjunto de polinomios

$$S_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \quad i = 1, \dots, n ,$$

que cumplen las condiciones anteriores. Veamos cómo podemos determinar estos polinomios.

Las condiciones (4.3) y (4.4), para los splines cúbicos, quedan de la forma:

$$\begin{aligned} S_i(x_{i-1}) &= f_{i-1} \\ S_i(x_i) &= f_i \end{aligned} \quad i = 1, \dots, n ,$$

$$\begin{aligned} S'_i(x_i) &= S'_{i+1}(x_i) \\ S_i(x_i) &= S_{i+1}(x_i) \end{aligned} \quad i = 1, \dots, n-1 ,$$

Hemos de determinar $4n$ coeficientes y se tienen $2n + 2(n-1) = 4n - 2$ condiciones. Harán falta, pues, dos condiciones extra que se suelen obtener imponiendo que $S''_1(x_0) = A$ y $S''_n(x_n) = B$, donde A y B son dos números reales. (En el caso que $A = B = 0$, el spline obtenido se llama spline natural).

Para obtener los polinomios, partimos de que $S''_i(x)$ son polinomios de grado 1 que cumplen $S''_i(x_i) = S''_{i+1}(x_i)$, $i = 1, \dots, n-1$. Así

$$S''_i(x) = M_{i-1} \frac{(x_i - x)}{(x_i - x_{i-1})} + M_i \frac{(x - x_{i-1})}{(x_i - x_{i-1})}, \quad i = 1, \dots, n. \quad (4.5)$$

Efectivamente

$$\begin{aligned} S''_i(x_i) &= M_i \\ S''_{i+1}(x_i) &= M_i \frac{(x_{i+1} - x_i)}{(x_{i+1} - x_i)} + M_{i+1} \frac{(x_i - x_i)}{(x_{i+1} - x_i)} = M_i. \end{aligned} \quad (4.6)$$

Tendremos que calcular sólo M_1, \dots, M_{n-1} ya que la expresión (4.5) nos permite afirmar que

$$A = S''_1(x_0) = M_0 ,$$

y la expresión (4.6)

$$B = S''_n(x_n) = M_n .$$

Integrando la expresión (4.5) obtenemos

$$S_i(x) = M_{i-1} \frac{(x_i - x)^3}{6(x_i - x_{i-1})} + M_i \frac{(x - x_{i-1})^3}{6(x_i - x_{i-1})} + \alpha_i x + \beta_i ,$$

donde α_i y β_i son constantes de integración que se determinan imponiendo

$$\begin{aligned} S_i(x_{i-1}) &= f_{i-1} \\ S_i(x_i) &= f_i \quad i = 1, \dots, n , \end{aligned}$$

con lo que se obtiene:

$$\begin{aligned} \alpha_i &= \frac{f_i - f_{i-1}}{x_i - x_{i-1}} - \frac{1}{6}(M_i - M_{i-1})(x_i - x_{i-1}) \\ \beta_i &= f_i - \alpha_i x_i - \frac{1}{6}M_i(x_i - x_{i-1})^2 . \end{aligned}$$

Introduciendo la notación $h_i \equiv x_i - x_{i-1}$ podemos escribir:

$$\begin{aligned} S_i(x) &= \frac{M_{i-1}}{6h_i}(x_i - x)^3 + \frac{M_i}{6h_i}(x - x_{i-1})^3 + (f_i - \frac{1}{6}M_i h_i^2) \frac{(x - x_{i-1})}{h_i} + \\ &+ (f_{i-1} - \frac{1}{6}M_{i-1} h_i^2) \frac{(x_i - x)}{h_i} . \end{aligned} \quad (4.7)$$

Falta calcular los M_i , para ello, se hace uso de la condición que nos queda $S_i(x_i) = S_{i+1}(x_i)$ $i = 1, \dots, n - 1$ obteniendo el sistema de ecuaciones

$$a_i M_{i-1} + M_i + c_i M_{i+1} = d_i ,$$

donde los coeficientes vienen dados por las expresiones

$$\begin{aligned} a_i &= \frac{h_i}{2(h_i + h_{i+1})} ; \quad c_i = \frac{h_{i+1}}{2(h_i + h_{i+1})} , \\ d_i &= \frac{1}{2(h_i + h_{i+1})} \left\{ \frac{f_{i+1} - f_i}{h_{i+1}} - \frac{f_i - f_{i-1}}{h_i} \right\} . \end{aligned}$$

Este sistema es un sistema tridiagonal, cuya matriz de coeficientes es de la forma:

$$[A] = \begin{bmatrix} 1 & c_1 & 0 & \dots & 0 \\ a_2 & 1 & c_2 & \dots & 0 \\ 0 & \dots & \dots & \dots & \vdots \\ \vdots & & & & \vdots \\ 0 & \dots & & a_{n-1} & 1 \end{bmatrix} ,$$

y, una vez resuelto, basta sustituir en (4.7) para obtener los distintos polinomios del spline cúbico.

Ejemplo 4.5 Consideremos la tabla de valores

Tabla 4.7.-

x	0	1	2	3	4	5
$f(x)$	0	1	4	9	16	25

cuyos puntos representamos en la figura 4.4.

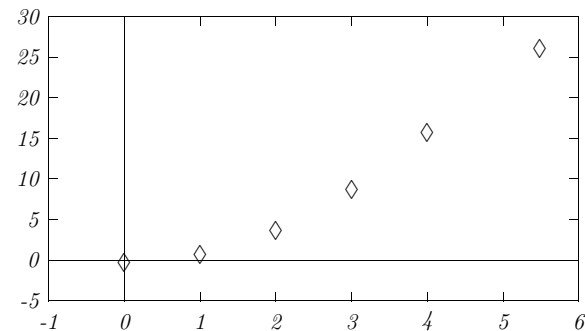


Fig. 4.4.- Gráfica de los puntos de la tabla 4.7.

Suponiendo que $M_0 = M_5 = 0$ y teniendo en cuenta que $h_j = 1$, $j = 1, \dots, 5$, calculamos los valores $a_i = 0,25$, $c_i = 0,25$, $d_i = 0,5$, $i = 1, \dots, 4$, que dan lugar a la matriz

$$[A] = \begin{bmatrix} 1 & c_1 & 0 & 0 \\ a_2 & 1 & c_2 & 0 \\ 0 & a_3 & 1 & c_3 \\ 0 & 0 & a_4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0,25 & 0 & 0 \\ 0,25 & 1 & 0,25 & 0 \\ 0 & 0,25 & 1 & 0,25 \\ 0 & 0 & 0,25 & 1 \end{bmatrix} .$$

Resolvemos el sistema

$$[A] \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix}$$

y obtenemos

$$\begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \end{bmatrix} = \begin{bmatrix} 0,421 \\ 0,316 \\ 0,316 \\ 0,421 \end{bmatrix}.$$

Sustituyendo en (4.7) obtenemos el spline

$$\begin{aligned} S_1(x) &= \frac{0,421}{6}x^3 + \left(1 - \frac{0,421}{6}\right)x, \\ S_2(x) &= \frac{0,421}{6}(2-x)^3 + \frac{0,316}{6}(x-1)^3 + \left(4 - \frac{0,316}{6}\right)(x-1) \\ &\quad + \left(1 - \frac{0,421}{6}\right)(2-x), \\ S_3(x) &= \frac{0,316}{6}(3-x)^3 + \frac{0,316}{6}(x-2)^3 + \left(9 - \frac{0,316}{6}\right)(x-2) \\ &\quad + \left(4 - \frac{0,316}{6}\right)(3-x), \\ S_4(x) &= \frac{0,316}{6}(4-x)^3 + \frac{0,421}{6}(x-3)^3 + \left(16 - \frac{0,421}{6}\right)(x-3) \\ &\quad + \left(9 - \frac{0,316}{6}\right)(4-x), \\ S_5(x) &= \frac{0,421}{6}(5-x)^3 + 25(x-4) + \left(16 - \frac{0,421}{6}\right)(5-x). \end{aligned}$$

En la figura 1.4, representamos los puntos de la tabla junto con la gráfica del spline obtenido.

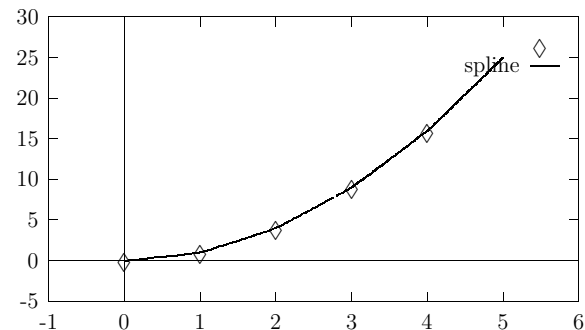


Fig. 1.4.- Puntos de la tabla 1.8 y el spline obtenido.

4.8. Ejercicios

- ¿Cuál es el polinomio de interpolación de la función $f(x) = 1$ en los puntos x_0, x_1, \dots, x_n ?
- Calcular un polinomio interpolador tal que en los puntos $x_0 = -1$, $x_1 = 0$ y $x_2 = 1$ tome los mismos valores que la función $f(x) = 3^x$.
- Escribir un polinomio de grado no mayor que 2 que tome los valores 1, 2, -1 en los puntos 0, 1, -2.
- Dada la tabla

x	0	1	2	3	4
$f(x)$	-1	2	13	44	107

determinar el polinomio interpolador de Newton asociado a estos datos.

- Dada la tabla

x	-2	1	2	4
$f(x)$	25	-8	-15	-23

determinar el polinomio interpolador de Lagrange asociado a estos datos, dar una estimación de $f(0)$.

6. Calcular el polinomio interpolador de Newton para la función tabulada

x	2	4	6	8
$f(x)$	3	11	27	50

Si sabemos a posteriori que $f(10) = 83$, determinar el nuevo polinomio interpolador de Newton.

7. Se tiene la siguiente tabla de datos

x	0	0.2	0.4	0.6	0.8	1.0
$\text{Si}(x)$	0	0.19956	0.39646	0.58813	0.77210	0.94608

donde $\text{Si}(x)$ es la función 'seno integral' dada por

$$\text{Si}(x) = \int_0^x \frac{\text{sen}(t)}{t} dt .$$

Calcular el polinomio interpolador correspondiente a los puntos de la tabla y obtener el valor de x tal que $\text{Si}(x) = 0,45$.

8. La presión y el volumen de una masa gaseosa están ligados por cierta función. Con los siguientes datos obtenidos experimentalmente

Presión	0.5	1.0	1.5	2.0	2.5
Volumen	1.65	1.03	0.74	0.61	0.53

- Determinar un polinomio interpolador de grado menor o igual que 4 para estos datos.
- Utilizar la información del apartado anterior para determinar un polinomio que tome el valor 0.45 para el volumen a una presión de 3.
- Calcular volumen aproximado para una presión de 0.6. Estudiar esta aproximación utilizando ambos polinomios.

9. La tabla adjunta muestra las puntuaciones de Álgebra y Física de 5 estudiantes

Álgebra	7.5	8	9.3	6.5	8.7
Física	8.2	7.8	8.6	7.2	9.1

- Hallar un polinomio de grado a lo sumo 4 que interpole estos valores.

- Calcular la nota que, aproximadamente, se espera que obtenga un alumno de Física si en Álgebra tiene un 7.

10. El número de bacterias por unidad de volumen presentes en un cultivo después de un número determinado de horas, viene dado por

Horas	0	2	4	6
Bacterias	32	65	132	275

- Hallar un polinomio interpolador de grado 3 a lo sumo para estos datos.
- A posteriori se analiza también el cultivo para diferentes tiempos, obteniendo

Horas	1	3	5
Bacterias	47	92	190

Calcular el polinomio interpolador para esta tabla, de grado menor o igual que 2.

- Tomando todos los datos, construir el polinomio interpolador de grado máximo.
- Comparar los distintos valores que se obtienen al considerar un periodo de tiempo igual a dos horas y media.

11. Construir un spline cúbico natural para aproximar $f(x) = \cos(\pi x)$ usando los valores dados por $f(x)$ en $x = 0, 0,25, 0,5, 0,75, 1,0$. Integrar el spline en el intervalo $[0, 1]$, y comparar con el resultado

$$\int_0^1 \cos(\pi x) dx = 0.$$

Usar la derivada del spline para aproximar $f'(0,5)$ y $f''(0,5)$.

12. Calcular el spline cúbico que aproxime a la función $f(x) = 3xe^x - e^{2x}$ en $x = 1,03$, utilizando los datos de la siguiente tabla

x	1.0	1.02	1.04	1.06
$f(x)$	0.76578939	0.79536678	0.82268817	0.84752226

Capítulo 5

Ajuste de curvas

En muchos problemas de tipo práctico, se dispone de gran cantidad de datos que se pretenden ajustar mediante una función con pocos parámetros. Para este tipo de problemas hay que desarrollar una estrategia distinta de la seguida en el problema de la interpolación.

Supongamos que se pretende ajustar los datos de una tabla de la forma

Tabla 5.1.- Tabla de datos.

x_0	x_1	\cdots	\cdots	x_n
y_0	y_1	\cdots	\cdots	y_n

mediante una función que depende de k parámetros

$$y = f(x, a_1, a_2, \dots, a_k) .$$

Así, por un lado tenemos los datos que proporciona la tabla

$$y_1, y_2, \dots, y_n , \quad (5.1)$$

y por otro los valores que predice la función

$$f(x_1), f(x_2), \dots, f(x_n) , \quad (5.2)$$

que se pretende que sean lo más cercanos posibles. Se pueden definir distintos errores que miden la distancia entre estos dos conjuntos de puntos. Así el *error máximo* se define como

$$E_\infty = \max_{0 \leq i \leq n} |f(x_i) - y_i| , \quad (5.3)$$

el error medio se define como

$$E_1 = \frac{1}{n+1} \sum_{i=0}^n |f(x_i) - y_i| , \quad (5.4)$$

y el error cuadrático medio

$$E_2 = \left[\frac{1}{n+1} \sum_{i=0}^n (f(x_i) - y_i)^2 \right]^{1/2} . \quad (5.5)$$

El problema del ajuste se puede plantear como el problema de buscar el valor de los parámetros a_1, a_2, \dots, a_k que hace mínima la distancia entre los conjuntos de puntos (5.1) y (5.2). Si la distancia entre los dos conjuntos se obtiene usando el error cuadrático medio, el problema del ajuste se denomina ajuste o aproximación de mínimos cuadrados. Veamos algunos ejemplos.

5.1. Recta de mínimos cuadrados

Supongamos que se quiere ajustar los datos de la tabla 2.1 mediante una recta de la forma

$$f(x, a, b) = ax + b .$$

El error cuadrático medio para este caso es de la forma

$$E_2 = \left[\frac{1}{n+1} \sum_{i=0}^n (y_i - ax_i - b)^2 \right]^{1/2} .$$

Para calcular a y b que hagan mínimo el error, se plantean las ecuaciones normales

$$\frac{\partial E}{\partial a} = \left(\frac{1}{n+1} \right)^{1/2} \frac{1}{2} \left[\sum_{i=0}^n (y_i - ax_i - b)^2 \right]^{-1/2} 2 \sum_{i=0}^n (y_i - ax_i - b) (-x_i) = 0 ,$$

$$\frac{\partial E}{\partial b} = \left(\frac{1}{n+1} \right)^{1/2} \frac{1}{2} \left[\sum_{i=0}^n (y_i - ax_i - b)^2 \right]^{-1/2} (-2) \sum_{i=0}^n (y_i - ax_i - b) = 0 ,$$

con lo que se obtiene el sistema de ecuaciones

$$\sum_{i=0}^n y_i x_i - a \sum_{i=0}^n x_i^2 - b \sum_{i=0}^n x_i = 0 ,$$

$$\sum_{i=0}^n y_i - a \sum_{i=0}^n x_i - b \sum_{i=0}^n 1 = 0 .$$

Utilizando la Regla de Cramer obtenemos las soluciones

$$a = \frac{\begin{vmatrix} \sum_{i=0}^n y_i x_i & \sum_{i=0}^n x_i \\ \sum_{i=0}^n y_i & \sum_{i=0}^n 1 \end{vmatrix}}{\begin{vmatrix} \sum_{i=0}^n x_i^2 & \sum_{i=0}^n x_i \\ \sum_{i=0}^n x_i & \sum_{i=0}^n 1 \end{vmatrix}} = \frac{(n+1) \sum_{i=0}^n y_i x_i - \left(\sum_{i=0}^n y_i \right) \left(\sum_{i=0}^n x_i \right)}{(n+1) \sum_{i=0}^n x_i^2 - \left(\sum_{i=0}^n x_i \right) \left(\sum_{i=0}^n x_i \right)} \quad (5.6)$$

$$b = \frac{\begin{vmatrix} \sum_{i=0}^n x_i^2 & \sum_{i=0}^n y_i x_i \\ \sum_{i=0}^n x_i & \sum_{i=0}^n y_i \end{vmatrix}}{\begin{vmatrix} \sum_{i=0}^n x_i^2 & \sum_{i=0}^n x_i \\ \sum_{i=0}^n x_i & \sum_{i=0}^n 1 \end{vmatrix}} = \frac{\left(\sum_{i=0}^n x_i^2 \right) \left(\sum_{i=0}^n y_i \right) - \left(\sum_{i=0}^n x_i \right) \left(\sum_{i=0}^n y_i x_i \right)}{(n+1) \sum_{i=0}^n x_i^2 - \left(\sum_{i=0}^n x_i \right) \left(\sum_{i=0}^n x_i \right)} \quad (5.7)$$

Veamos un ejemplo. Dada la tabla de datos siguiente:

Tabla 5.2.- Tabla de datos.

x_i	1	2	3	4	5	6	7	8	9	10
y_i	1.3	3.5	4.2	5.0	7.0	8.8	10.1	12.5	13.0	15.6

Trataremos de ajustar la nube puntos mediante una recta usando la aproximación de mínimos cuadrados. Basta para ello, utilizar las expresiones (5.6)

y (5.7) para calcular los parámetros de la recta, obteniendo

$$a = 1,538, \quad b = -0,360.$$

En la figura 5.1 se muestra la nube de puntos y la recta ajustada.

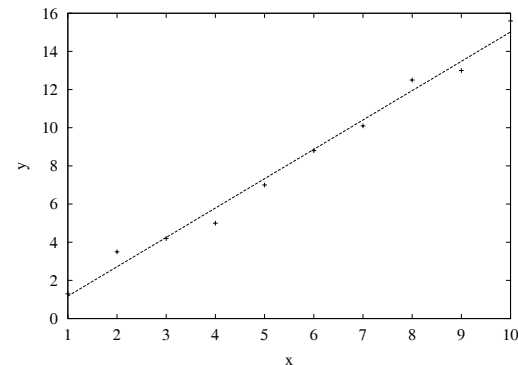


Fig. 5.1.- Nube de puntos asociada a la tabla 2.2 y su recta de mínimos cuadrados.

Hay distintas funciones que se pueden reducir a una recta mediante un cambio de variable y, por tanto, se pueden usar para ajustar los datos de una tabla siguiendo una metodología similar a la que se ha expuesto. Como ejemplo, se muestran las siguientes funciones:

$$y = a \frac{1}{x} + b \quad \rightarrow \quad y = at + b, \quad t = \frac{1}{x},$$

$$y = \frac{d}{x+c} \quad \rightarrow \quad x = d \frac{1}{y} - c,$$

$$y = \frac{x}{ax+b} \quad \rightarrow \quad \frac{x}{y} = ax + b,$$

$$y = ce^{ax} \quad \rightarrow \quad \ln(y) = \ln(c) + ax,$$

$$y = cx^a \quad \rightarrow \quad \ln(y) = \ln(c) + a \ln(x),$$

$$y = \frac{1}{(ax+b)^2} \quad \rightarrow \quad \frac{1}{\sqrt{y}} = ax + b,$$

5.2. Parábola de mínimos cuadrados

Supongamos ahora que se quiere ajustar los datos de la tabla 2.1 mediante una parábola de la forma

$$f(x, a, b) = a_0 + a_1x + a_2x^2 . \quad (5.8)$$

El error cuadrático medio para este caso es de la forma

$$E_2 = \left[\frac{1}{n+1} \sum_{i=0}^n (y_i - a_0 - a_1x_i - a_2x_i^2)^2 \right]^{1/2} .$$

Los valores de los parámetros a_0 , a_1 y a_2 se obtienen resolviendo las ecuaciones normales

$$\frac{\partial E_2}{\partial a_0} = 0 , \quad \frac{\partial E_2}{\partial a_1} = 0 , \quad \frac{\partial E_2}{\partial a_2} = 0 .$$

que dan lugar al sistema de ecuaciones

$$\begin{aligned} \sum_{i=0}^n y_i &= a_0 \sum_{i=0}^n 1 + a_1 \sum_{i=0}^n x_i + a_2 \sum_{i=0}^n x_i^2 , \\ \sum_{i=0}^n x_i y_i &= a_0 \sum_{i=0}^n x_i + a_1 \sum_{i=0}^n x_i^2 + a_2 \sum_{i=0}^n x_i^3 , \\ \sum_{i=0}^n y_i x_i^2 &= a_0 \sum_{i=0}^n x_i^2 + a_1 \sum_{i=0}^n x_i^3 + a_2 \sum_{i=0}^n x_i^4 , \end{aligned} \quad (5.9)$$

La solución del sistema (5.9) nos proporciona los parámetros de la parábola de mínimos cuadrados.

Consideremos los datos de la tabla siguiente

Tabla 5.3.- Tabla de datos.

x_i	0	0.25	0.5	0.75	1.0
y_i	1.0000	1.2840	1.6487	2.1170	2.7183

Ajustaremos los puntos mediante una parábola. Planteando las ecuaciones (5.9) y resolviéndolas, obtenemos

$$a_0 = 1,0052, \quad a_1 = 0,8641, \quad a_2 = 0,8437.$$

En la figura 5.2 se muestra la nube de puntos y la parábola ajustada.

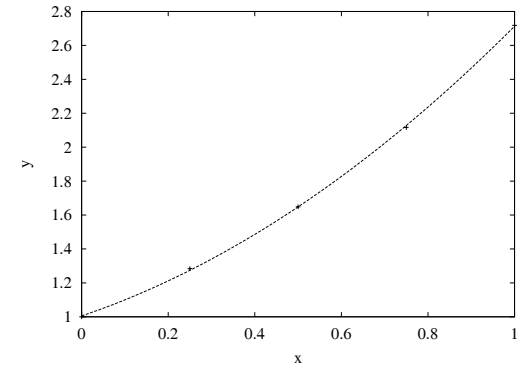


Fig. 5.2.- Nube de puntos asociada a la tabla 2.3 y su parábola de mínimos cuadrados.

5.3. Ajuste no lineal

Veremos algunos ejemplos donde al aplicar la metodología del ajuste por mínimos cuadrados se obtienen ecuaciones no lineales, que en general no tienen una solución analítica. No obstante va a ser posible estimar los valores de los parámetros de las funciones utilizando métodos numéricos que estudiaremos en el próximo capítulo.

Hay muchos procesos que se pueden modelizar mediante una ecuación de la forma

$$y(x) = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x} . \quad (5.10)$$

Piénsese, por ejemplo, en la solución de una ecuación diferencial lineal de la forma

$$y'' + ay' + by = 0 ,$$

donde no se conozcan los parámetros a , b y c . Su solución general será de la forma (5.10) con

$$\lambda_{1,2} = \frac{-a \pm \sqrt{a^2 - 4b}}{2} .$$

Nos planteamos pues, ajustar una función de la forma (5.10) a los datos de la tabla 2.1. El error cuadrático medio en este caso es de la forma

$$E_2 = \left[\frac{1}{n+1} \sum_{i=0}^n (y_i - C_1 e^{\lambda_1 x_i} - C_2 e^{\lambda_2 x_i})^2 \right]^{1/2} .$$

Las ecuaciones normales

$$\begin{aligned} \frac{\partial E_2}{\partial C_1} &= \frac{\sum_{i=0}^n (y_i - C_1 e^{\lambda_1 x_i} - C_2 e^{\lambda_2 x_i}) (e^{\lambda_1 x_i})}{(n+1)^{1/2} \left[\sum_{i=0}^n (y_i - C_1 e^{\lambda_1 x_i} - C_2 e^{\lambda_2 x_i})^2 \right]^{1/2}} = 0 , \\ \frac{\partial E_2}{\partial C_2} &= \frac{\sum_{i=0}^n (y_i - C_1 e^{\lambda_1 x_i} - C_2 e^{\lambda_2 x_i}) (e^{\lambda_2 x_i})}{(n+1)^{1/2} \left[\sum_{i=0}^n (y_i - C_1 e^{\lambda_1 x_i} - C_2 e^{\lambda_2 x_i})^2 \right]^{1/2}} = 0 , \\ \frac{\partial E_2}{\partial \lambda_1} &= - \frac{\sum_{i=0}^n (y_i - C_1 e^{\lambda_1 x_i} - C_2 e^{\lambda_2 x_i}) (C_1 x_i e^{\lambda_1 x_i})}{(n+1)^{1/2} \left[\sum_{i=0}^n (y_i - C_1 e^{\lambda_1 x_i} - C_2 e^{\lambda_2 x_i})^2 \right]^{1/2}} = 0 , \\ \frac{\partial E_2}{\partial \lambda_2} &= - \frac{\sum_{i=0}^n (y_i - C_1 e^{\lambda_1 x_i} - C_2 e^{\lambda_2 x_i}) (C_2 x_i e^{\lambda_2 x_i})}{(n+1)^{1/2} \left[\sum_{i=0}^n (y_i - C_1 e^{\lambda_1 x_i} - C_2 e^{\lambda_2 x_i})^2 \right]^{1/2}} = 0 , \end{aligned}$$

o lo que es lo mismo

$$\begin{aligned} \sum_{i=0}^n y_i e^{\lambda_1 x_i} - C_1 \sum_{i=0}^n e^{2\lambda_1 x_i} - C_2 \sum_{i=0}^n e^{(\lambda_1 + \lambda_2) x_i} &= 0 , \\ \sum_{i=0}^n y_i e^{\lambda_2 x_i} - C_1 \sum_{i=0}^n e^{(\lambda_1 + \lambda_2) x_i} - C_2 \sum_{i=0}^n e^{2\lambda_2 x_i} &= 0 , \\ \sum_{i=0}^n y_i x_i C_1 e^{\lambda_1 x_i} - C_1^2 \sum_{i=0}^n e^{2\lambda_1 x_i} - C_1 C_2 \sum_{i=0}^n x_i e^{(\lambda_1 + \lambda_2) x_i} &= 0 , \\ \sum_{i=0}^n y_i x_i C_2 e^{\lambda_2 x_i} - C_1 C_2 \sum_{i=0}^n e^{(\lambda_1 + \lambda_2) x_i} - C_2^2 \sum_{i=0}^n x_i e^{2\lambda_2 x_i} &= 0 . \end{aligned}$$

que es un sistema de ecuaciones no lineales, cuya solución veremos cómo se puede abordar en el próximo capítulo.

Ejemplo 5.1 Se sabe que una tabla de datos de la forma

t_0	t_1	\cdots	\cdots	t_n
x_0	x_1	\cdots	\cdots	x_n

corresponde a la evolución del oscilador

$$\frac{d^2 x}{dt^2} = -\omega^2 x , \quad x(0) = \frac{\pi}{4} , \quad x'(0) = 0 .$$

Escribir la ecuación normal que permite obtener la frecuencia ω a partir de los datos de la tabla.

Solución.- La solución general de la ecuación del oscilador es de la forma

$$x = C_1 \cos(\omega t) + C_2 \sin(\omega t) .$$

Imponiendo que se satisfagan las condiciones iniciales se obtiene

$$C_1 = \frac{\pi}{4} , \quad C_2 = 0 .$$

Luego los datos de la tabla hay que ajustarlos a una función de la forma

$$x = \frac{\pi}{4} \cos(\omega t) .$$

El error cuadrático será

$$E_2^2 = \sum_{i=0}^n \left(x_i - \frac{\pi}{4} \cos(\omega t_i) \right)^2 .$$

La ecuación normal se obtiene derivando respecto de ω e igualando a cero,

$$\sum_{i=0}^n \left(y_i - \frac{\pi}{4} \cos(\omega t_i) \right) \sin(\omega t_i) t_i = 0 .$$

5.4. Ejercicios

1. Escribir el sistema análogo al (5.9) que se obtiene si en vez de la parábola (5.8), consideramos un polinomio de grado k .

2. Se tomaron los siguientes datos del coeficiente de atenuación en función del espesor de una muestra de taconite

Espesor (cm)	Coeficiente (db/cm)
0.040	26.5
0.041	28.1
0.055	25.2
0.056	26.0
0.062	24.0
0.071	25.0
0.078	27.2
0.082	25.6
0.090	25.0
0.092	26.8

Obtener la recta de mínimos cuadrados asociada a estos datos.

3. Se conoce que la relación existente entre el peso vivo de las larvas de la mariposa nocturna W (g) y el oxígeno consumido por la larva R en ml/h es aproximadamente de la forma

$$R = bW^a .$$

Obtener los valores de a y b a partir de los datos de la siguiente tabla

W	R
0.017	0.154
0.087	0.296
0.174	0.363
1.11	0.531
1.74	2.23
4.09	3.58
5.45	3.52
5.96	2.40

4. Bajo ciertas condiciones se conoce que la evolución de una población con el tiempo se puede modelizar mediante una ecuación logística de la forma

$$P(t) = \frac{1000}{1 + Ce^{At}} .$$

Obtener C y A para la siguiente tabla de datos

$P(t)$	200	400	650	850	950
t	0	1	2	3	4

5. Dada la tabla de mineralización del nitrógeno

Tiempo (días)	Nmin (mg/kg)
7	9.466
14	8.211
27	15.590
41	17.615
55	20.215
83	21.734

Obtener las ecuaciones normales para el modelo $Nmin = N0(1 - \exp(-kt))$.

Capítulo 6

Introducción a las ecuaciones diferenciales en derivadas parciales

6.1. Introducción

Muchos fenómenos se modelizan utilizando ecuaciones en derivadas parciales. Estas ecuaciones en pocas ocasiones admiten soluciones analíticas, por tanto, es interesante disponer de métodos aproximados para resolverlas. En este capítulo veremos algunos problemas asociados con algunos tipos de ecuaciones en derivadas parciales y daremos algunos métodos sencillos para su resolución.

Una de las ecuaciones más conocidas es la ecuación de Poisson, que para un sistema bidimensional se escribe de la forma

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = -f, \quad (x, y) \in D \subset \mathbb{R}^2. \quad (6.1)$$

Asociadas a esta ecuación se pueden tener diversas condiciones de contorno

$$\begin{aligned} u(x_1, x_2) &= \phi(x_1, x_2), & \text{Dirichlet,} \\ \vec{n} \nabla u &= 0, & \text{Neumann,} \\ \vec{n} \nabla u + \alpha u &= \gamma, & \text{Cauchy.} \end{aligned}$$

para (x_1, x_2) pertenecientes a la frontera de D , y \vec{n} es un vector normal a la frontera de D en cada punto.

Para modelizar procesos de difusión transitorios se suele utilizar la ecuación del calor

$$\frac{\partial u}{\partial t} = \alpha \left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} \right) + f. \quad (6.2)$$

Una ecuación más complicada que se utiliza para estudiar procesos en los que se tiene difusión y convección es la ecuación de convección difusión

$$\frac{\partial u}{\partial t} + \vec{b} \nabla u = \vec{\nabla} \cdot (a \vec{\nabla} u) + f. \quad (6.3)$$

Otros problemas interesantes son los que estudian la propagación de las ondas, que utilizan ecuaciones de ondas, que tienen la estructura

$$\frac{\partial^2 u}{\partial t^2} = v^2 \nabla^2 u. \quad (6.4)$$

A continuación veremos algunos métodos de diferencias finitas para resolver problemas asociados con estas ecuaciones.

6.2. Métodos de diferencias finitas

Para obtener un método aproximado para resolver un problema asociado a las ecuaciones en derivadas parciales se suele dividir el dominio donde se ha de resolver la ecuación en una serie de celdas o nodos. Para un problema unidimensional tendríamos

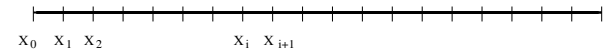


Figura 6.1.- Mallado para un problema 1D.

Los métodos en diferencias finitas se basan en aproximaciones locales de las derivadas parciales. Para obtener estas aproximaciones, se hace uso de la fórmula de Taylor, que para funciones de una variable se escribe como

$$u(x + \Delta x) = u(x) + \Delta x u'(x) + \frac{1}{2!} \Delta x^2 u'' + \frac{1}{3!} \Delta x^3 u''' + \dots, \quad (6.5)$$

de este modo, podemos tomar la aproximación

$$u'(x) \approx \frac{u(x + \Delta x) - u(x)}{\Delta x} + o(\Delta x), \quad (6.6)$$

donde $o(\Delta x)$ hace referencia a que los términos que no se han tenido en cuenta son al menos proporcionales a Δx .

Si escribimos,

$$u(x - \Delta x) = u(x) - \Delta x u'(x) + \frac{1}{2!} \Delta x^2 u'' - \frac{1}{3!} \Delta x^3 u''' + \dots, \quad (6.7)$$

obtenemos

$$u''(x) \approx \frac{u(x + \Delta x) - 2u(x) + u(x - \Delta x)}{\Delta x^2} + o(\Delta x^2). \quad (6.8)$$

Otra aproximación que se puede tomar es

$$\frac{d}{dx} \left(a(x) \frac{du}{dx} \right) (x_i) \approx \frac{a_{i+1/2} (u_{i+1} - u_i) - a_{i-1/2} (u_i - u_{i-1})}{\Delta x^2} + o(\Delta x^2), \quad (6.9)$$

donde $u_i = u(x_i)$, $a_{i+1/2}$ es el valor de a en el punto medio entre el nodo i y el nodo $i + 1$. Del mismo modo, $a_{i-1/2}$ es el valor de a en el punto medio entre el nodo i y el nodo $i - 1$.

Veamos algunos problemas que se pueden tratar con estas aproximaciones.

6.2.1. Ecuación de Poisson 1D

Supongamos que se tiene el problema

$$\begin{aligned} -u''(x) &= f(x), \quad x \in (0, 1), \\ u(0) &= u(1) = 0. \end{aligned}$$

Para resolverlo, discretizamos el intervalo $[0, 1]$ en subintervalos igualmente espaciados

$$x_i = i\Delta x, \quad i = 0, 1, \dots, n + 1.$$

Para los nodos interiores podemos usar la aproximación (6.8), para la derivada segunda obteniendo la relación

$$-u_{i-1} + 2u_i - u_{i+1} = -\Delta x^2 f_i.$$

Escribiendo esta ecuación para $i = 1, 2, \dots, n$, obtenemos un sistema de ecuaciones de la forma

$$Au = f,$$

donde

$$A = \begin{bmatrix} 2 & -1 & 0 & \dots \\ -1 & 2 & -1 & \dots \\ & & \dots & \\ & & \dots & -1 & 2 & -1 \\ & & & \dots & 0 & -1 & 2 \end{bmatrix}, \quad f = \begin{bmatrix} -\Delta x^2 f_1 \\ \vdots \\ -\Delta x^2 f_n \end{bmatrix}.$$

La matriz de este sistema es tridiagonal y se puede resolver de un modo eficiente utilizando el algoritmo de Thomas.

6.2.2. Ecuación de la difusión dependiente del tiempo

La ecuación de la difusión dependiente del tiempo que consideraremos tiene la forma

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2}. \quad (6.10)$$

Supondremos que se satisfacen unas condiciones de contorno en los extremos de dominio y que la distribución espacial de la u en el instante inicial es conocida.

El primer paso para obtener una aproximación numérica para resolver esta ecuación es discretizar el tiempo y el espacio en intervalos igualmente espaciados, $t = n\Delta t$, $n = 1, 2, \dots$, y $x = x_0 + i\Delta x$. Se toma una aproximación para la derivada temporal de la forma

$$\frac{\partial u}{\partial t} \approx \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} + o(\Delta t).$$

Para la derivada espacial se toma

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u(x - \Delta x, t) - 2u(x, t) + u(x + \Delta x, t)}{\Delta x^2} + o(\Delta x^2).$$

Se suele utilizar la notación $u(n\Delta t, x_0 + i\Delta x) = u_i^n$, y se escribe la aproximación de la ecuación (6.10) como

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = \alpha \frac{u_{i-1}^n - 2u_i^n + u_{i+1}^n}{\Delta x^2},$$

o sea,

$$u_i^{n+1} = u_i^n + r (u_{i-1}^n - 2u_i^n + u_{i+1}^n), \quad r = \frac{\alpha \Delta t}{\Delta x^2}.$$

El método obtenido es un método explícito, ya que los valores de u_i^{n+1} se pueden calcular directamente sabiendo los valores de u_i^n .

Para garantizar la estabilidad del esquema explícito, se puede ver que es necesario que se cumpla la condición

$$0 < \frac{\alpha \Delta t}{\Delta x^2} < 0,5 ,$$

que se conoce como *la condición de Courant*, y que limita la longitud del paso temporal que es necesario elegir una vez se ha elegido un paso espacial.

Para evitar problemas de estabilidad, se puede evaluar la derivada segunda espacial en el instante $(n+1)\Delta t$, en vez de hacerlo en el instante $n\Delta t$, obteniendo de este modo la aproximación

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = \alpha \frac{u_{i-1}^{n+1} - 2u_i^{n+1} + u_{i+1}^{n+1}}{\Delta x^2} ,$$

o sea,

$$-ru_{i-1}^{n+1} + (1+2r)u_i^{n+1} - ru_{i+1}^{n+1} = u_i^n ,$$

que es un método implícito, ya que si hacemos variar $i = 1, \dots, N_x$, para cada paso de tiempo, se ha de resolver un sistema de ecuaciones de la forma

$$\begin{bmatrix} (1+2r) & -r & & & & & & & \\ -r & (1+2r) & -r & & & & & & \\ & & \dots & & & & & & \\ & & & & -r & (1+2r) & -r & & \\ & & & & -r & (1+2r) & & & \end{bmatrix} \begin{bmatrix} u_1^{n+1} \\ \vdots \\ u_{N_x}^{n+1} \end{bmatrix} = \begin{bmatrix} u_1^n + ru_0^{n+1} \\ \vdots \\ u_{N_x}^n + ru_{N_x}^{n+1} \end{bmatrix} .$$

Otro método que se puede obtener que no tiene problemas de estabilidad y es más preciso que el método implícito, es el método de Crank-Nicolson, que viene dado por la ecuación

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = \frac{\alpha}{2} \left(\frac{u_{i-1}^n - 2u_i^n + u_{i+1}^n}{\Delta x^2} + \frac{u_{i-1}^{n+1} - 2u_i^{n+1} + u_{i+1}^{n+1}}{\Delta x^2} \right) .$$

6.2.3. La ecuación de ondas

Otra ecuación que aparece con frecuencia es la ecuación de ondas, que es de la forma

$$c^2 \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial t^2} . \quad (6.11)$$

Utilizando las aproximaciones

$$\frac{\partial^2 u}{\partial x^2}(x_i, t_n) \approx \frac{u_{i-1}^n - 2u_i^n + u_{i+1}^n}{\Delta x^2} ,$$

y

$$\frac{\partial^2 u}{\partial t^2}(x_i, t_n) \approx \frac{u_i^{n-1} - 2u_i^n + u_i^{n+1}}{\Delta t^2} ,$$

se llega a una aproximación de la ecuación (6.11) de la forma

$$c^2 \frac{u_{i-1}^n - 2u_i^n + u_{i+1}^n}{\Delta x^2} = \frac{u_i^{n-1} - 2u_i^n + u_i^{n+1}}{\Delta t^2} ,$$

que se puede reescribir como

$$u_i^{n+1} = 2u_i^n \left(1 - \frac{c^2 \Delta t^2}{\Delta x^2} \right) + (u_{i-1}^n + u_{i+1}^n) \left(\frac{c^2 \Delta t^2}{\Delta x^2} \right) - u_i^{n-1} .$$

6.2.4. Ecuación de convección-difusión

Supongamos ahora, que se quiere resolver un problema estacionario de la forma

$$\begin{aligned} -au'' + bu' &= 0 , \quad 0 < x < L , \\ u(0) &= 0 , \quad u(L) = 1 . \end{aligned}$$

Para obtener la solución analítica de este problema introducimos $u' = y$ con lo que queda la ecuación

$$y' = \frac{b}{a} y ,$$

con lo que

$$y = K \exp\left(\frac{b}{a}x\right) ,$$

y, por tanto,

$$u = \tilde{K} \exp\left(\frac{b}{a}x\right) .$$

Imponiendo las condiciones de contorno se tiene

$$u(x) = \frac{1 - e^{RLx}}{1 - e^R} , \quad (6.12)$$

donde R es el número de Péclet, definido como

$$R = \frac{bL}{a} .$$

Para obtener una aproximación por diferencias finitas, primero obtenemos una aproximación $o(\Delta x^2)$ de $u'(x)$. Para ello se usan los desarrollos (6.5) y (6.7), legando a que

$$u'(x) \approx \frac{u(x + \Delta x) - u(x - \Delta x)}{2\Delta x} + o(\Delta x^2) . \quad (6.13)$$

Utilizando (6.8) y (6.13), se tiene la relación

$$b \frac{u_{i+1} - u_{i-1}}{2\Delta x} - a \frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2} = 0 , \quad (6.14)$$

que, definiendo

$$c = \frac{R\Delta x}{2} ,$$

se reescribe como

$$-(1-c)u_{i+1} + 2u_i - (1+c)u_{i-1} = 0 . \quad (6.15)$$

Esta es una ecuación en diferencias que se suele resolver probando una solución de la forma $u_i = r^i$,

$$-(1-c)r^{i+1} + 2r^i - (1+c)r^{i-1} = 0 ,$$

esto es,

$$r^{i-1} (-(1-c)r^2 + 2r - (1+c)) = 0 ,$$

luego r tendrá que cumplir que

$$(1-c)r^2 - 2r + (1+c) = 0 .$$

Las soluciones de esta ecuación son

$$r_1 = 1 , \quad r_2 = \frac{1+c}{1-c} .$$

La solución general de la (6.15) será de la forma

$$u_i = \alpha + \beta \left(\frac{1+c}{1-c} \right)^i ,$$

como $u_0 = 0$, $\beta = -\alpha$ y como $u_{n+1} = 1$, se cumplirá

$$\alpha = \frac{1}{1 - r_2^{n+1}} ,$$

así, la solución es de la forma

$$u_i = \frac{1 - r_2^i}{1 - r_2^{n+1}} .$$

Cuando $c > 1$, ($\Delta x > \frac{2}{R}$) entonces $r_2 < 0$, y la solución u_i oscila. Esto contrasta con el comportamiento de la solución analítica (6.12), que es creciente.

Para resolver este problema hay que utilizar esquemas de primer orden, que tengan en cuenta el signo de la velocidad b . Si $b > 0$, la derivada primera se aproxima como

$$u'_i \approx \frac{u_i - u_{i-1}}{\Delta x} ,$$

obteniendo una ecuación de la forma

$$b \frac{u_i - u_{i-1}}{\Delta x} - a \frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2} = 0 . \quad (6.16)$$

Si $b < 0$, la derivada primera se aproxima como

$$u'_i \approx \frac{u_{i+1} - u_i}{\Delta x} ,$$

obteniendo una ecuación de la forma

$$b \frac{u_{i+1} - u_i}{\Delta x} - a \frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2} = 0 . \quad (6.17)$$

Estas dos posibilidades se escriben de una forma compacta como

$$b \left(\frac{1}{2\Delta x} \left(1 - \frac{b}{|b|} \right) \frac{1}{\Delta x} (u_{i+1} - u_i) + \frac{1}{2\Delta x} \left(1 + \frac{b}{|b|} \right) \frac{1}{\Delta x} (u_i - u_{i-1}) \right) - a \frac{1}{\Delta x^2} (u_{i+1} - 2u_i + u_{i-1}) = 0 .$$

Este esquema se conoce como un esquema 'up-wind' de primer orden para la ecuación de convección-difusión.

6.2.5. Ecuación de Poisson 2D

Supongamos ahora, que se pretende resolver un problema de la forma

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -f, \quad (x, y) \in (0, l_1) \times (0, l_2), \quad (6.18)$$

$$u(x, y) = 0, \quad \text{para } x = 0; x = l_1; y = 0; x = l_2. \quad (6.19)$$

El primer paso que realizaremos consistirá en discretizar el rectángulo $[0, l_1] \times [0, l_2]$ mediante un conjunto de nodos igualmente espaciados, como se muestra en la figura 6.1.

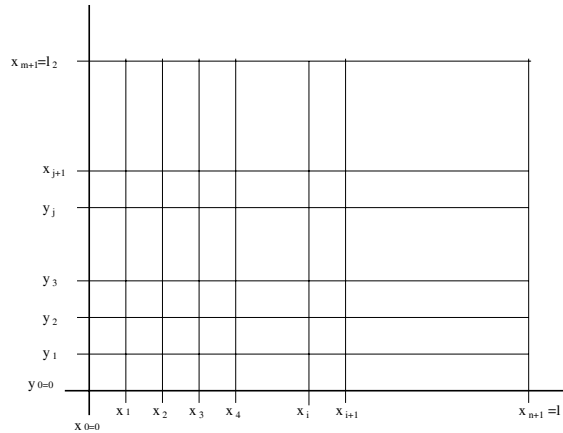


Figura 6.1.- Malla para un problema 2D rectangular.

$$x_i = i\Delta x, \quad i = 0, 1, 2, \dots, n + 1,$$

$$y_j = j\Delta y, \quad j = 0, 1, 2, \dots, m + 1.$$

Utilizando una aproximación similar a la de (6.8), para las derivadas segundas, se tiene que

$$\frac{\partial^2 u}{\partial x^2}(u_i, u_j) \approx \frac{u_{i-1j} - 2u_{ij} + u_{i+1j}}{\Delta x^2},$$

$$\frac{\partial^2 u}{\partial y^2}(u_i, u_j) \approx \frac{u_{i-1j} - 2u_{ij} + u_{i+1j}}{\Delta y^2},$$

donde $u_{ij} = u(x_i, y_j)$. De este modo, se tiene que la ecuación (6.18) se puede aproximar de la forma

$$\frac{1}{\Delta x^2}(u_{i-1j} - 2u_{ij} + u_{i+1j}) + \frac{1}{\Delta y^2}(u_{i-1j} - 2u_{ij} + u_{i+1j}) = -f_{ij}. \quad (6.20)$$

Para poder escribir las ecuaciones resultantes para $i = 1, \dots, n, j = 1, \dots, m$, se han de ordenar los nodos del malla de algún modo. Una posibilidad es utilizar el orden dado por

$$l = i + n(j - 1).$$

Cuando se escriben las ecuaciones para $i = 1, \dots, n, j = 1, \dots, m$, y se tienen en cuenta las condiciones de contorno (6.19), se obtiene un sistema de ecuaciones de la forma

$$Au = f,$$

donde A tiene una estructura en banda similar a la de la figura 6.2.

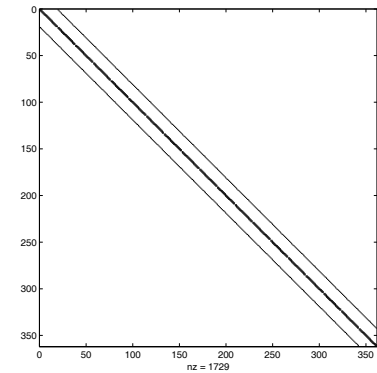


Figura 6.2.- Matriz asociada a la ecuación de Poisson.

6.3. Ejercicios

1. Dado el problema

$$y'' = 4(y - x); \quad 0 \leq x \leq 1, \quad y(0) = 0, \quad y(1) = 2,$$

que tiene por solución

$$y(x) = e^2 (e^4 - 1)^{-1} (e^{2x} - e^{-2x}) + x .$$

Obtén un esquema en diferencias finitas para calcular una solución aproximada del problema. Utiliza una malla con 5 nodos para el cálculo de la solución y compara el resultado con la solución exacta.

2. Sea u el potencial electrostático entre dos esferas concéntricas de metal, con radios R_1 y R_2 ($R_1 < R_2$), de manera que el potencial de la esfera interna se mantiene constante a V_1 V y el de la esfera externa es de 0 V. El potencial entre las dos esferas está regido por la ecuación

$$\frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r} = 0 .$$

Supón que $V_1 = 110$ V, $R_1 = 4$ cm y $R_2 = 8$ cm y obtén una aproximación de u para $r = 5$ cm y compáralo con la solución analítica

$$u(r) = \frac{V_1 R_1}{r} \left(\frac{R_2 - r}{R_2 - R_1} \right) .$$

3. Utilizando un mallado de 5×5 nodos y un método explícito, obtén una solución aproximada del problema

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} ,$$

con las condiciones

$$\begin{aligned} u(0, t) &= 0 , & u(1, t) &= 0 , & 0 \leq t \leq 1 . \\ u(x, 0) &= \text{sen}(\pi x) , & 0 \leq x \leq 1 , \\ \frac{\partial u}{\partial t}(x, 0) &= 0 , & 0 \leq x \leq 1 . \end{aligned}$$

4. Utilizando una malla 4×4 para el rectángulo $[0, 1] \times [0, 1]$. Obtén la solución aproximada de la ecuación de Poisson

$$\nabla^2 u = y , \quad (x, y) \in (0, 1) \times (0, 1) ,$$

con las condiciones

$$\begin{aligned} u(x, 0) &= x^3 , & u(x, 1) &= x^3 , & 0 \leq x \leq 1 . \\ u(0, y) &= 0 , & u(1, y) &= 1 , & 0 \leq y \leq 1 . \end{aligned}$$

Bibliografía

- [1] A. Aubanell, A. Benseny, A. Delshams, *Útiles de cálculo numérico*. Editorial Labor Universitaria, Barcelona, (1993).
- [2] R.L. Burden, J.D. Faires, *Análisis numérico*. Grupo Editorial Iberoamérica, (1985).
- [3] J.H. Mathews, *Numerical Methods, for mathematics, science and engineering*. 2nd Edition. Prentice Hall International Editions. (1992).
- [4] W.H. Press, S.A. Teukolski, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in Fortran 77. The art of scientific computing. Second edition*. Cambridge University Press, (1992).
- [5] Y. Saad, *Iterative methods for sparse linear systems. Second edition*. SIAM, (2003).
- [6] J.M. Sanz-Serna, *Diez lecciones de cálculo numérico*, Secretariado de Publicaciones e Intercambio Científico. Universidad de Valladolid, (1998).
- [7] J. Stoer, R. Burlich, *Introduction to Numerical Analysis. Second Edition*, Springer Verlag, (1993).