

$$\begin{aligned} f(x) &= 0 \text{ .} \\ x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \text{ .} \end{aligned}$$

Métodos numéricos

Problemas resueltos y prácticas

$$P_n(x) = \sum_{k=0}^n f(x_k) \prod_{\substack{j=0 \\ j \neq k}}^n \frac{x - x_j}{x_k - x_j} \text{ .}$$

$$\begin{aligned} \frac{dy}{dx} &= f(x, y) \text{ , } y(x_0) = y_0 \text{ .} \\ y(x_{i+1}) &\approx y_{i+1} = y_i + hf(x_i, y_i) \text{ .} \end{aligned}$$

Isaac A. García
Susanna Maza

$$\begin{aligned} f(x) &= 0 \text{ .} \\ x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \text{ .} \end{aligned}$$

$$\int_a^b f(x) \, dx \approx \sum_{i=0}^n \alpha_i f(x_i) \text{ .}$$

$$P_n(x) = \sum_{k=0}^n f(x_k) \prod_{\substack{j=0 \\ j \neq k}}^n \frac{x - x_j}{x_k - x_j} \text{ .}$$

$$\begin{aligned} \frac{dy}{dx} &= f(x, y) \text{ , } y(x_0) = y_0 \text{ .} \\ y(x_{i+1}) &\approx y_{i+1} = y_i + hf(x_i, y_i) \text{ .} \end{aligned}$$

$$\begin{aligned} \frac{dy}{dx} &= f(x, y) \text{ , } y(x_0) = y_0 \text{ .} \\ y(x_{i+1}) &\approx y_{i+1} = y_i + hf(x_i, y_i) \text{ .} \end{aligned}$$

$$\begin{aligned} f(x) &= 0 \text{ .} \\ x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \text{ .} \end{aligned}$$

eines 62

Métodos numéricos

Problemas resueltos y prácticas

Isaac A. García y Susanna Maza

Seminari de Sistemes Dinàmics
Departament de Matemàtica
Universitat de Lleida

ISBN: 978-84-8409-329-9

© Edicions de la Universitat de Lleida, 2009

© Els autors

Maquetación

Edicions i Publicacions de la UdL

Diseño de portada

cat & cas

La reproducción total o parcial de esta obra por cualquier procedimiento, incluidos la reprografía y el tratamiento informático, y la distribución de ejemplares mediante alquiler o préstamo público, queda rigurosamente prohibida sin la autorización de los titulares del “copy-right”, y será sometida a las sanciones establecidas por la Ley.
--

Índice general

1. Prólogo	11
2. Errores, Estabilidad y Condicionamiento	15
2.1. Resumen de teoría	15
2.2. Problemas resueltos	16
3. Sistemas Lineales	27
3.1. Resumen de teoría	27
3.1.1. Algoritmo de eliminación Gaussiana	27
3.2. Problemas resueltos	31
4. Interpolación Polinómica	43
4.1. Resumen de teoría	43
4.2. Problemas resueltos	49
5. Derivación Numérica	71
5.1. Resumen de teoría	71
5.2. Problemas resueltos	72
6. Integración Numérica	85
6.1. Resumen de teoría	85
6.2. Problemas resueltos	90
7. Ecuaciones Diferenciales Ordinarias	125
7.1. Resumen de teoría	125
7.2. Problemas resueltos	127
8. Solución de Ecuaciones no Lineales	153
8.1. Resumen de teoría	153
8.2. Problemas resueltos	154
9. Aproximación de Funciones	195
9.1. Resumen de teoría	195
9.2. Problemas resueltos	197

10.Introducción a la Programación	217
10.1. Velocidad del Sonido en el Aire en Función de la Temperatura . .	217
10.2. Ecuación de Segundo Grado	218
10.3. Método de Newton: Cálculo de Raíces Cuadradas	220
10.4. Algoritmo de Arquimedes y el Número π	222
10.5. El problema $3x + 1$	225
11.El Periodo del Péndulo Simple	227
11.1. Introducción	227
11.2. Algoritmo de Integración Gaussiana con 5 Nodos	229
11.3. Resultados Numéricos	231
12.Sistemas Lineales de Ecuaciones: Eliminación Gaussiana	233
12.1. Realización de la Práctica	233
12.1.1. Solución en FORTRAN	234
12.1.2. Solución en MATHEMATICA	236
12.2. Aplicación al Cálculo de Intensidades en Redes Eléctricas	238
13.La Ecuación del Calor Unidimensional	241
13.1. Introducción	241
13.2. Método de Crank-Nicholson	241
13.3. Separación de Variables: Series de Fourier	243
13.3.1. Búsqueda de las soluciones fundamentales $T_n(x, t)$ con condiciones de contorno homogéneas	243
13.3.2. Realización de la práctica	245
13.3.3. Representación gráfica de los resultados	246
13.3.4. Comparación con el método de Fourier	247
14.Interpolación de Newton	249
14.1. Introducción	249
14.2. Diferencias Divididas de Newton	250
14.3. Fenómeno Runge	251
14.4. Minimización del Error de Interpolación	252
15.Interpolación por Splines Cúbicos Naturales	255
15.1. Introducción	255
15.2. Realización de la Práctica	257
15.3. Solución en FORTRAN	258
15.4. Solución en MATHEMATICA	260
16.Derivación Numérica	263
16.1. Introducción	263
16.2. Realización de la Práctica	264
16.2.1. Apéndice: generación de números aleatorios	265

17.Integración Romberg	267
17.1. Introducción	267
17.2. Realización de la Práctica	268
17.3. La Capacidad Calorífica de los Sólidos	270
18.Ecuaciones Integrales: Transmisión de Calor	271
18.1. Introducción	271
18.2. Realización de la Práctica	272
18.3. Existencia de Solución para Ecuaciones de Fredholm	273
19.Aproximación Trigonométrica y Coeficientes de Fourier	275
19.1. Introducción	275
19.2. Realización de la Práctica	277
20.Distribución de Velocidades en un Fluido	279
20.1. Introducción	279
20.2. Integración Gauss-Chebychev	280
20.3. Realización de la Práctica	281
21.La Condensación de Bose-Einstein	283
21.1. Introducción	283
21.1.1. La física estadística cuántica	284
21.2. Un Gas de Bosones Ideal	285
21.3. Realización de la Práctica	286
21.3.1. Apéndice: Integrales de Bose	286
22.Ecuaciones Diferenciales Ordinarias: Método RK4	287
22.1. Introducción	287
22.2. El Método RK4 en \mathbb{R}^n	288
22.3. Ciclos Límite	289
22.4. El Problema del Centro-Foco	291
22.5. El Péndulo de Wilberforce	293
22.6. El Modelo de Lorenz: Convección de Fluidos	294
22.6.1. Las ecuaciones de Lorenz	295
22.6.2. Realización de la práctica	296
23.Raíces de Ecuaciones: Método de Newton-Raphson	299
23.1. Introducción	299
23.2. Ecuaciones de Estado de los Gases	300
23.2.1. Ecuación de Dieterici	301
23.2.2. El Módulo de elasticidad en la ecuación de van der Waals	301
23.3. Un Problema Sobre la Profundidad de Absorción	302
23.3.1. Realización de la práctica	302
23.4. Método de Newton para Sistemas de Ecuaciones	303
23.4.1. Equilibrio en sistema masa-muelle	305

24.Método de los Elementos Finitos	307
24.1. Introducción: Modelo Matemático	307
24.2. Solución Numérica	308
24.2.1. Mallado del dominio	308
24.2.2. Las ecuaciones de los elementos: método variacional . . .	308
24.2.3. Aproximación de la solución	310
24.2.4. Método variacional de Galerkin	310
24.2.5. Notas computacionales	311
24.2.6. Encadenamiento de los elementos	312
24.2.7. La solución del sistema encadenado	314
24.3. Desplazamientos Longitudinales Elásticos	314
24.4. Fluidos Viscosos en Canalización Cilíndrica	316
25.Problemas Variacionales con Fronteras Fijas	317
25.1. Introducción	317
25.2. Un método numérico	317
25.3. El problema de la braquistocrona	319
26.Introducción a MATHEMATICA	321
26.1. Introducción	321
26.2. Estructuras de datos: Listas	321
26.3. Funciones	322
26.4. Representación Gráfica	323
26.5. Matrices y Sistemas Lineales	325
26.6. El Lenguaje de Programación	325
26.6.1. Tipos de números	325
26.6.2. Entradas de datos	326
26.6.3. Asignación de variables	327
26.6.4. Iteraciones y bucles	327
26.6.5. Condicionales y operadores lógicos	328
26.6.6. Sentencias de control	328
26.7. Cálculo Numérico	328
26.7.1. Errores	328
26.7.2. Polinomio interpolador	330
26.7.3. Aproximación lineal por mínimos cuadrados	330
26.7.4. Integración	331
26.7.5. Raíces de ecuaciones	331
26.7.6. Ecuaciones diferenciales ordinarias	333
27.Introducción a OCTAVE	337
27.1. Introducción a Matlab–Octave	337
27.1.1. Algunas convenciones	337
27.1.2. Primeros cálculos: funciones elementales	338
27.1.3. Matrices y vectores	339
27.1.4. Funciones	342
27.1.5. Gráficos	344

27.1.6. Generación de números aleatorios	347
27.1.7. Polinomios	347
27.2. El Lenguaje de Programación	348
27.2.1. Iteraciones y bucles	348
27.2.2. Condicionales y operadores lógicos	350
27.2.3. Errores	351
27.2.4. Comparando la eficiencia de algoritmos	352
27.3. Métodos Numéricos	353
27.3.1. Polinomio interpolador	353
27.3.2. Splines Cúbicos	354
27.3.3. Integración Numérica	355
27.3.4. Ecuaciones Diferenciales Ordinarias	355
27.3.5. Raíces de Ecuaciones	357
27.3.6. Aproximación Polinomial por Mínimos Cuadrados	359

Capítulo 1

Prólogo

Los Métodos Numéricos son una herramienta fundamental en el campo de las ciencias aplicadas que tratan de diseñar métodos que aproximen, de forma eficiente, las soluciones de problemas previamente formulados matemáticamente. En la mayoría de los casos, el problema matemático se deriva de un problema práctico en áreas experimentales como la Física, Química, Ingeniería, Economía, etc... Sobre él se aplican, típicamente, dos tipos de estrategias generales:

1. Se dan por supuestas algunas hipótesis de carácter simplificador que permiten llegar a una formulación matemática resoluble.
2. Se prescinde de alguna de estas hipótesis para llegar a una formulación matemática más complicada, que no se puede resolver explícitamente, pero cuya solución puede calcularse de forma aproximada.

Aunque de ninguna de las dos formas anteriores obtenemos la solución del problema original, a menudo resulta más apropiado utilizar la segunda. Las dificultades que aparecen en la resolución explícita de los modelos hacen que no se puedan hallar efectivamente las soluciones, en otros casos que el proceso de solución precise de infinitas operaciones. Es entonces cuando los métodos numéricos proporcionan soluciones aproximadas a tales problemas.

En este libro se aplican algoritmos y métodos constructivos que nos permitan obtener la solución de un problema con una precisión arbitraria en un número finito de pasos. Se pretende poner al alcance de los estudiantes una introducción a los Métodos Numéricos con el objetivo de ir aprendiendo las técnicas que hace servir la Matemática Aplicada no sólo a nivel teórico, sino también de una forma práctica mediante la resolución de problemas concretos. Dichos problemas son una forma de aproximarse a situaciones reales relacionadas con la ingeniería.

El presente libro consta de dos partes. En primer lugar, se dan problemas resueltos y corresponde a los temas básicos de un primer curso de introducción a los Métodos Numéricos. El libro comprende los siguientes temas: Errores, Estabilidad y Condicionamiento, Sistemas Lineales, Interpolación Polinómica, Derivación Numérica, Integración Numérica, Ecuaciones Diferenciales Ordinarias, Solución de Ecuaciones no Lineales y Aproximación de Funciones. A parte

de un breve resumen de teoría en cada capítulo, el objetivo principal es que el alumno pueda seguir paso a paso la resolución de numerosos problemas de los temas mencionados como ejemplificación de los conceptos y resultados teóricos, complementando así los realizados en clase. El presente libro es un buen complemento del libro de teoría [4]. Se ha procurado presentar las soluciones en la forma más práctica y directa. Cabe destacar que, a medida que avanzan los capítulos, los problemas propuestos pueden tener preguntas de capítulos anteriores entrelazadas con las del tema actual.

En la segunda parte, las prácticas propuestas consisten en la confección de *programas* que realicen los métodos explicados en las clases teóricas para la resolución de algún problema en concreto. De esta forma, en dichas prácticas el alumno se ha de familiarizar con el ordenador. Un programa es una secuencia de instrucciones que se dan a la computadora para procesar *datos*. El resultado del procesamiento es la *salida de datos*, que puede ocurrir en varias formas, como por ejemplo una lista impresa o un gráfico. Las operaciones realizadas por una computadora pueden clasificarse en *cálculos aritméticos* y *decisiones lógicas* adoptadas sobre la base de los valores calculados, *lectura y escritura de datos*.

Uno de los lenguajes utilizados en las prácticas será el FORTRAN (el nombre que proviene de las primeras letras de las palabras inglesas FORMula y TRANslation) que es un lenguaje de traducción de fórmulas, ver por ejemplo [3]. El nombre sugiere la aplicación más importante de dicho lenguaje: calcular expresiones matemáticas. Es, por lo tanto, un lenguaje de programación muy usual en cálculo numérico. Se utilizará también el programa de libre distribución OCTAVE que es también muy adecuado para cálculos numéricos. Otra opción que se utiliza consiste en usar también manipuladores algebraicos (en concreto MATHEMATICA) que tanto se han aconsejado para otras asignaturas de matemáticas de carreras técnicas, de manera que en muchas de las prácticas propuestas se pide realizar programas que se podrían sustituir por la utilización de un único comando de tales manipuladores. Pero, sin embargo, creemos que debe de ser un objetivo, cuando se trata de asignaturas de carácter numérico, el utilizar un lenguaje de programación y profundizar más en la propia estructura de los métodos. Si hacemos esto es porque estamos convencidos, y la experiencia con nuestros alumnos así nos lo demuestra, de que sólo cuando uno se enfrenta a la programación efectiva de los métodos es capaz de entenderlos en profundidad. De cualquier forma, en el estudio de los métodos de aproximación puede ser sumamente útil contar con un paquete de cómputo simbólico. Por ejemplo, muchos métodos numéricos tienen límites de error que exigen acotar una derivada ordinaria o parcial de una función. A veces se trata de una tarea tediosa y de poco valor didáctico una vez dominados los métodos del cálculo. Estas derivadas pueden obtenerse rápidamente en forma simbólica, y un poco de ingenio permitirá que una computación simbólica facilite también el proceso de acotación. Hemos escogido MATHEMATICA para este uso debido a que los estudiantes ya lo han utilizado en prácticas realizadas en las asignaturas de Cálculo y Álgebra Lineal de primer curso. Una guía completa sobre las posibilidades de MATHEMATICA se encuentra en el manual del programa contenido en la página web

<http://documents.wolfram.com/v5/TheMathematicaBook/index.html>.

El principal objetivo de las prácticas es fijar conceptos y resultados teóricos mediante la realización de una serie de ejercicios con la ayuda del ordenador. Además, dichas prácticas pretenden ser una aproximación a problemas reales de la ingeniería en los cuales su solución pasa por la programación de algún método numérico. Nos gustaría que este libro de problemas facilitase el aprendizaje de la asignatura y, agradeceríamos cualquier sugerencia o comentario que pueda mejorarlo dirigiéndose a las direcciones electrónicas: `garcia@matematica.udl.cat` o bien `smaza@matematica.udl.cat`.

Isaac A. García y Susanna Maza, Lleida 2009.

Capítulo 2

Errores, Estabilidad y Condicionamiento

2.1. Resumen de teoría

Habitualmente se requiere el uso del ordenador en el tratamiento numérico de los problemas planteados debido a la gran extensión en los cálculos. De esta forma, al ejecutarse los algoritmos en un ordenador, las operaciones básicas de suma, multiplicación y división se realizan con un cierto error. Uno de los problemas más importantes cuando se realizan cálculos numéricos es conocer el grado de precisión que se puede obtener en su realización. La limitación en la precisión de los resultados es consecuencia del hecho que los cálculos se han realizado con un número finito de dígitos, por lo que los resultados de las operaciones van afectados por un *error*. De los errores hay que conocer su procedencia, magnitud y comportamiento en el transcurso de los cálculos. A la hora de diseñar un buen algoritmo numérico hay que tener en cuenta estas limitaciones. Hemos de intentar evitar, mediante una organización cuidadosa de las operaciones, la propagación y acumulación de los errores.

Sea x_i^* un valor aproximado del valor x_i . Denotamos el *error absoluto* por $\Delta_{x_i} := |x_i - x_i^*|$ y al *error relativo* de la forma $\epsilon_{x_i} := |\Delta_{x_i}/x_i|$ siempre que $x_i \neq 0$.

Sea $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ con $f \in C^1(D)$ tal que $y = f(x_1, x_2, \dots, x_n)$. Las fórmulas de propagación de errores son:

$$\textbf{Error Absoluto: } \Delta_y \approx \sum_{j=1}^n \left| \frac{\partial f(x)}{\partial x_j} \right| \Delta_{x_j}.$$

$$\textbf{Error Relativo: } \epsilon_y \approx \sum_{j=1}^n \left| \frac{x_j}{f(x)} \frac{\partial f(x)}{\partial x_j} \right| \epsilon_{x_j}.$$

Los factores de amplificación $|(x_j/f(x))\partial f(x)/\partial x_j|$ de los errores relativos son llamados *números de condición*. Si cualquier número de condición presente en un problema tiene una magnitud grande decimos que estamos ante un *prob-*

lema mal condicionado.

Se sabe que, con el ordenador, no podemos trabajar con los números reales, sino sólo con un subconjunto de los números racionales. Como existe un número representable por el ordenador máximo en valor absoluto, cuando el ordenador intenta utilizar uno de mayor valor absoluto se produce un error llamado *overflow*. También se tiene el caso contrario que consiste en utilizar un número no nulo pero menor en valor absoluto que el mínimo número representable en valor absoluto. Entonces el ordenador devuelve un error llamado *underflow*. Se dispone del concepto *epsilon de la máquina*, que es el mayor valor positivo ϵ que verifique

$$1 + x = 1 \quad \forall x \in (0, \epsilon) .$$

Tenemos pues que la precisión de la aritmética en coma flotante se caracteriza por ϵ y dependerá del ordenador y del tipo de variable.

2.2. Problemas resueltos

Problema 2.1 Es conocido que el periodo T de un péndulo simple de longitud ℓ viene dado por $T = 2\pi\sqrt{\ell/g}$, siendo g la gravedad. Se pretende medir la aceleración de la gravedad g usando un péndulo. Aproximar el error relativo cometido si las medidas efectuadas son $\ell = 1 \pm 0,001$ metros y $T = 2 \pm 0,01$ segundos.

Solución. La aceleración de la gravedad es

$$g(\ell, T) = 4\pi^2 \frac{\ell}{T^2} .$$

Según la fórmula de propagación de errores relativos se tiene

$$\epsilon_g \approx \left| \frac{\ell}{g(\ell, T)} \frac{\partial g}{\partial \ell} \right| \epsilon_\ell + \left| \frac{T}{g(\ell, T)} \frac{\partial g}{\partial T} \right| \epsilon_T .$$

Calculando las derivadas parciales

$$\frac{\partial g}{\partial \ell} = \frac{4\pi^2}{T^2} , \quad \frac{\partial g}{\partial T} = -\frac{8\pi^2 \ell}{T^3} ,$$

se obtiene

$$\epsilon_g \approx \epsilon_\ell + 2\epsilon_T .$$

Según los datos del enunciado

$$\epsilon_\ell = \frac{0,001}{1} = 0,001 , \quad \epsilon_T = \frac{0,01}{2} = 0,005 ,$$

de manera que

$$\epsilon_g \approx 0,001 + 2 \times 0,005 = 0,011 .$$

Problema 2.2 La velocidad límite v de una gota esférica de radio R compuesta de un fluido de viscosidad μ_1 y densidad ρ_1 desplazándose bajo la acción de la gravedad constante g en otro fluido de viscosidad $\mu_2 \neq 0$ y densidad ρ_2 viene dada por

$$v = \frac{2R^2g(\rho_1 - \rho_2)(\mu_1 + \mu_2)}{3\mu_2(2\mu_2 + 3\mu_1)}.$$

- (i) Si se conocen de manera exacta los siguientes datos en el sistema internacional $\rho_1 = 10$, $\rho_2 = 15$ y $\mu_2 = 0,3$, pero sin embargo $\mu_1 = 0,2 \pm 0,1$ y $R = 1,5 \pm 0,2$, aproximar el error absoluto con el que obtiene la velocidad v .
- (ii) Suponiendo que sólo se tienen errores en la medida de R , averiguar si el problema de evaluar v está mal condicionado para alguna relación entre las variables del problema.

Solución. (i) Puesto que las únicas variables del problema que contienen errores son μ_1 y R , la fórmula de propagación de errores absolutos permite aproximar el error absoluto Δ_v con el que obtiene la velocidad v de la forma

$$\Delta_v = \left| \frac{\partial v}{\partial \mu_1} \right| \Delta_{\mu_1} + \left| \frac{\partial v}{\partial R} \right| \Delta_R,$$

siendo Δ_{μ_1} y Δ_R los errores absolutos de las magnitudes μ_1 y R respectivamente. Calculando las anteriores derivadas parciales de v a partir de la expresión dada en el enunciado se tiene

$$\Delta_v = \left| \frac{-2gR^2(\rho_1 - \rho_2)}{3(3\mu_1 + 2\mu_2)^2} \right| \Delta_{\mu_1} + \left| \frac{4gR(\mu_1 + \mu_2)(\rho_1 - \rho_2)}{3\mu_2(3\mu_1 + 2\mu_2)} \right| \Delta_R.$$

Sustituyendo los datos del enunciado y tomando la gravedad como $g = 9,8$ se tiene

$$\begin{aligned} \Delta_v &= \left| \frac{-2 \times 9,8 \times 1,5^2(10 - 15)}{3(3 \times 0,2 + 2 \times 0,3)^2} \right| \times 0,1 \\ &\quad + \left| \frac{4 \times 9,8 \times 1,5(0,2 + 0,3)(10 - 15)}{3 \times 0,3(3 \times 0,2 + 2 \times 0,3)} \right| \times 0,2 \\ &= |51,0417| \times 0,1 + |-136,111| \times 0,2 = 32,3264. \end{aligned}$$

(ii) Suponiendo que sólo se tienen errores en la medida de R , el problema de evaluar v estará mal condicionado si existe alguna relación entre las variables del problema para la cual el número de condición κ_R asociado a la variable R sea de una magnitud considerable. Por definición de número de condición se tiene

$$\kappa_R := \left| \frac{R}{v} \frac{\partial v}{\partial R} \right| = \left| \frac{R}{\frac{2R^2g(\rho_1 - \rho_2)(\mu_1 + \mu_2)}{3\mu_2(2\mu_2 + 3\mu_1)}} \frac{4gR(\mu_1 + \mu_2)(\rho_1 - \rho_2)}{3\mu_2(3\mu_1 + 2\mu_2)} \right| = 2.$$

Por lo tanto κ_R es una constante independiente de los valores de las variables del problema y en consecuencia el problema de evaluar v teniendo únicamente errores en R está bien condicionado siempre. Esto es debido a que, por la fórmula de propagación de los errores relativos se tiene que, en este caso, el error relativo ϵ_v en la medida de la velocidad v vale aproximadamente $\epsilon_v \approx |\kappa_R|\epsilon_R = 2\epsilon_R$, siendo ϵ_R el error relativo asociado a R .

Problema 2.3 *Bajo ciertas hipótesis, la velocidad v del sonido en el aire en función de la temperatura T viene expresada por*

$$v(T) = 331\sqrt{1 + \frac{T}{273}} ,$$

donde T y v son medidos en $^{\circ}C$ y m/s respectivamente. Averiguar si existe algún rango de temperaturas para el cual la evaluación de la función v está mal condicionada.

Solución. Se sabe, de la fórmula de propagación de errores relativos, que los errores relativos ϵ_T y ϵ_v en la temperatura y en la velocidad ϵ_T y ϵ_v están relacionados por

$$\epsilon_v \approx \kappa_T \epsilon_T ,$$

siendo κ_T el número de condición asociado a la variable T y definido como

$$\kappa_T := \left| \frac{T}{v(T)} \frac{dv}{dT} \right| .$$

Puesto que

$$v(T) = 331\sqrt{1 + \frac{T}{273}} ,$$

y su derivada es

$$\frac{dv}{dT} = \frac{331}{273} \frac{1}{2\sqrt{1 + \frac{T}{273}}} ,$$

se tiene que

$$\kappa_T = \left| \frac{T}{2(273 + T)} \right| .$$

En definitiva, pequeños errores relativos en la temperatura ϵ_T pueden provocar grandes errores relativos en la velocidad ϵ_v si $\kappa_T \gg 1$, es decir, si la temperatura está en un rango de valores tal que $T \approx -273$. Por supuesto, por motivos termodinámicos, se debe detener $T \gtrsim -273^{\circ}C$.

Problema 2.4 Sean x e y la solución del sistema lineal

$$\begin{aligned} 2x + ay &= 1, \\ x + by &= 0, \end{aligned}$$

con $2b - a \neq 0$.

- (i) Si $a = 1 \pm 0,1$ y $b = 2 \pm 0,2$, aproximar con qué exactitud se puede calcular $x + y$.
- (ii) Obtener el rango de valores de a y de b para los cuales el problema de hallar $x + y$ está mal condicionado.

Solución. (i) La solución del sistema lineal es

$$x = \frac{b}{2b - a}, \quad y = \frac{-1}{2b - a},$$

de manera que, definimos su suma como la función

$$f(a, b) = x + y = \frac{b - 1}{2b - a}.$$

Aplicando la fórmula de propagación de errores absolutos, se tiene que el error absoluto Δ_f al evaluar la suma $x + y$ está relacionado con los errores absolutos de a y de b mediante

$$\Delta_f \approx \left| \frac{\partial f}{\partial a} \right| \Delta_a + \left| \frac{\partial f}{\partial b} \right| \Delta_b.$$

Puesto que las derivadas parciales de f de primer orden son

$$\frac{\partial f}{\partial a} = \frac{b - 1}{(a - 2b)^2}, \quad \frac{\partial f}{\partial b} = \frac{2 - a}{(a - 2b)^2},$$

obtenemos

$$\Delta_f \approx \left| \frac{b - 1}{(a - 2b)^2} \right| \Delta_a + \left| \frac{2 - a}{(a - 2b)^2} \right| \Delta_b.$$

Tomando, según los datos del problema, $a = 1$, $\Delta_a = 0,1$, $b = 2$ y $\Delta_b = 0,2$ se obtiene

$$\Delta_f \approx 0,03333.$$

(ii) Para obtener el rango de valores de a y de b para los cuales el problema de hallar $x + y$, es decir hallar $f(a, b)$, está mal condicionado, calcularemos los dos números de condición κ_a y κ_b asociados a la función f y posteriormente estudiaremos para qué valores de a y de b dichos números de condición son grandes.

Por definición de número de condición se tiene que

$$\kappa_a = \left| \frac{a}{f(a, b)} \frac{\partial f}{\partial a} \right|, \quad \kappa_b = \left| \frac{b}{f(a, b)} \frac{\partial f}{\partial b} \right|,$$

por lo tanto

$$\kappa_a = \left| \frac{a}{2b-a} \right|, \quad \kappa_b = \left| \frac{b(a-2)}{(a-2b)(b-1)} \right|.$$

De aquí se deduce que, si $a - 2b \approx 0$ o bien $b - 1 \approx 0$ entonces el problema de hallar $x + y$ está mal condicionado.

Problema 2.5 *Estudiar el condicionamiento de la suma de n números reales. ¿Es bueno desde el punto de vista numérico restar dos números reales muy próximos a la vista del resultado anterior?*

Solución. Sean x_i , con $i = 1, \dots, n$, los números que se desean sumar. Hemos de estudiar pues el condicionamiento de la función $S(x_1, \dots, x_n) = \sum_{i=1}^n x_i$. Puesto que $S : \mathbb{R}^n \rightarrow \mathbb{R}$ es de clase $C^1(\mathbb{R}^n)$ se tiene, por la fórmula de propagación de errores, que el error relativo en la suma ϵ_S viene dado aproximadamente por

$$\epsilon_S \approx \sum_{j=1}^n \kappa_{x_j} \epsilon_{x_j},$$

siendo ϵ_{x_j} y κ_{x_j} el error relativo y el número de condición asociado a la magnitud x_j . Además se sabe que

$$\kappa_{x_j} = \left| \frac{x_j}{S} \frac{\partial S}{\partial x_j} \right|.$$

En nuestro caso, $\partial S / \partial x_j = 1$ para todo $j = 1, \dots, n$. De este modo se obtienen los números de condición

$$\kappa_{x_j} = \left| \frac{x_j}{\sum_{i=1}^n x_i} \right|, \quad j = 1, \dots, n.$$

Es claro que si se verifica la condición

$$\sum_{i=1}^n x_i \approx 0,$$

entonces los números de condición asociados al problema de sumar n números son grandes y por lo tanto el problema está mal condicionado.

Existe otra forma alternativa de obtener los números de condición para la suma de n números reales. Sean las perturbaciones $x_i^* = x_i \pm \Delta_i$, siendo Δ_i el error absoluto cometido en la aproximación $x_i \approx x_i^*$. El error absoluto Δ en la suma es

$$\Delta = \left| \sum_{i=1}^n x_i - \sum_{i=1}^n x_i^* \right| = \left| \sum_{i=1}^n (x_i - x_i^*) \right| = \left| \sum_{i=1}^n \Delta_i \right|.$$

Una cota del error relativo ϵ cometido en la suma es la siguiente

$$\epsilon = \left| \frac{\Delta}{\sum_{j=1}^n x_j} \right| = \frac{|\sum_{i=1}^n \Delta_i|}{|\sum_{j=1}^n x_j|} \leq \frac{\sum_{i=1}^n |\Delta_i|}{|\sum_{j=1}^n x_j|} = \sum_{i=1}^n \frac{|x_i|}{|\sum_{j=1}^n x_j|} \left| \frac{\Delta_i}{x_i} \right|$$

$$= \sum_{i=1}^n \frac{|x_i|}{\left| \sum_{j=1}^n x_j \right|} \epsilon_i ,$$

donde ϵ_i es el error relativo cometido en la aproximación $x_i \approx x_i^*$. Así, los números de condición son

$$\kappa_{x_i} = \frac{|x_i|}{\left| \sum_{j=1}^n x_j \right|} , i = 1, \dots, n.$$

Por lo que se refiere a la pregunta ¿Es bueno desde el punto de vista numérico restar dos números reales muy próximos? la respuesta es NO a la vista del resultado anterior puesto que $x_1 + x_2 \approx 0$ lo cual implica $\kappa_{x_i} \gg 1$.

Problema 2.6 *Demostrar que, si x está próximo al valor 1 entonces, pequeños errores relativos en x pueden provocar grandes errores relativos en $\arcsin x$.*

Solución. Definimos la función $f(x) = \arcsin x$. Sea ϵ_x el error relativo en la variable x . Se sabe, de la fórmula de propagación de errores relativos, que el error relativo en f viene aproximado por $\epsilon_f \approx \kappa_x \epsilon_x$, siendo el número de condición

$$\kappa_x = \left| \frac{x f'(x)}{f(x)} \right| = \left| \frac{x}{\sqrt{1-x^2} \arcsin x} \right| .$$

Puesto que para x próximo a 1 se tiene $\arcsin x \approx \pi/2$, entonces

$$\kappa_x \approx \left| \frac{2x}{\pi \sqrt{1-x^2}} \right| ,$$

es decir, $\lim_{x \rightarrow 1} \kappa_x = \infty$. De este modo se comprueba que, para x cercano a 1, el número de condición κ_x es muy grande con lo que, aunque ϵ_x sea pequeño, ϵ_f puede ser muy grande.

Problema 2.7 *Calcular aproximadamente con qué precisión se debe conocer el número $\sqrt{2}$ si se pretende calcular $\ln(1/\sqrt{2})$ con 4 cifras decimales de precisión.*

Solución. Consideremos la función $f(x) = \ln(1/x)$. Sea Δ_x y Δ_f los errores absolutos, es decir, precisiones, con los que se conocen x y $f(x)$ respectivamente. Se sabe de la teoría de propagación de errores que $\Delta_f \approx |f'(x)| \Delta_x$. En nuestro caso particular, esta fórmula reduce a

$$\Delta_f \approx \left| -\frac{1}{x} \right| \Delta_x ,$$

de donde, usando las propiedades del valor absoluto y despejando se llega a

$$\Delta_x \approx |x| \Delta_f .$$

Finalmente, en el caso que nos ocupa, $x = \sqrt{2}$ y $\Delta_f \leq 10^{-4}$ de modo que

$$\Delta_{\sqrt{2}} \leq \sqrt{2} 10^{-4} \approx 0,00014 < 10^{-3} .$$

Problema 2.8 Un bloque de masa m desliza bajo la acción de la gravedad por un plano inclinado formando un ángulo θ respecto de la horizontal. Se puede demostrar que, si la fuerza de rozamiento F_r entre el bloque y el plano viene dada por $F_r = -kmv^2$ con v la velocidad del bloque y k un coeficiente de rozamiento, entonces el tiempo T requerido para que el bloque recorra una distancia D partiendo del reposo está relacionado de la forma

$$e^{kD} = \cosh[T\sqrt{kg \sin \theta}] ,$$

siendo $g = 9,8$ la aceleración de la gravedad. Si $\theta = \pi/4$, $k = 0,5 \pm 0,1$ y $T = 2 \pm 0,2$, hallar la precisión con la que se conoce D .

Solución. A partir de la relación

$$e^{kD} = \cosh[T\sqrt{kg \sin \theta}] ,$$

despejamos D y obtenemos

$$D = f(k, T, \theta) = \frac{1}{k} \ln \left(\cosh[T\sqrt{kg \sin \theta}] \right) .$$

El valor aproximado de D será

$$D^* = f(0,5, 2, \pi/4) = 6,060 .$$

Puesto que en las magnitudes k y T se tienen asociadas errores absolutos Δ_k y Δ_T respectivamente, se obtiene la siguiente aproximación del error absoluto en la magnitud D

$$\begin{aligned} \Delta_D &\approx \left| \frac{\partial D}{\partial k} \right| \Delta_k + \left| \frac{\partial D}{\partial T} \right| \Delta_T \\ &= \left| -\frac{\ln \left(\cosh[T\sqrt{kg \sin \theta}] \right)}{k^2} + \frac{gT \sin \theta \tanh[T\sqrt{kg \sin \theta}]}{2k\sqrt{kg \sin \theta}} \right| \Delta_k + \\ &\quad \left| \frac{\sqrt{kg \sin \theta} \tanh[T\sqrt{kg \sin \theta}]}{k} \right| \Delta_T . \end{aligned}$$

Sustituyendo en esta expresión los valores $\Delta_k = 0,1$, $\Delta_T = 0,2$, $g = 9,8$, $T = 2$ y $k = 0,5$ se obtiene

$$\Delta_D \approx |-4,684| 0,1 + |3,718| 0,2 = 1,212 .$$

En definitiva se obtiene

$$D = D^* \pm \Delta_D \approx 6,060 \pm 1,212 .$$

Problema 2.9 Averiguar si el cálculo de raíces cuadradas es una operación computacionalmente segura atendiendo a la propagación de errores.

Solución. Sea \mathbb{R}^+ el conjunto de todos los números reales positivos. Sea $x^* \in \mathbb{R}^+$ una aproximación de $x \in \mathbb{R}^+$. Nos preguntamos cómo se propaga el error relativo si aproximamos \sqrt{x} por $\sqrt{x^*}$. Puesto que la función $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ definida mediante $f(x) = \sqrt{x}$ es una función de clase $C^1(\mathbb{R}^+)$, se sabe que $\epsilon_f \approx \kappa_x \epsilon_x$, siendo $\epsilon_x = |x - x^*|/|x|$ y $\epsilon_f = |f(x) - f(x^*)|/|f(x)|$ los errores relativos en las aproximaciones $x \approx x^*$ y $f(x) \approx f(x^*)$ respectivamente y

$$\kappa_x = \left| \frac{xf'(x)}{f(x)} \right|$$

el número de condición. En nuestro caso, $f(x) = \sqrt{x}$ y $f'(x) = 1/(2\sqrt{x})$, de modo que

$$\kappa_x = \frac{1}{2},$$

es decir, los errores relativos se reducen aproximadamente a la mitad cuando se evalúa una raíz cuadrada y por lo tanto el cálculo de raíces cuadradas es una operación computacionalmente segura.

Problema 2.10 Consideremos la función $f(x) = \sin x$. Se realiza con un ordenador una tabla donde se obtiene la aproximación dada por la fórmula de diferencias hacia adelante de $f'(1)$ con longitudes de paso h cada vez más pequeñas obteniéndose

$f'(1)$	0.5402980	0.5403023	0.5551115	0
h	10^{-5}	10^{-9}	10^{-15}	10^{-16}

Dar una explicación del comportamiento obtenido en términos de errores de truncamiento y errores de redondeo. ¿Se obtendrá un comportamiento similar para cualquier otra función $f(x)$?

Solución. La fórmula de diferencias hacia adelante aproxima la derivada $f'(x)$ de una función mediante

$$f'(x) \approx \frac{f(x+h) - f(x)}{h},$$

con un error truncamiento $\Delta = O(h)$. Esto significa que, si se realizan los cálculos de manera exacta (es decir, sin errores de redondeo) entonces el error Δ disminuye cuando lo hace la longitud de paso h y además $\lim_{h \rightarrow 0} \Delta = 0$. Sin embargo, cuando se realizan los cálculos con un ordenador, este comete errores de redondeo de forma inevitable. En el caso que nos ocupa, debido al epsilon de la máquina ϵ , cuando h se hace suficientemente pequeña $0 < h < \epsilon$, el ordenador aproxima $x+h \approx x$ de modo que $f(x+h) - f(x) \approx 0$. Este comportamiento explica el valor incorrecto $f'(1) = 0$ de la tabla dada en el enunciado.

En resumen, cuando el ordenador realiza los cálculos, el valor óptimo de h en la fórmula de diferencias hacia adelante debe ser pequeño para minimizar el error de truncamiento propio de la fórmula, pero sin ser del orden de magnitud de ϵ para evitar los errores de redondeo. En el caso particular que nos ocupa,

$f(x) = \sin x$ de modo que $f'(1) = \cos(1) = 0,5403023$ observándose en la tabla el comportamiento descrito anteriormente.

Este comportamiento lo verificará cualquier función $f(x)$ con la única excepción de las funciones constantes $f(x) = c \in \mathbb{R}$ para todo x . El motivo es que la función f no depende de su argumento x de modo que $f(x+h) = f(x)$ para todo x y h y además $f'(x) = 0$.

Problema 2.11 *Se puede demostrar que la sucesión $\{x_n\}_{n \in \mathbb{N}} \subset \mathbb{R}$ definida de la forma recurrente siguiente*

$$x_2 = 2, \quad x_{n+1} = 2^{n-\frac{1}{2}} \sqrt{1 - \sqrt{1 - 4^{1-n} x_n^2}}$$

converge hacia π . Explica el siguiente comportamiento al usar un ordenador para calcular los primeros términos de la sucesión.

n	2	6	10	25	550
x_n	2	3.1365	3.1415	3.1425	0

Solución. A pesar de que un cálculo muestra que $\lim_{n \rightarrow \infty} x_n = \pi \approx 3,14159$, en la tabla adjunta se observa que, al ir calculando con un ordenador los términos x_n de la sucesión para valores crecientes de n , el comportamiento es de un acercamiento hacia π para $n \leq 10$. Sin embargo, al seguir aumentando n , la sucesión comienza a alejarse de π hasta estabilizarse en el valor 0 para $n \geq 550$. La explicación de este comportamiento anómalo se debe a los errores de redondeo cometidos por el ordenador. Notemos que, para n suficientemente grande, 4^{1-n} es suficientemente cercano a cero como para que, debido al épsilon de la máquina, se cometa la aproximación por redondeo $1 - 4^{1-n} x_n^2 \approx 1$. De este modo se tiene que $\sqrt{1 - \sqrt{1 - 4^{1-n} x_n^2}} \approx 0$ con lo cual $x_{n+1} = 2^{n-\frac{1}{2}} \sqrt{1 - \sqrt{1 - 4^{1-n} x_n^2}} \approx 0$ mostrando de esta forma el comportamiento indeseado debido a errores de redondeo.

Problema 2.12 *Un disco de radio R y grosor despreciable está cargado eléctricamente con una densidad de carga σ constante. Consideremos su eje de simetría, es decir, la recta perpendicular al disco que pasa por su centro. Entonces, se puede demostrar que la intensidad de campo eléctrico E que genera el disco en un punto de su eje de simetría que está a distancia x del disco viene dado por*

$$E = \frac{\sigma}{2\epsilon_0} \left(1 - \frac{x}{\sqrt{x^2 + R^2}} \right),$$

siendo $\epsilon_0 = 8,8542 \times 10^{-12}$ la permitividad del vacío en el sistema internacional. Tomando unidades del sistema internacional, realizamos las siguientes mediciones: $\sigma = 10^{-10}$, $x = 1 \pm 0,1$ y $R = 3 \pm 0,1$. Aproximar el error absoluto con el que se conoce E .

Solución. Puesto que las únicas magnitudes que contienen errores son x y R ,

tomaremos la función

$$E(x, R) = \frac{\sigma}{2\epsilon_0} \left(1 - \frac{x}{\sqrt{x^2 + R^2}} \right).$$

Denotamos Δ_E , Δ_x y Δ_R a los errores absolutos asociados a las cantidades E , x y R . Utilizando la fórmula de propagación de errores absolutos se tiene que

$$\Delta_E \approx \left| \frac{\partial E}{\partial x} \right| \Delta_x + \left| \frac{\partial E}{\partial R} \right| \Delta_R = \left| \frac{-R^2 \sigma}{2\epsilon_0 (x^2 + R^2)^{3/2}} \right| \Delta_x + \left| \frac{R \sigma x}{2\epsilon_0 (x^2 + R^2)^{3/2}} \right| \Delta_R.$$

Tomamos $\epsilon_0 = 8,8542 \times 10^{-12}$, $\sigma = 10^{-10}$ y recordamos que $x = 1 \pm 0,1$ y $R = 3 \pm 0,1$ de modo que $\Delta_x = \Delta_R = 0,1$. Particularizando para estos valores se obtiene

$$\Delta_E \approx |-1,60718| \times 0,1 + |0,535725| \times 0,1 = 0,21429.$$

Capítulo 3

Sistemas Lineales

3.1. Resumen de teoría

El sistema lineal se escribe en la forma matricial $Ax = b$, siendo $A = (a_{ij})$ una matriz cuadrada $n \times n$ y $b = (b_i)^T$ un vector columna. Supondremos que el sistema es compatible y determinado, es decir la matriz A es *no singular* y por lo tanto $\det A \neq 0$, lo que nos garantiza una única solución $x^T = (x_1^0, \dots, x_n^0)$. Desde un punto de vista teórico, la resolución de tales sistemas de ecuaciones puede hacerse a través de la regla de Cramer (cálculo de determinantes) o bien mediante el cálculo de la matriz inversa $x = A^{-1}b$. Sin embargo, ambos métodos no son recomendables desde el punto de vista numérico para sistemas lineales con muchas ecuaciones e incógnitas bien por que son muy costosos o por que son numéricamente inestables.

Los procedimientos de resolución de sistemas de ecuaciones lineales se dividen fundamentalmente en dos grupos

- *Métodos directos*, que son algoritmos finitos para el cálculo de las soluciones de un sistema lineal. El método de eliminación Gaussiana es un ejemplo de este tipo.
- *Métodos iterativos*, los cuales permiten obtener las soluciones de un sistema lineal con una exactitud dada mediante procesos infinitos convergentes.

3.1.1. Algoritmo de eliminación Gaussiana

La eliminación Gaussiana es uno de los métodos más utilizados a la hora de resolver numéricamente un sistema de ecuaciones lineales. Llamaremos *transformaciones elementales* a las siguientes manipulaciones

- (1) Permutar ecuaciones.
- (2) Multiplicar una ecuación por un número diferente de cero.

(3) Sumar a una ecuación una combinación lineal de las ecuaciones restante.

Es conocido que las transformaciones elementales no modifican la solución del sistema lineal.

El algoritmo Gaussiano consiste en la utilización repetida de la segunda y tercera transformación elemental con el objetivo de transformar el sistema en uno triangular superior. Una vez se ha conseguido este paso, la solución del sistema triangular superior vendrá dada por el algoritmo de sustitución hacia atrás.

Consideremos el sistema $Ax = b$, con $\det A \neq 0$. Más explícitamente

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix}. \quad (3.1)$$

Paso 1. Denotemos por $A^{(0)} = A$ y supongamos que $a_{11} \neq 0$. El primer paso de la eliminación Gaussiana consiste en eliminar x_1 de las $n - 1$ últimas ecuaciones. Para hacerlo, le restamos a la i -ésima ecuación la primera ecuación multiplicada por el factor $m_{i1} = a_{i1}/a_{11}$ para $i = 2, \dots, n$. De esta forma el sistema lineal queda $A^{(1)}x = b^{(1)}$, es decir

$$\begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & \cdots & a_{3n}^{(1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & a_{n3}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(0)} \\ b_2^{(1)} \\ b_3^{(1)} \\ \vdots \\ b_n^{(1)} \end{pmatrix}, \quad (3.2)$$

donde hemos definido $a_{ij}^{(1)} = a_{ij}^{(0)} - m_{i1}a_{1j}^{(0)}$ y $b_i^{(1)} = b_i^{(0)} - m_{i1}b_1^{(0)}$ para $i, j = 2, \dots, n$.

Paso 2. Supongamos que $a_{22}^{(1)} \neq 0$. El segundo paso del algoritmo Gaussiana consiste en eliminar x_2 de las $n - 2$ últimas ecuaciones del sistema (3.2). Para conseguirlo, restamos a la i -ésima ecuación la segunda ecuación multiplicada por el factor $m_{i2} = a_{i2}^{(1)}/a_{22}^{(1)}$ para $i = 3, \dots, n$, consiguiendo que el sistema (3.2) adopte la forma matricial $A^{(2)}x = b^{(2)}$ o más detallado

$$\begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(0)} \\ b_2^{(1)} \\ b_3^{(2)} \\ \vdots \\ b_n^{(2)} \end{pmatrix}, \quad (3.3)$$

siendo $a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i2}a_{2j}^{(1)}$ y $b_i^{(2)} = b_i^{(1)} - m_{i2}b_2^{(1)}$ para $i, j = 3, \dots, n$.

Paso k -ésimo. Después de $k - 1$ pasos, el sistema tiene la siguiente forma

$$\begin{pmatrix} a_{11}^{(0)} & \cdots & a_{1k-1}^{(0)} & a_{1k}^{(0)} & \cdots & a_{1n}^{(0)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & a_{k-1k-1}^{(k-2)} & a_{k-1k}^{(k-2)} & \cdots & a_{k-1n}^{(k-2)} \\ 0 & \cdots & 0 & a_{kk}^{(k-1)} & \cdots & a_{kn}^{(k-1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nk}^{(k-1)} & \cdots & a_{nn}^{(k-1)} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_{k-1} \\ x_k \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(0)} \\ \vdots \\ b_{k-1}^{(k-2)} \\ b_k^{(k-1)} \\ \vdots \\ b_n^{(k-1)} \end{pmatrix}, \quad (3.4)$$

de manera que para conseguir los ceros subdiagonales de la columna k -ésima hacemos (siempre y cuando $a_{kk}^{(k-1)} \neq 0$) la transformación siguiente

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - m_{ik}a_{kj}^{(k-1)}, \quad b_i^{(k)} = b_i^{(k-1)} - m_{ik}b_k^{(k-1)}, \quad (3.5)$$

siendo $m_{ik} = a_{ik}^{(k-1)} / a_{kk}^{(k-1)}$ y con $i, j = k + 1, \dots, n$.

Los elementos $a_{kk}^{(k-1)}$ con $k = 1, 2, \dots$ que van apareciendo en este proceso se llaman *pivotes*. Si todos los pivotes son no nulos, la aplicación descrita en (3.5) desde $k = 1, \dots, n - 1$ da lugar a la triangularización superior de la matriz de los coeficientes del sistema

$$\begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(n-1)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(0)} \\ b_2^{(1)} \\ \vdots \\ b_n^{(n-1)} \end{pmatrix}, \quad (3.6)$$

de manera que hemos reducido el problema a la resolución de un sistema triangular.

Estrategia de *pivotación parcial*: En el paso k -ésimo, en el cual pasamos de la matriz $A^{(k-1)} \rightarrow A^{(k)}$, podemos tener el pivote $a_{kk}^{(k-1)}$ o bien nulo o bien muy pequeño. Entonces, tomaremos como pivote al elemento de la columna k -ésima diagonal y subdiagonal $a_{kk}^{(k-1)}, a_{k+1k}^{(k-1)}, \dots, a_{nk}^{(k-1)}$ más grande en valor absoluto e intercambiaremos la fila k -ésima y la fila del pivote elegido. Es decir, sea r el número entero más pequeño que verifique

$$a_{rk}^{(k-1)} = \max_{i=k, \dots, n} |a_{ik}^{(k-1)}|, \quad (3.7)$$

e intercambiaremos las filas (ecuaciones) k y r .

Existen casos particulares para los cuales el algoritmo de eliminación Gaussiana aplicado a la resolución del sistema lineal $Ax = b$ es numéricamente estable y no es necesaria ninguna estrategia de pivote. Alguno de estos casos son cuando

- La matriz cuadrada A es *diagonal dominante por filas*, es decir

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad \text{para } i = 1, \dots, n.$$

- La matriz cuadrada A es *simétrica* ($A = A^t$) y *definida positiva* ($x^t A x > 0$ para toda $x \neq 0$ con $x \in \mathbb{R}^n$). A este respecto es útil el llamado *Criterio de Sylvester*: “una matriz simétrica es definida positiva si y sólo si todos sus determinantes menores principales Δ_k son positivos”.

Del estudio realizado en la eliminación Gaussiana sin pivotación, se desprende que los elementos de la matriz triangular superior resultante vienen dados por

$$a_{ij}^{(i-1)} = \begin{cases} a_{ij} - m_{i1}a_{1j} - m_{i2}a_{2j}^{(1)} - \dots - m_{ii-1}a_{i-1j}^{(i-2)} & \text{si } i \leq j, \\ 0 & \text{si } i > j, \end{cases}$$

de lo que se deduce que

$$a_{ij} = m_{i1}a_{1j} + m_{i2}a_{2j}^{(1)} + \dots + m_{ii-1}a_{i-1j}^{(i-2)} + a_{ij}^{(i-1)}.$$

Esta última ecuación nos permite asegurar que la matriz original de coeficientes del sistema puede descomponer en el producto de dos matrices triangulares de la forma

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} = \quad (3.8)$$

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & m_{n3} & \cdots & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n-1)} \end{pmatrix}$$

Respecto a la unicidad de esta descomposición tenemos el siguiente teorema.

Teorema 3.1 Sea A una matriz $n \times n$. Denotemos por A_k las submatrices cuadradas $k \times k$ formadas por las primeras k filas y columnas de A . Si los determinantes menores principales $\det A_k \neq 0$ para $k = 1, \dots, n$, entonces existen matrices triangulares únicas $L = (\ell_{ij})$ y $U = (u_{ij})$, inferior y superior respectivamente, con $\ell_{ii} = 1$ para $i = 1, \dots, n$, tales que $A = LU$.

La relación entre la eliminación Gaussiana y la descomposición LU descrita en el Teorema 3.1 es la siguiente

$$u_{ij} = \begin{cases} 0 & \text{si } i > j, \\ a_{ij}^{(i-1)} & \text{si } i \leq j, \end{cases}, \quad \ell_{ij} = \begin{cases} 0 & \text{si } i < j, \\ 1 & \text{si } i = j, \\ m_{ij} & \text{si } i > j, \end{cases}$$

donde el paso k -ésimo de la eliminación Gaussiana de una matriz cuadrada de orden n consiste en la transformación elemental

$$f_i \rightarrow f_i - m_{ik}f_k, \quad \text{para } i = k+1, \dots, n,$$

utilizando la notación f_i para la fila i -ésima de la matriz.

Finalmente, recordemos que el *número de condición* de la matriz A respecto de una norma dada es

$$\kappa(A) := \|A\| \|A^{-1}\|. \quad (3.9)$$

Sea $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$. Algunas normas matriciales son:

- $\|A\|_1 := \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$.
- $\|A\|_\infty := \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$.

3.2. Problemas resueltos

Problema 3.1 Consideremos el sistema lineal $Ax = b$ definido de la forma

$$\begin{pmatrix} 6 & -2 & 2 & 4 \\ 12 & -8 & 6 & 10 \\ 3 & -13 & 9 & 3 \\ -6 & 4 & 1 & -18 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 12 \\ 34 \\ 27 \\ -38 \end{pmatrix}.$$

- (i) Resolver el sistema lineal utilizando el algoritmo de eliminación Gaussiana sin pivotación.
- (ii) A partir de los cálculos efectuados en el apartado (i), hallar la descomposición LU de la matriz A y calcular el valor del determinante de A .
- (iii) Demostrar que la descomposición LU hallada es única.

Solución. (i) Realizando transformaciones elementales por filas a la matriz ampliada A^* del sistema lineal se tiene $A^* \sim A^{(1)} \sim A^{(2)} \sim A^{(3)}$ como detallamos a continuación

$$\begin{pmatrix} 6 & -2 & 2 & 4 & 12 \\ 12 & -8 & 6 & 10 & 34 \\ 3 & -13 & 9 & 3 & 27 \\ -6 & 4 & 1 & -18 & -38 \end{pmatrix} \sim \begin{pmatrix} 6 & -2 & 2 & 4 & 12 \\ 0 & -4 & 2 & 2 & 10 \\ 0 & -12 & 8 & 1 & 21 \\ 0 & 2 & 3 & -14 & -26 \end{pmatrix} \sim$$

$$\begin{pmatrix} 6 & -2 & 2 & 4 & 12 \\ 0 & -4 & 2 & 2 & 10 \\ 0 & 0 & 2 & -5 & -9 \\ 0 & 0 & 4 & -13 & -21 \end{pmatrix} \sim \begin{pmatrix} 6 & -2 & 2 & 4 & 12 \\ 0 & -4 & 2 & 2 & 10 \\ 0 & 0 & 2 & -5 & -9 \\ 0 & 0 & 0 & -3 & -3 \end{pmatrix}.$$

Ahora podemos resolver el sistema por substitución hacia atrás y obtener

$$(x_1, x_2, x_3, x_4) = (1, -3, -2, 1) .$$

(ii) Sabemos que $U = A^{(3)}$ es triangular superior y que $L = (\ell_{ij})$ es triangular inferior tal que

$$\ell_{ij} = \begin{cases} 0 & \text{si } i < j \\ 1 & \text{si } i = j \\ m_{ij} & \text{si } i > j \end{cases}$$

Las transformaciones elementales (por filas) realizadas en el apartado (i) han sido:

$$\begin{aligned} (1) \quad & \begin{cases} f_2 \rightarrow f_2 - 2f_1 & \Rightarrow m_{21} = 2 \\ f_3 \rightarrow f_3 - 1/2f_1 & \Rightarrow m_{31} = 1/2 \\ f_4 \rightarrow f_4 + f_1 & \Rightarrow m_{41} = -1 \end{cases} \\ (2) \quad & \begin{cases} f_3 \rightarrow f_3 - 3f_2 & \Rightarrow m_{32} = 3 \\ f_4 \rightarrow f_4 + 1/2f_2 & \Rightarrow m_{42} = -1/2 \\ f_4 \rightarrow f_4 + f_1 & \Rightarrow m_{41} = -1 \end{cases} \\ (3) \quad & f_4 \rightarrow f_4 - 2f_3 \Rightarrow m_{43} = 2 \end{aligned}$$

de manera que la descomposición LU de la matriz A viene dada por las matrices

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1/2 & 3 & 1 & 0 \\ -1 & -1/2 & 2 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 6 & -2 & 2 & 4 \\ 0 & -4 & 2 & 2 \\ 0 & 0 & 2 & -5 \\ 0 & 0 & 0 & -3 \end{pmatrix}.$$

Finalmente, el determinante de la matriz A es

$$\det A = \det(LU) = \det L \det U = \det U = 6(-4)2(-3) = 144 .$$

(iii) Definimos Δ_k como el determinante menor principal de orden k de la matriz A . Que la descomposición LU hallada es única se desprende de la condición suficiente $\Delta_k \neq 0$ para $k = 1, 2, 3, 4$. En concreto tenemos

$$\Delta_1 = 6, \quad \Delta_2 = -24, \quad \Delta_3 = -48, \quad \Delta_4 = \det A = 144 .$$

Problema 3.2 Supongamos que se quiere resolver el sistema lineal $Ax = b$, donde $x, b \in \mathbb{R}^2$, $b \neq 0$ y A es la matriz cuadrada 2×2 siguiente

$$A = \begin{pmatrix} 1 & 1 + \epsilon \\ 1 - \epsilon & 1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}),$$

con $\epsilon > 0$. Utilizando la norma del infinito y suponiendo que $\epsilon \leq 0,01$, demostrar que una pequeña perturbación relativa del término independiente b puede producir una perturbación relativa 40401 veces mayor en la solución del sistema lineal.

Solución. Sea Δb una perturbación del vector de términos independientes del sistema $Ax = b$. Se sabe que el error relativo que se producirá en la solución del sistema debido a la perturbación viene acotado por

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

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

En primer lugar calculamos la inversa de la matriz A , por ejemplo mediante la utilización de la fórmula

$$A^{-1} = \frac{1}{\det A} (A^*)^T,$$

donde A^* es la matriz de adjuntos de la matriz A y viene dada por

$$A^* = \begin{pmatrix} 1 & \epsilon - 1 \\ -1 - \epsilon & 1 \end{pmatrix}.$$

Como $\det A = \epsilon^2$, se concluye que

$$A^{-1} = \frac{1}{\epsilon^2} \begin{pmatrix} 1 & -\epsilon - 1 \\ \epsilon - 1 & 1 \end{pmatrix}.$$

Finalmente se calcula, utilizando la norma del infinito, el número de condición de la matriz A

$$\begin{aligned} \kappa(A) &= \|A\|_{\infty} \|A^{-1}\|_{\infty} \\ &= \max\{1 + |\epsilon - 1|, 1 + \epsilon + 1\} \max\left\{\frac{1 + |\epsilon - 1|}{\epsilon^2}, \frac{1 + 1 + \epsilon}{\epsilon^2}\right\} \\ &= (2 + \epsilon) \frac{2\epsilon}{\epsilon^2} = \left(\frac{2 + \epsilon}{\epsilon}\right)^2. \end{aligned}$$

Si tomamos $\epsilon = 0,01$ se tiene $\kappa(A) = 40401$, de manera que

$$\frac{\|\Delta x\|}{\|x\|} \leq 40401 \frac{\|\Delta b\|}{\|b\|},$$

como se quería demostrar.

Problema 3.3 Considerar la matriz $A \in \mathcal{M}_4(\mathbb{R})$ definida de la forma

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

- (i) Hallar la descomposición LU de la matriz A y calcular el valor del determinante de A .
- (ii) Averiguar si la descomposición LU hallada es única.

Solución. (i) Realizamos las siguientes transformaciones elementales a la matriz A

$$A = \begin{pmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 2 \\ -1 & 2 & 3 & -1 \end{pmatrix} \sim A^{(1)} = \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & -4 & -1 & -7 \\ 0 & 3 & 3 & 2 \end{pmatrix} \sim$$

$$A^{(2)} = \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{pmatrix} = A^{(3)}.$$

Sabemos que $U = A^{(3)}$ es triangular superior y que $L = (\ell_{ij})$ es triangular inferior tal que

$$\ell_{ij} = \begin{cases} 0 & \text{si } i < j \\ 1 & \text{si } i = j \\ m_{ij} & \text{si } i > j \end{cases}$$

Las transformaciones elementales (por filas) realizadas han sido:

$$\begin{aligned} (1) \quad & \begin{cases} f_2 \rightarrow f_2 - 2f_1 & \implies m_{21} = 2 \\ f_3 \rightarrow f_3 - 3f_1 & \implies m_{31} = 3 \\ f_4 \rightarrow f_4 + f_1 & \implies m_{41} = -1 \end{cases} \\ (2) \quad & \begin{cases} f_3 \rightarrow f_3 - 4f_2 & \implies m_{32} = 4 \\ f_4 \rightarrow f_4 + 3f_2 & \implies m_{42} = -3 \end{cases} \\ (3) \quad & f_4 \rightarrow f_4 - 0f_3 \implies m_{43} = 0 \end{aligned}$$

de manera que la descomposición LU de la matriz A es

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{pmatrix}.$$

Finalmente, el determinante de la matriz A es

$$\det A = \det(LU) = \det L \det U = \det U = (1)(-1)(3)(-13) = 39.$$

(ii) Definimos Δ_k como el determinante menor principal de orden k de la matriz A . Que la descomposición LU hallada es única se desprende de la condición suficiente $\Delta_k \neq 0$ para $k = 1, 2, 3, 4$. En concreto tenemos

$$\Delta_1 = 1, \quad \Delta_2 = -1, \quad \Delta_3 = -3, \quad \Delta_4 = \det A = 39.$$

Problema 3.4 Averiguar si es posible resolver el siguiente sistema lineal por descomposición LU .

$$\begin{pmatrix} 1 & 2 & 6 \\ 4 & 8 & -1 \\ -2 & 3 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 9 \\ 100 \\ 45 \end{pmatrix}$$

Solución. Recordemos que, dada una matriz $A \in \mathcal{M}_n(\mathbb{R})$, si todos sus determinantes menores principales Δ_k con $k = 1, 2, \dots, n$, son diferentes de cero, entonces existen matrices triangulares únicas $L = (\ell_{ij})$ y $U = (u_{ij})$ triangulares inferiores y superiores respectivamente con $\ell_{ii} = 1$ para $i = 1, 2, \dots, n$ tales que $A = LU$.

En nuestro caso, si A es la matriz de coeficientes del sistema lineal dado, sus determinantes menores principales son

$$\Delta_1 = 1 \neq 0, \quad \Delta_2 = \begin{vmatrix} 1 & 2 \\ 4 & 8 \end{vmatrix} = 0, \quad \Delta_3 = \det A = 175 \neq 0.$$

Como $\Delta_2 = 0$ no podemos aplicar el resultado que hemos comentado en un principio. Así, hemos de proceder a su construcción explícita. En concreto, suponiendo que $A = LU$, se tiene

$$\begin{pmatrix} 1 & 2 & 6 \\ 4 & 8 & -1 \\ -2 & 3 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}.$$

Efectuando el producto de las dos matrices anteriores e igualando elemento a elemento se tiene:

(a) En la primera columna:

$$\begin{aligned} 1 &= u_{11}, \\ 4 &= m_{21}u_{11} = m_{21}, \\ -2 &= m_{31}u_{11} = m_{31}. \end{aligned}$$

(b) En la segunda columna:

$$\begin{aligned} 2 &= u_{12}, \\ 8 &= m_{21}u_{12} + u_{22} = 4 \times 2 + u_{22}, \quad u_{22} = 0, \\ 3 &= m_{31}u_{12} + m_{32}u_{22} = -2 \times 2 + m_{32} \times 0 = -4, \end{aligned}$$

donde esta última igualdad es una contradicción. Concluimos que A no admite descomposición LU y por lo tanto el sistema del enunciado no puede ser resuelto por este método.

Problema 3.5 Sean x e y la solución del sistema lineal

$$\begin{pmatrix} b & 2 \\ 1 & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

- (i) Averiguar para qué valores de los parámetros a y b es posible hallar la descomposición LU de la matriz de coeficientes del sistema a partir del algoritmo de eliminación Gaussiana sin pivotación. Para esos valores de los parámetros obtener las matrices L y U . ¿Es única dicha descomposición si $a = 2$ y $b = 7$?
- (ii) Utilizando la norma L^1 , calcular el número de condición de la matriz de coeficientes del sistema cuando $a = 1$ y $b = 3$.
- (iii) Si $a = 2 \pm 0,5$ y $b = 5 \pm 0,1$, aproximar la precisión con la que se puede determinar $x + y$.

Solución. (i) Sea A la matriz de coeficientes del sistema lineal, es decir,

$$A = \begin{pmatrix} b & 2 \\ 1 & a \end{pmatrix}.$$

Para hallar la descomposición LU de la matriz A mediante el algoritmo de eliminación Gaussiana sin pivotación, el primer paso de dicho algoritmo consiste en realizar la transformación elemental por filas $f_2 \rightarrow f_2 - m_{21}f_1$, siendo el multiplicador $m_{21} = 1/b$. Obviamente este paso es sólo posible cuando el pivote utilizado $b \neq 0$. Bajo esta hipótesis, se realiza la anterior transformación elemental y se obtiene la matriz $A^{(1)}$ semejante a A y dada por

$$A^{(1)} = \begin{pmatrix} b & 2 \\ 0 & a - \frac{2}{b} \end{pmatrix}.$$

Puesto que $A^{(1)}$ ya es triangular superior, el algoritmo finaliza. En resumen, si $b \neq 0$ entonces

$$A = LU = LA^{(1)} = \begin{pmatrix} 1 & 0 \\ \frac{1}{b} & 1 \end{pmatrix} \begin{pmatrix} b & 2 \\ 0 & a - \frac{2}{b} \end{pmatrix}.$$

Finalmente, como para los valores $a = 2$ y $b = 7$ los determinantes menores principales de A son no nulos, es decir,

$$\Delta_1 = 7 \neq 0, \quad \Delta_2 = 12 \neq 0,$$

se concluye que la descomposición LU es única.

(ii) Si $a = 1$ y $b = 3$ la matriz A es

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

Tomando la norma L^1 , el número de condición $\kappa(A)$ de la matriz A es $\kappa(A) = \|A\|_1 \|A^{-1}\|_1$. Finalmente, puesto que la matriz inversa de A es

$$A^{-1} = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix},$$

las normas involucradas son

$$\|A\|_1 = \max\{|3|+|1|, |2|+|11|\} = 4, \quad \|A^{-1}\|_1 = \max\{|1|+|-1|, |-2|+|3|\} = 5.$$

En definitiva $\kappa(A) = 20$.

(iii) Puesto que $a = 2 \pm 0,5$ y $b = 5 \pm 0,1$, es obvio que $\det A = ab - 2 \neq 0$ y por lo tanto el sistema lineal es compatible y determinado. Además, su solución es

$$x = \frac{2(a-3)}{ab-2}, \quad y = \frac{3b-2}{ab-2}.$$

Se tiene por lo tanto que la función f suma de las soluciones del sistema lineal, es decir $f = x + y$, vale

$$f(a, b) = \frac{2a + 3b - 8}{ab - 2}.$$

Puesto que a y b tienen errores absolutos Δ_a y Δ_b respectivamente, se puede aproximar el error absoluto Δ_f al calcular la función f de la forma

$$\begin{aligned} \Delta_f &\approx \left| \frac{\partial f}{\partial a}(a, b) \right| \Delta_a + \left| \frac{\partial f}{\partial b}(a, b) \right| \Delta_b \\ &= \left| \frac{(3b-2)(2-b)}{(ab-2)^2} \right| \Delta_a + \left| \frac{2(a-3)(1-a)}{(ab-2)^2} \right| \Delta_b. \end{aligned}$$

Sustituyendo en la anterior expresión $a = 2$, $b = 5$, $\Delta_a = 0,5$ y $\Delta_b = 0,1$ se obtiene $\Delta_f \approx 0,307813$.

Problema 3.6 Si $\epsilon = 10^{-5}$, averiguar si la siguiente matriz $A \in \mathcal{M}_2(\mathbb{R})$ está bien condicionada utilizando la norma L^1

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 + \epsilon \end{pmatrix}.$$

Solución. Sea $A = (a_{ij}) \in \mathcal{M}_2(\mathbb{R})$. Entonces

$$\|A\|_1 := \max_{1 \leq j \leq 2} \sum_{i=1}^2 |a_{ij}|.$$

La matriz inversa de A es

$$A^{-1} = \frac{1}{\epsilon} \begin{pmatrix} 2 + \epsilon & -2 \\ -1 & 1 \end{pmatrix}.$$

De esta forma, se tiene que

$$\begin{aligned}\|A\|_1 &= \max\{2, 4 + \epsilon\} = 4 + \epsilon, \\ \|A^{-1}\|_1 &= \frac{1}{\epsilon} \max\{3 + \epsilon, 3\} = \frac{3 + \epsilon}{\epsilon}.\end{aligned}$$

Se concluye que, el número de condición $\kappa(A)$ asociado a A en la norma L^1 vale

$$\kappa(A) = \|A\|_1 \|A^{-1}\|_1 = (4 + \epsilon) \frac{3 + \epsilon}{\epsilon}.$$

Para el caso $\epsilon = 10^{-5}$ se obtiene

$$\kappa(A) = 1.20001 \times 10^6 \gg 1,$$

de manera que A está muy mal condicionada.

Problema 3.7 Consideremos el sistema lineal $Ax = b$, siendo

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

- (i) Averiguar si la matriz A admite una única descomposición LU .
- (ii) Utilizar dicha descomposición para resolver el sistema lineal del enunciado.

Solución. (i) Puesto que todos los determinantes menores principales de la matriz A

$$\Delta_1 = 3 \neq 0, \quad \Delta_2 = 7 \neq 0, \quad \Delta_3 = \det A = -7 \neq 0$$

son no nulos, podemos concluir que la matriz A admite una única descomposición LU .

(ii) Para hallar la descomposición LU de la matriz A , utilizaremos el algoritmo de eliminación gaussiana sin pivotación.

$$A \sim \begin{pmatrix} 3 & -1 & 2 \\ 0 & 7/3 & 7/3 \\ 0 & -4/3 & -7/3 \end{pmatrix} \sim \begin{pmatrix} 3 & -1 & 2 \\ 0 & 7/3 & 7/3 \\ 0 & 0 & -1 \end{pmatrix} = U,$$

donde las transformaciones elementales por filas realizadas han sido: 1) $f_i \rightarrow f_i - m_{i1}f_1$, con $i = 2, 3$, siendo $m_{21} = 1/3$ y $m_{31} = 2/3$; 2) $f_3 \rightarrow f_3 - m_{32}f_2$, con $m_{32} = -4/7$.

En definitiva $A = LU$, siendo

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 2/3 & -4/7 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 3 & -1 & 2 \\ 0 & 7/3 & 7/3 \\ 0 & 0 & -1 \end{pmatrix}.$$

Descomponiendo el sistema lineal $Ax = b$ en los dos sistemas triangulares $Ly = b$, $Ux = y$ se tiene que, la solución del primer sistema se puede hallar por sustitución hacia adelante obteniéndose $y = (0, 1, 18/7)^t$ mientras que la solución del segundo sistema por sustitución hacia atrás es $x = (19/7, 3, -18/7)^t$.

Problema 3.8 Considerar un conjunto de osciladores armónicos acoplados con un grado de libertad. En concreto el sistema tiene 3 masas y cuatro muelles enganchados a dos paredes de forma consecutiva. Las distancias de las masas a la pared izquierda son x_i , con $i = 1, 2, 3$. Las paredes están separadas una distancia d . Los cuatro muelles tienen asociados las constantes elásticas k_i y una longitud natural ℓ_i , con $i = 1, 2, 3, 4$. Es fácil demostrar que, las posiciones de equilibrio x_i^* con $i = 1, 2, 3$ de los osciladores deben verificar el sistema lineal tridiagonal

$$\begin{pmatrix} -(k_1 + k_2) & k_2 & 0 \\ k_2 & -(k_2 + k_3) & k_3 \\ 0 & k_3 & -(k_3 + k_4) \end{pmatrix} \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \end{pmatrix} = \begin{pmatrix} -k_1\ell_1 + k_2\ell_2 \\ -k_2\ell_2 + k_3\ell_3 \\ -k_3\ell_3 + k_4(\ell_4 - d) \end{pmatrix}.$$

- (i) Supongamos que se han de hallar las posiciones de equilibrio x_i^* con $i = 1, 2, 3$ de los osciladores en 1000 experimentos diferentes, donde la única variación entre los experimentos es la longitud d entre las paredes. Razonar qué método se programaría en un ordenador para resolver los 1000 sistemas lineales asociados a cada experimento.
- (ii) Si todos los muelles son iguales de constante elástica k y longitud natural ℓ , utilizar el método elegido en el apartado anterior para hallar las posiciones de equilibrio de los osciladores.

Solución. (i) Se han de resolver 1000 sistemas lineales $Ax = b_i$ para $i = 1, \dots, 1000$ con $A \in \mathcal{M}_3(\mathbb{R})$, $b_i \in \mathbb{R}^3$ y $x = (x_1^*, x_2^*, x_3^*)^t \in \mathbb{R}^3$. Ya que la única variación entre los sistemas lineales a resolver es el vector de términos independientes b_i , es natural utilizar para su resolución el método de descomposición LU de la matriz A de coeficientes. Observar además que, los determinantes menores principales de A son no nulos debido a que $k_i \neq 0$ para $i = 1, 2, 3, 4$, de modo que la descomposición LU de la matriz A existe y es única.

(ii) Si todos los muelles son iguales de constante elástica k y longitud natural ℓ , el sistema lineal a resolver es

$$\begin{pmatrix} -2k & k & 0 \\ k & -2k & k \\ 0 & k & -2k \end{pmatrix} \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -kd \end{pmatrix}.$$

Denotando por A a la matriz de coeficientes de este sistema, hallaremos su descomposición LU a través del algoritmo de eliminación Gaussiana. En concreto

se tiene

$$\begin{aligned} A &= \begin{pmatrix} -2k & k & 0 \\ k & -2k & k \\ 0 & k & -2k \end{pmatrix} \sim \begin{pmatrix} -2k & k & 0 \\ 0 & -3/2k & k \\ 0 & k & -2k \end{pmatrix} \\ &\sim \begin{pmatrix} -2k & k & 0 \\ 0 & -3/2k & k \\ 0 & 0 & -4/3k \end{pmatrix} = U. \end{aligned}$$

Las transformaciones elementales por filas realizadas han sido en primer lugar $f_i \rightarrow f_i - m_{i1}f_1$ para $i = 2, 3$, con $m_{21} = -1/2$ y $m_{31} = 0$ y en segundo lugar $f_3 \rightarrow f_3 - m_{32}f_2$ con $m_{32} = -2/3$. De este modo se tiene que

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -2/3 & 1 \end{pmatrix}.$$

Aprovechando la descomposición $A = LU$, el sistema lineal a resolver $Ax = b$ con $b = (0, 0, -kd)^t$ se descompone en dos sistemas triangulares $Ly = b$, $Ux = y$. La solución del primer sistema por sustitución hacia adelante es $y = (0, 0, -kd)^t$, mientras que la solución del segundo sistema por sustitución hacia atrás es $x = (d/4, d/2, 3d/4)^t$. Se tiene pues que las posiciones de equilibrio de los osciladores son

$$x_1^* = \frac{d}{4}, \quad x_2^* = \frac{d}{2}, \quad x_3^* = \frac{3d}{4}.$$

Problema 3.9 Resolver el siguiente sistema lineal utilizando el algoritmo de eliminación Gaussiana con pivotación parcial.

$$\begin{pmatrix} 6 & -2 & 2 & 4 \\ 12 & -8 & 6 & -10 \\ 3 & -13 & 9 & 3 \\ -6 & 4 & 1 & -18 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 12 \\ 34 \\ 27 \\ -38 \end{pmatrix}.$$

Solución. Tomamos la matriz ampliada A^* del sistema

$$A^* = \begin{pmatrix} 6 & -2 & 2 & 4 & 12 \\ 12 & -8 & 6 & -10 & 34 \\ 3 & -13 & 9 & 3 & 27 \\ -6 & 4 & 1 & -18 & -38 \end{pmatrix}.$$

Con el objetivo de tener el pivote mayor en valor absoluto, permutamos las filas 1 y 2 de modo que

$$A^* \sim A^{(1)} = \begin{pmatrix} 12 & -8 & 6 & -10 & 34 \\ 6 & -2 & 2 & 4 & 12 \\ 3 & -13 & 9 & 3 & 27 \\ -6 & 4 & 1 & -18 & -38 \end{pmatrix}.$$

Ahora realizamos las transformaciones elementales por filas $f_i \rightarrow f_i - m_{i1}f_1$, para $i = 2, 3, 4$, siendo $m_{21} = 1/2$, $m_{31} = 1/4$ y $m_{41} = -1/2$. Se tiene de esta forma

$$A^{(1)} \sim A^{(2)} = \begin{pmatrix} 12 & -8 & 6 & -10 & 34 \\ 0 & 2 & -1 & 9 & -5 \\ 0 & -11 & 7,5 & 5,5 & 18,5 \\ 0 & 0 & 4 & -23 & -21 \end{pmatrix}.$$

Permutamos ahora la fila 2 y la 3 con lo cual

$$A^{(2)} \sim A^{(3)} = \begin{pmatrix} 12 & -8 & 6 & -10 & 34 \\ 0 & -11 & 7,5 & 5,5 & 18,5 \\ 0 & 2 & -1 & 9 & -5 \\ 0 & 0 & 4 & -23 & -21 \end{pmatrix}.$$

Realizando las transformaciones elementales por filas $f_i \rightarrow f_i - m_{i2}f_2$, para $i = 3, 4$, siendo $m_{32} = -2/11$ y $m_{42} = 0$ se llega a

$$A^{(3)} \sim A^{(4)} = \begin{pmatrix} 12 & -8 & 6 & -10 & 34 \\ 0 & -11 & 7,5 & 5,5 & 18,5 \\ 0 & 0 & 0,3636 & 10 & -1,6364 \\ 0 & 0 & 4 & -23 & -21 \end{pmatrix}.$$

Permutamos a continuación la fila 3 y la 4 de modo que

$$A^{(4)} \sim A^{(4)} = \begin{pmatrix} 12 & -8 & 6 & -10 & 34 \\ 0 & -11 & 7,5 & 5,5 & 18,5 \\ 0 & 0 & 4 & -23 & -21 \\ 0 & 0 & 0,3636 & 10 & -1,6364 \end{pmatrix}.$$

Finalmente, la transformaciones elemental $f_4 \rightarrow f_4 - m_{43}f_3$ con $m_{43} = 0,3636/4$ lleva el sistema a la forma triangular

$$A^{(4)} \sim A^{(5)} = \begin{pmatrix} 12 & -8 & 6 & -10 & 34 \\ 0 & -11 & 7,5 & 5,5 & 18,5 \\ 0 & 0 & 4 & -23 & -21 \\ 0 & 0 & 0 & 12,09 & 0,2726 \end{pmatrix}.$$

Utilizando a continuación el algoritmo de sustitución hacia atrás se obtiene la solución

$$(x, y, z, t) = (1,97118, -5,16165, -5,1203, 0,0225564).$$

Capítulo 4

Interpolación Polinómica

4.1. Resumen de teoría

*Interpol*ar una función $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ en un conjunto de $n + 1$ puntos $\{x_0, x_1, \dots, x_n\} \subset I$ es encontrar otra función ϕ de manera que sobre estos puntos, la nueva función tome los mismos valores que la función original. Es decir, verificando

$$\phi(x_i) = f(x_i) = f_i, \quad \text{para } i = 0, 1, \dots, n.$$

Son muchos los problemas en ingeniería en los que hay que estudiar funciones que no se conocen explícitamente. En su lugar sólo se conoce una tabulación de los valores que toma la función en ciertos puntos (por ejemplo como resultado de medidas experimentales). Dichos problemas son buenos candidatos para la utilización de las técnicas interpolatorias que explicaremos a continuación. Nosotros sólo vamos a centrarnos en la interpolación polinómica que es la más sencilla.

Teorema 4.1 *Supongamos conocido el valor de una función $f(x)$ en un conjunto de puntos distintos dos a dos $\{x_0, x_1, \dots, x_n\}$. Entonces, existe un único polinomio $P_n(x)$ de grado menor o igual que n que interpola a la función en estos puntos, es decir $P_n(x_i) = f(x_i)$ para $i = 0, \dots, n$.*

Presentamos ahora una método para calcular el polinomio interpolador $P_n(x)$ asociado a una tabla de datos (x_i, f_i) con $i = 0, \dots, n$. Este método es el más explícito, aunque no el más eficaz por lo que se refiere al número de operaciones requeridas. La *fórmula de interpolación de Lagrange* es

$$P_n(x) = \sum_{k=0}^n f_k \ell_k(x), \quad \ell_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^n \frac{x - x_j}{x_k - x_j}, \quad k = 0, \dots, n. \quad (4.1)$$

Los polinomios $\ell_k(x)$ reciben el nombre de *polinomios básicos de Lagrange*.

Otro método alternativo para calcular $P_n(x)$ es el llamado de las *diferencias divididas de Newton*. En concreto, el polinomio interpolador adopta la forma

$$P_n(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \cdots + f[x_0, x_1, \dots, x_n](x - x_0)(x - x_1) \cdots (x - x_{n-1}), \quad (4.2)$$

donde, las diferencias divididas vienen definidas recurrentemente de la forma siguiente

$$\begin{aligned} f[x_i] &= f_i, \\ f[x_i, x_{i+1}, \dots, x_{i+j}] &= \frac{f[x_{i+1}, x_{i+2}, \dots, x_{i+j}] - f[x_i, x_{i+1}, \dots, x_{i+j-1}]}{x_{i+j} - x_i}. \end{aligned} \quad (4.3)$$

Una forma efectiva de realizar esta recurrencia es el siguiente esquema triangular (dado para el caso $n = 3$)

x_0	$f[x_0]$			
		$f[x_0, x_1]$		
x_1	$f[x_1]$		$f[x_0, x_1, x_2]$	
		$f[x_1, x_2]$		$f[x_0, x_1, x_2, x_3]$
x_2	$f[x_2]$		$f[x_1, x_2, x_3]$	
		$f[x_2, x_3]$		
x_3	$f[x_3]$			

Teorema 4.2 Sea f una función de clase $C^{n+1}[a, b]$, y sea P_n un polinomio de grado menor o igual que n que interpola a la función f en los siguientes $n + 1$ puntos $\{x_0, x_1, \dots, x_n\} \subset [a, b]$ distintos dos a dos. Entonces, para cualquier $x \in [a, b]$, existe un punto $\xi_x \in (a, b)$ tal que

$$f(x) - P_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{i=0}^n (x - x_i). \quad (4.4)$$

Si, además de interpolar a la función f en $n + 1$ puntos $\{x_0, x_1, \dots, x_n\}$ queremos buscar un polinomio $H_{2n+1}(x)$ tal que $f'(x_i) = H'_{2n+1}(x_i)$ entonces, el proceso se llama interpolación de Hermite y se tiene el siguiente resultado.

Teorema 4.3 Sea $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ una función derivable en $[a, b]$. Consideremos $n + 1$ puntos $\{x_0, x_1, \dots, x_n\} \subset [a, b]$ distintos dos a dos. Existe un único polinomio $H_{2n+1}(x)$ de grado menor o igual que $2n + 1$ tal que $H_{2n+1}(x_k) = f(x_k)$ y $H'_{2n+1}(x_k) = f'(x_k)$, $k = 0, 1, \dots, n$. La expresión explícita de $H_{2n+1}(x)$ es la siguiente:

$$H_{2n+1}(x) = \sum_{k=0}^n f_k h_k(x) + \sum_{k=0}^n f'_k \tilde{h}_k(x) ,$$

siendo $h_k(x) = [1 - 2\ell'_k(x_k)(x - x_k)]\ell_k^2(x)$, $\tilde{h}_k(x) = (x - x_k)\ell_k^2(x)$, donde $\ell_k(x)$ denota al polinomio k -ésimo básico de Lagrange. Además, si $f \in C^{2n+2}[a, b]$, entonces

$$f(x) - H_{2n+1}(x) = \frac{f^{(2n+2)}(\xi_x)}{(2n+2)!} \prod_{k=0}^n (x - x_k)^2 ,$$

para todo $x \in [a, b]$.

Como se ha visto en la expresión (4.4), el error en la interpolación depende de los nodos a través del polinomio $\pi_n(x) = \prod_{i=0}^n (x - x_i)$ con $x_i \in [a, b]$. Nos preguntamos si es posible realizar una elección de los nodos de manera que, la norma del máximo o del infinito de $\pi_n(x)$ en el intervalo $[a, b]$ dada por

$$\|\pi_n(x)\|_\infty := \max_{a \leq x \leq b} \left\{ \prod_{i=0}^n |x - x_i| \right\}$$

sea la más pequeña de entre cualquier otra posible elección de nodos. La respuesta a esta pregunta es afirmativa y su solución se basa en una teoría desarrollada por el matemático ruso P.L. Chebychev que exponemos a continuación.

Definición 4.4 Consideremos la sucesión de polinomios $\{T_n(x)\}_{n \in \mathbb{N}}$ definidos de la forma recurrente siguiente

$$T_0(x) = 1 , \quad T_1(x) = x , \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) .$$

$T_n(x)$ es llamado el polinomio de Chebychev de grado n .

Algunos de los primeros polinomios de Chebychev son $T_2(x) = 2x^2 - 1$, $T_3(x) = 4x^3 - 3x$, etc... Veamos alguna propiedad importante de estos polinomios.

Proposición 4.5 *Los polinomios de Chebychev verifican lo siguiente:*

(i) $T_n(x)$ es un polinomio de grado n del tipo $T_n(x) = 2^{n-1}x^n + \dots$, donde los puntos suspensivos denotan términos de grado inferior a n .

(ii) Para todo $x \in [-1, 1]$ se verifica

$$T_n(x) = \cos(n \arccos x) .$$

(iii) $|T_n(x)| \leq 1$ para todo $x \in [-1, 1]$ y para todo $n \in \mathbb{N}$.

(iv) $T_n(\cos(k\pi/n)) = (-1)^k$ para $k = 0, 1, \dots, n$.

Demostración. La parte (i) se prueba fácilmente por inducción a partir de la Definición 4.4. Demostraremos a continuación la parte (ii). Definamos la función $g_n(x) = \cos(n \arccos x)$ para todo $n \in \mathbb{N}$ y $x \in [-1, 1]$. Vamos a ver que la sucesión de funciones $\{T_n(x)\}_{n \in \mathbb{N}}$ satisface la relación de recurrencia de la Definición 4.4 de modo que $g_n(x) = T_n(x)$. Para ello, recordemos las relaciones trigonométricas siguientes

$$\begin{aligned} \cos[(n+1)\theta] &= \cos(n\theta)\cos(\theta) - \sin(n\theta)\sin(\theta) , \\ \cos[(n-1)\theta] &= \cos(n\theta)\cos(\theta) + \sin(n\theta)\sin(\theta) . \end{aligned}$$

Tomando $\theta = \arccos x$ y sumando las dos anteriores ecuaciones se tiene que

$$g_{n+1}(x) + g_{n-1}(x) = 2 \cos(\theta) \cos(n\theta) = 2xg_n(x) ,$$

es decir, las funciones $g_n(x)$ verifican la misma recurrencia que los $T_n(x)$. Además, satisfacen las dos condiciones iniciales de la recurrencia puesto que

$$g_0(x) = \cos 0 = 1 , \quad g_1(x) = \cos(\arccos x) = x .$$

Entonces, $g_n(x) = T_n(x)$.

Las partes (iii) y (iv) son consecuencias inmediatas de la parte (ii). ■

Un polinomio se llama *mónico* si el coeficiente de la indeterminada de mayor grado es la unidad. Por ejemplo, a partir de la sentencia (i) de la Proposición 4.5, es claro que el polinomio $2^{1-n}T_n(x)$ es mónico siendo $T_n(x)$ el polinomio de Chebychev de grado n . Para este tipo de polinomios se tiene el siguiente resultado.

Proposición 4.6 *Sea $p(x)$ un polinomio mónico de grado n . Entonces,*

$$\|p(x)\|_{\infty} := \max_{-1 \leq x \leq 1} \{|p(x)|\} \geq 2^{1-n} .$$

Demostración. Por reducción al absurdo, supongamos que $|p(x)| < 2^{1-n}$ para todo $x \in [-1, 1]$ y lleguemos a una contradicción. Sea $q(x) = 2^{1-n}T_n(x)$ siendo $T_n(x)$ el polinomio de Chebychev de grado n . Entonces, $q(x)$ es mónico y de

grado n . Definamos los puntos $x_k \in [-1, 1]$ de la forma $x_k = \cos(k\pi/n)$, con $k = 0, 1, \dots, n$. Entonces, se tiene la siguiente cadena de desigualdades

$$(-1)^k p(x_k) \leq |p(x_k)| < 2^{1-n} = (-1)^k q(x_k) ,$$

donde la primera desigualdad es por definición de valor absoluto, la segunda desigualdad es debida a nuestra hipótesis inicial y la última igualdad proviene de utilizar la sentencia (iv) de la Proposición 4.5. Se tiene pues que

$$(-1)^k [q(x_k) - p(x_k)] > 0 , \quad k = 0, 1, \dots, n .$$

De este modo, se comprueba que el polinomio $q(x) - p(x)$ cambia de signo en $n+1$ puntos distintos del intervalo $[a, b]$ y, por el Teorema de Bolzano, $q(x) - p(x)$ tiene al menos n raíces reales en el intervalo $[-1, 1]$. Pero esto es imposible puesto que, al ser $p(x)$ y $q(x)$ mónicos, el polinomio $q(x) - p(x)$ tiene grado $n-1$ a lo sumo. ■

Teorema 4.7 *Sea f una función de clase $C^{n+1}[-1, 1]$, y sea P_n un polinomio de grado menor o igual que n que interpola a la función f en los siguientes $n+1$ nodos $\{x_0, x_1, \dots, x_n\} \subset [-1, 1]$ distintos dos a dos. Entonces, si se eligen los nodos x_k como las raíces del polinomio de Chebychev $T_{n+1}(x)$, es decir, $x_k = \cos((2k+1)\pi/(2n+2))$, con $k = 0, 1, \dots, n$, se minimiza la norma del infinito del error de interpolación $\max_{-1 \leq x \leq 1} \{|f(x) - P_n(x)|\}$ respecto de cualquier otra elección de nodos en $[-1, 1]$.*

Demostración. Se sabe a partir del Teorema 4.2 que, para todo $x \in [-1, 1]$, existe $\xi_x \in (-1, 1)$ tal que

$$f(x) - P_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \pi_n(x) ,$$

siendo $\pi_n(x) = \prod_{i=0}^n (x - x_i)$. Entonces,

$$\max_{-1 \leq x \leq 1} \{|f(x) - P_n(x)|\} \leq \frac{M}{(n+1)!} \max_{-1 \leq x \leq 1} \{|\pi_n(x)|\} ,$$

siendo $M := \max_{-1 \leq x \leq 1} \{|f^{(n+1)}(x)|\}$. Si queremos minimizar el error de interpolación $\max_{-1 \leq x \leq 1} \{|f(x) - P_n(x)|\}$ eligiendo los nodos adecuadamente, hemos de elegirlos de manera que se minimice $\max_{-1 \leq x \leq 1} \{|\pi_n(x)|\}$. Veamos cómo conseguirlo.

De la Proposición 4.6, se sabe que, para cualquier conjunto de nodos obtenemos

$$\max_{-1 \leq x \leq 1} \{|\pi_n(x)|\} \geq 2^{-n} , \quad (4.5)$$

puesto que $\pi_n(x)$ es un polinomio mónico de grado $n+1$ por construcción. Además, de la sentencia (iii) de la Proposición 4.5, se tiene que $|T_{n+1}(x)| \leq 1$

para todo $x \in [-1, 1]$ y para todo $n \in \mathbb{N}$. De este modo obtenemos que el polinomio $2^{-n}T_{n+1}(x)$ mónico de grado $n + 1$ verifica

$$|2^{-n}T_{n+1}(x)| \leq 2^{-n}, \quad \text{para todo } x \in [-1, 1]. \quad (4.6)$$

A partir de (4.5) y (4.6) se tiene que el valor mínimo de $\max_{-1 \leq x \leq 1} \{|\pi_n(x)|\}$ se obtiene cuando

$$\pi_n(x) := \prod_{i=0}^n (x - x_i) = 2^{-n}T_{n+1}(x),$$

es decir, los nodos x_k a elegir son las raíces del polinomio de Chebychev $T_{n+1}(x)$. Es además obvio que dichas raíces son $x_k = \cos((2k + 1)\pi/(2n + 2))$ a partir de la expresión $T_{n+1}(x) = \cos((n + 1) \arccos x)$. ■

Notemos que el Teorema 4.7 está formulado para realizar una interpolación con nodos únicamente en el intervalo $[-1, 1]$. Sin embargo, no existe pérdida de generalidad, puesto que el intervalo $x \in [-1, 1]$ se transforma en el intervalo $y \in [a, b]$ mediante el cambio lineal de variables

$$y = \frac{a+b}{2} + \frac{b-a}{2}x.$$

De este modo es fácil mostrar la siguiente generalización del Teorema 4.7

Teorema 4.8 *Dada una función $f \in C^{n+1}[a, b]$, el menor error en la norma del máximo en la interpolación de f se comete cuando se toman las siguientes abscisas de interpolación $\{y_0, y_1, \dots, y_n\} \subset [a, b]$*

$$y_j = \frac{a+b}{2} + \frac{b-a}{2} \cos \frac{(2j+1)\pi}{2(n+1)}, \quad j = 0, 1, \dots, n.$$

En general, no es aconsejable efectuar interpolación polinómica con muchas abscisas de interpolación x_0, x_1, \dots, x_n . Dicho de otro modo, no es conveniente tener un polinomio de interpolación $P_n(x)$ cuyo grado n sea elevado. Una de las razones de esta afirmación es la siguiente. Supongamos que la función $f(x)$ es continua y que $P_n(x)$ es su polinomio interpolador en las abscisas $a = x_0, x_1, \dots, x_n = b$. Entonces, en general, no es cierto que se verifique la convergencia puntual siguiente

$$\lim_{n \rightarrow \infty} P_n(x) = f(x) \quad \forall x \in [a, b].$$

Esto se conoce como *fenómeno Runge*. Una forma de evitarlo es la llamada interpolación de splines.

Una función *spline* está formada por varios polinomios, cada uno definido en un subintervalo, que se unen entre sí obedeciendo a ciertas condiciones de continuidad.

Más concretamente, supongamos que tenemos definidas $n + 1$ abscisas (llamadas *nudos*) $x_0 < x_1 < \dots < x_n$. Definimos una *función spline de grado k* sobre los nudos x_0, x_1, \dots, x_n como una función S satisfaciendo las siguientes propiedades

- (i) S es un polinomio de grado menor o igual que k en cada intervalo $[x_{i-1}, x_i]$.
- (ii) $S \in C^{k-1}[x_0, x_n]$.

Dada una tabulación $(x_i, f(x_i))$ con $i = 0, 1, \dots, n$ la función Spline debe interpolar a la función f en dichos puntos, es decir, $S(x_i) = f(x_i)$. Entonces

$$S(x) = \begin{cases} S_0(x) & x \in [x_0, x_1] , \\ S_1(x) & x \in [x_1, x_2] , \\ \vdots & \vdots \\ S_{n-1}(x) & x \in [x_{n-1}, x_n] , \end{cases} \quad (4.7)$$

En la práctica los splines más utilizados son los cúbicos ($k = 3$). Se puede ver que, en la construcción de splines cúbicos existen 2 grados de libertad. Si se toma $S''(x_0) = S''(x_n) = 0$ los splines se llaman *splines cúbicos naturales*.

4.2. Problemas resueltos

Problema 4.1 *A partir de la siguiente tabulación de la función $f(x) = \ln x$*

x_i	0.40	0.50	0.70	0.80
f_i	-0.916291	-0.693147	-0.356675	-0.223144

- (i) *Aproximar el valor de $\ln 0,6$ utilizando interpolación cúbica de Lagrange.*
- (ii) *Acotar el error cometido en la aproximación realizada en el apartado anterior (sin utilizar la calculadora para averiguar el valor exacto de $\ln 0,6$).*

Solución. (i) La fórmula de interpolación de Lagrange viene dada por

$$P_n(x) = \sum_{k=0}^n f_k \ell_k(x) , \quad \ell_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^n \frac{x - x_j}{x_k - x_j} , \quad k = 0, \dots, n.$$

En este problema $n = 3$, de manera que los polinomios de Lagrange son

$$\begin{aligned} \ell_0(x) &= \frac{(x - 0,5)(x - 0,7)(x - 0,8)}{(0,4 - 0,5)(0,4 - 0,7)(0,4 - 0,8)} \\ &= -83,3333x^3 + 166,667x^2 - 109,167x + 23,3333 \\ \ell_1(x) &= \frac{(x - 0,4)(x - 0,7)(x - 0,8)}{(0,5 - 0,4)(0,5 - 0,7)(0,5 - 0,8)} \\ &= 166,667x^3 - 316,667x^2 + 193,333x - 37,3333 \end{aligned}$$

$$\begin{aligned}
\ell_2(x) &= \frac{(x-0,4)(x-0,5)(x-0,8)}{(0,7-0,4)(0,7-0,5)(0,7-0,8)} \\
&= -166,667x^3 + 283,333x^2 - 153,333x + 26,6667 \\
\ell_3(x) &= \frac{(x-0,4)(x-0,5)(x-0,7)}{(0,8-0,4)(0,8-0,5)(0,8-0,7)} \\
&= 83,3333x^3 - 133,333x^2 + 69,1667x - 11,6667 .
\end{aligned}$$

Tenemos que el polinomio interpolador es

$$P_3(x) = -0,916291\ell_0(x) - 0,693147\ell_1(x) - 0,356675\ell_2(x) - 0,223144\ell_3(x) .$$

Finalmente, realizamos la aproximación

$$\begin{aligned}
\ln 0,6 \approx P_3(0,6) &= -0,916291 \left(\frac{-1}{6} \right) - 0,693147 \left(\frac{2}{3} \right) \\
&\quad - 0,356675 \left(\frac{2}{3} \right) - 0,223144 \left(\frac{-1}{6} \right) = -0,509975 .
\end{aligned}$$

(ii) Sabemos que existe $\xi \in (0,4, 0,8)$ tal que el error de interpolación viene dado por

$$\ln 0,6 - P_3(0,6) = \frac{f^{(iv)}(\xi)}{4!} \prod_{i=0}^3 (0,6 - x_i) .$$

Como $f^{(iv)}(\xi) = \frac{-6}{\xi^4}$, obtenemos un error

$$\ln 0,6 - P_3(0,6) = \frac{-0,0004}{4} \frac{1}{\xi^4} .$$

Finalmente, debido a que

$$0,4 < \xi < 0,8 \implies \frac{1}{0,8} < \frac{1}{\xi} < \frac{1}{0,4} \implies \frac{1}{0,8^4} < \frac{1}{\xi^4} < \frac{1}{0,4^4} ,$$

podemos acotar el error de la forma

$$|\ln 0,6 - P_3(0,6)| = \frac{0,0004}{4} \left| \frac{1}{\xi^4} \right| < \frac{0,0004}{4} \frac{1}{0,4^4} = 0,00390625 .$$

Problema 4.2 Considerar una función $f : [a, b] \rightarrow \mathbb{R}$ de clase $\mathcal{C}^2[a, b]$. Sean tres puntos distintos $x_0, x_1, x_2 \in [a, b]$. Demostrar que existe un único polinomio $P(x)$ de grado menor o igual que tres verificando

$$P(x_0) = f(x_0) , \quad P(x_2) = f(x_2) , \quad P''(x_1) = f''(x_1) , \quad P''(x_2) = f''(x_2) .$$

Solución. Sea $P(x) \in \mathbb{R}_3[x]$ un polinomio de grado menor o igual que tres, de manera que $P(x) = a_0 + a_1x + a_2x^2 + a_3x^3$, $P'(x) = a_1 + 2a_2x + 3a_3x^2$ y

$P''(x) = 2a_2 + 6a_3x$. Imponiendo las condiciones del enunciado obtenemos el siguiente sistema lineal de 4 ecuaciones para las incógnitas a_0, a_1, a_2, a_3

$$\begin{aligned} P(x_0) = f(x_0) &\implies a_0 + a_1x_0 + a_2x_0^2 + a_3x_0^3 = f(x_0) \\ P(x_2) = f(x_2) &\implies a_0 + a_1x_2 + a_2x_2^2 + a_3x_2^3 = f(x_2) \\ P''(x_1) = f''(x_1) &\implies 2a_2 + 6a_3x_1 = f''(x_1) \\ P''(x_2) = f''(x_2) &\implies 2a_2 + 6a_3x_2 = f''(x_2) \end{aligned}$$

El determinante de la matriz A de coeficientes del sistema es

$$\begin{aligned} \det A &= \begin{vmatrix} 1 & x_0 & x_0^2 & x_0^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 0 & 0 & 2 & 6x_1 \\ 0 & 0 & 2 & 6x_2 \end{vmatrix} = \begin{vmatrix} 1 & x_0 & x_0^2 & x_0^3 \\ 0 & x_2 - x_0 & x_2^2 - x_0^2 & x_2^3 - x_0^3 \\ 0 & 0 & 2 & 6x_1 \\ 0 & 0 & 0 & 6(x_2 - x_1) \end{vmatrix} \\ &= 12(x_2 - x_1)(x_2 - x_0) . \end{aligned}$$

Como los tres puntos x_0, x_1, x_2 son distintos dos a dos, tenemos que $\det A \neq 0$ y el sistema lineal es compatible y determinado con lo que demostramos la existencia y unicidad del polinomio $P(x)$.

Problema 4.3 Consideremos una función $f \in C^{n+1}[x_0, x_n]$. Sea $P(x)$ el polinomio que interpola a la función $f(x)$ en los nudos $x_0 < x_1 < \dots < x_n$. Consideremos un valor $\bar{x} \in (x_0, x_n)$ que no coincida con ninguno de los nudos anteriores, es decir, $\bar{x} \neq x_i$ para $i = 0, 1, \dots, n$.

(i) Utilizando el formalismo de diferencias divididas de Newton, demostrar que

$$f(\bar{x}) - P(\bar{x}) = f[x_0, x_1, \dots, x_n, \bar{x}] w(\bar{x}) , \quad \text{siendo} \quad w(x) = \prod_{i=0}^n (x - x_i) .$$

(ii) Hallar una relación entre la diferencia dividida $f[x_0, x_1, \dots, x_n]$ y la n -ésima derivada de la función f .

Solución. (i) Utilizando el formalismo de diferencias divididas de Newton, el polinomio interpolador P adopta la forma

$$P(x) = f[x_0] + f[x_0, x_1](x - x_0) + \dots + f[x_0, x_1, \dots, x_n](x - x_0) \cdots (x - x_{n-1}) .$$

Definamos $\bar{P}(x)$ como el polinomio que interpola a la función $f(x)$ en los nudos $x_0, x_1, \dots, x_n, \bar{x}$. Por las propiedades del polinomio interpolador de Newton se sabe que

$$\bar{P}(x) = P(x) + f[x_0, x_1, \dots, x_n, \bar{x}] w(x) .$$

Particularizando esta ecuación para $x = \bar{x}$ se obtiene

$$f(\bar{x}) = \bar{P}(\bar{x}) = P(\bar{x}) + f[x_0, x_1, \dots, x_n, \bar{x}] w(\bar{x}) ,$$

de lo que se deduce

$$f(\bar{x}) - P(\bar{x}) = f[x_0, x_1, \dots, x_n, \bar{x}] w(\bar{x}) .$$

(ii) Por el teorema del error interpolatorio se sabe que

$$f(\bar{x}) - P(\bar{x}) = \frac{f^{(n+1)}(\xi)}{(n+1)!} w(\bar{x}) ,$$

siendo $\xi \in (x_0, x_n)$. Comparando este resultado con el obtenido en el apartado

(i) concluimos que

$$f[x_0, x_1, \dots, x_n, \bar{x}] = \frac{f^{(n+1)}(\xi)}{(n+1)!} .$$

Tenemos pues, por la simetría de las diferencias divididas, que

$$f[x_0, x_1, \dots, x_n] = \frac{f^{(n)}(\xi)}{n!} .$$

Problema 4.4 Sea $P_n(x)$ el polinomio de grado menor o igual que n que interpola a la función $f(x) = \sinh x$ en cualquier conjunto de $n+1$ abscisas en el intervalo $[-1, 1]$.

- (i) Suponiendo que n es impar, hallar una cota del error absoluto $|f(x) - P_n(x)|$ para $x \in [-1, 1]$ que sólo dependa de n .
- (ii) Calcular, por el método de las diferencias divididas de Newton, el polinomio que interpola a la siguiente tabulación.

x_i	-1	0	1
$\sinh(x_i)$	-1.1752	0.0000	1.1752

Solución. (i) El error absoluto es

$$|f(x) - P_n(x)| = \left| \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{i=0}^n (x - x_i) \right| .$$

Como n es impar, $n+1$ es par y por lo tanto $f^{(n+1)}(\xi_x) = \sinh \xi_x$. Además, como $\xi_x \in (-1, 1)$ y la función $\sinh x$ es impar y creciente se tiene que $|\sinh \xi_x| \leq \sinh 1$ de manera que

$$|f(x) - P_n(x)| \leq \frac{\sinh 1}{(n+1)!} \left| \prod_{i=0}^n (x - x_i) \right| .$$

Finalmente, como $x, x_i \in [-1, 1]$ para $i = 0, 1, \dots, n$ se tiene que $|x - x_i| \leq 2$ para $i = 0, 1, \dots, n$ de manera que

$$|f(x) - P_n(x)| \leq \frac{\sinh 1}{(n+1)!} 2^{n+1} .$$

Notemos que $\lim_{n \rightarrow \infty} |f(x) - P_n(x)| = 0$.

(ii) Utilizando el formalismo de diferencias divididas de Newton, el polinomio interpolador P_2 adopta la forma

$$P_2(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) .$$

Realizando el esquema de diferencias divididas se obtiene

$$f[x_0] = -1,1752 , \quad f[x_0, x_1] = 1,1752 , \quad f[x_0, x_1, x_2] = 0 ,$$

de manera que

$$P_2(x) = -1,1752 + 1,1752(x + 1) = 1,1752x .$$

Problema 4.5 *Se quiere construir una tabulación equiespaciada de la función $f(x) = (x^4 - x)/12$ en el intervalo $[0, a]$ de manera que, si no se tienen en cuenta los errores de redondeo, el error absoluto cometido en cualquier interpolación lineal efectuada con dos abscisas consecutivas de dicha tabulación sea inferior a un valor dado $\epsilon > 0$. Calcular el mínimo número de abscisas en la tabulación n en función de a y de ϵ .*

Solución. Se ha de hallar la tabulación equiespaciada $\{0 = x_0, x_1, \dots, x_n = a\}$ de la función $f(x)$. Sean x_j y x_{j+1} dos puntos del intervalo $[0, a]$. Sea $x \in [x_j, x_{j+1}]$. El error absoluto Δ cometido en una interpolación lineal de la función f en las abscisas consecutivas x_j y x_{j+1} viene dado por

$$\Delta = \left| \frac{f''(\xi)}{2} (x - x_j)(x - x_{j+1}) \right| ,$$

siendo $\xi \in [x_j, x_{j+1}]$. Definimos la longitud del paso de tabulación $h = x_{j+1} - x_j$. Puesto que $x \in [x_j, x_{j+1}]$ es evidente que $|x - x_i| \leq h$ para $i = j, j + 1$. Como $f''(\xi) = \xi^2$, podemos acotar el error absoluto de la forma

$$\Delta = \left| \frac{f''(\xi)}{2} (x - x_j)(x - x_{j+1}) \right| = \frac{\xi^2}{2} |x - x_j| |x - x_{j+1}| \leq \frac{\xi^2}{2} h^2 .$$

Finalmente, como $\xi \in [x_j, x_{j+1}] \subseteq [0, a]$ es evidente que $\xi \leq a$ de manera que

$$\Delta \leq \frac{h^2 a^2}{2} < \epsilon ,$$

de lo que se deduce $h < \sqrt{2\epsilon}/a$. Finalmente, puesto que en la tabulación $\{0 = x_0, x_1, \dots, x_n = a\}$ se tiene $h = a/n$, el mínimo número de abscisas n viene dado por $n > a^2/\sqrt{2\epsilon}$. Es decir

$$n = E \left[\frac{a^2}{\sqrt{2\epsilon}} \right] + 1 ,$$

donde $E[\]$ denota la parte entera.

Problema 4.6 Estudiar de manera razonada la existencia y unicidad del polinomio interpolador $P(x)$ con coeficientes reales en los casos siguientes:

- (i) $\text{gr}(P) = 2$, $P(0) = 0$, $P(1) = 1$ y $P'(1/2) = 2$.
- (ii) $\text{gr}(P) = 3$, $P(0) = 0$, $P(1) = 1$ y $P'(1/2) = 2$.
- (iii) $\text{gr}(P) \leq n$, $P(x_k) = f_k$ con $x_k, f_k \in \mathbb{R}$ conocidos para $k = 0, 1, \dots, n$ y $x_i \neq x_j$ si $i \neq j$.
- (iv) $\text{gr}(P) = n + 1$, $P(x_k) = f_k$ con $x_k, f_k \in \mathbb{R}$ conocidos para $k = 0, 1, \dots, n$ y $x_i \neq x_j$ si $i \neq j$.
- (v) $\text{gr}(P) \leq 2n + 1$, $P(x_k) = f_k$, $P'(x_k) = f'_k$ con $x_k, f_k, f'_k \in \mathbb{R}$ conocidos para $k = 0, 1, \dots, n$.

Solución. (i) Sea $P(x) = a + bx + cx^2$ con $a, b, c \in \mathbb{R}$. Imponiendo la condición de interpolación $P(0) = 0$ se tiene $a = 0$. Del restante de condiciones $P(1) = 1$ y $P'(1/2) = 2$ se construye el sistema lineal

$$1 = b + c, \quad 2 = b + c,$$

que evidentemente es incompatible. Por lo tanto no existe el polinomio $P(x)$.

(ii) Sea $P(x) = a + bx + cx^2 + dx^3$ con $a, b, c, d \in \mathbb{R}$. Imponiendo $P(0) = 0$ se tiene $a = 0$. De imponer $P(1) = 1$ y $P'(1/2) = 2$ se obtiene el sistema lineal

$$1 = b + c + d, \quad 2 = b + c + \frac{3}{4}d,$$

que es compatible indeterminado. Sus soluciones son de la forma $d = -4$ y $b + c = 5$ de manera que existen infinitos polinomios $P(x)$ todos ellos de la forma

$$P(x) = (5 - c)x + cx^2 - 4x^3, \quad c \in \mathbb{R}.$$

(iii) En este caso está garantizada la existencia y la unicidad del polinomio $P(x)$ puesto que se está realizando una interpolación de Lagrange.

(iv) Puesto que si se sustituye la condición $\text{gr}(P) = n + 1$ por $\text{gr}(P) \leq n$ se está en las condiciones de interpolación de Lagrange y se sabe que en este caso el polinomio $P(x)$ existe y es único, es obvio que con la condición $\text{gr}(P) = n + 1$ también existirá pero no será único.

(v) En este caso está garantizada la existencia y la unicidad del polinomio $P(x)$ debido a que se está realizando una interpolación de Hermite.

Problema 4.7 Considerar la siguiente tabulación de la función $f(x) = \ln x$.

x_i	1	2	3
$f(x_i)$	0	0.693147	1.09861

Calcular una cota del error cometido si se aproxima $\ln 1,5$ a partir de una interpolación polinomial con todos los datos de la tabla adjunta.

Solución. Sea $P_2(x)$ el polinomio de grado menor o igual que 2 que interpola a todos los puntos de la tabulación del enunciado, es decir, $f(x_i) = P_2(x_i)$ para $i = 0, 1, 2$. Como $f(x) \in C^3[1, 3]$, si se realiza la aproximación $f(x) \approx P_2(x)$ para cualquier $x \in [1, 3]$, el error Δ cometido viene dado por

$$\Delta = |f(x) - P_2(x)| = \left| \frac{f^{(3)}(\xi_x)}{3!} (x-1)(x-2)(x-3) \right| ,$$

siendo $\xi_x \in (x_0, x_2) = (1, 3)$. Puesto que $f(x) = \ln x$, se tiene que $f^{(3)}(\xi_x) = 2/\xi_x^3$. Por otra parte, la aproximación que se quiere efectuar es $\ln 1,5 = f(1,5) \approx P_2(1,5)$, de manera que el error cometido es

$$\Delta = \left| \frac{2/\xi_x^3}{3!} (1,5-1)(1,5-2)(1,5-3) \right| = \left| \frac{0,375}{3\xi_x^3} \right| .$$

Como $1/\xi_x^3$ es una función decreciente en el intervalo $\xi \in (1, 3)$, se concluye que

$$\Delta < \frac{0,375}{3} = 0,125 .$$

Problema 4.8 Se ha de realizar una interpolación polinomial cuadrática de la función $f(x) = e^x \sin x$ para valores de $x \in [0, 1]$. Hallar una cota superior del error absoluto que se cometería al aproximar la función f por el polinomio interpolador.

Solución. Puesto que la función $f(x) = e^x \sin x$ es infinitas veces derivable en \mathbb{R} , en particular $f \in C^3[0, 1]$. Por lo tanto, el error absoluto Δ que se comete al realizar la aproximación $f(x) \approx P_2(x)$ para cualquier valor $x \in [0, 1]$, siendo $P_2(x)$ el polinomio de grado menor o igual que 2 que interpola a f en los nodos $\{x_0, x_1, x_2\} \subset [0, 1]$, viene dado por

$$\Delta = \left| \frac{f'''(\xi_x)}{3!} \pi(x) \right| ,$$

siendo $\pi(x) = (x-x_0)(x-x_1)(x-x_2)$ y $\xi_x \in (0, 1)$. Utilizando las propiedades del valor absoluto se tiene

$$\Delta = \frac{|f'''(\xi_x)|}{3!} |\pi(x)| .$$

Puesto que $x \in [0, 1]$, es claro que $|x - x_i| \leq 1$ para $i = 0, 1, 2$, de manera que

$$|\pi(x)| = |x - x_0||x - x_1||x - x_2| \leq 1 .$$

En definitiva

$$\Delta \leq \frac{|f'''(\xi_x)|}{3!},$$

donde la tercera derivada de la función f viene dada por

$$f'''(\xi_x) = 2e^{\xi_x}(\cos \xi_x - \sin \xi_x).$$

Es fácil ver que la función $f'''(\xi_x)$ es decreciente en el intervalo $(0, 1)$ puesto que $f^{iv}(\xi_x) = -4e^{\xi_x} \sin \xi_x < 0$ para todo $\xi_x \in (0, 1)$. De este modo se tiene que los extremos absolutos de $f'''(\xi_x)$ en el compacto $[0, 1]$ ocurren en los extremos de dicho intervalo. Puesto que $f'''(0) = 2$ y $f'''(1) = -1,63$ se tiene la siguiente acotación del error

$$\Delta < \frac{|f'''(0)|}{3!} = \frac{2}{3!} = \frac{1}{3}.$$

Problema 4.9 *Se han realizado las siguientes mediciones de la capacidad calorífica molar a presión constante c del nitrógeno para diferentes temperaturas T obteniéndose*

T_k (K)	300	700	1500
c_k (cal/gr. mol K)	6.81495	7.40495	8.25375

- (i) Hallar la expresión polinomial de la función $c(T)$ a partir de una interpolación polinomial mediante el método de las diferencias divididas de Newton con los datos de la tabla adjunta.
- (ii) Si el termómetro utilizado tiene una escala que varía entre 300 y 1500 grados Kelvin, obtener las 3 temperaturas donde se realizarían las mediciones de c sabiendo que posteriormente se hará una interpolación polinomial con esas mediciones.

Solución. (i) Con los datos de la tabla adjunta, es decir, la tabulación $c(T_k)$ para $k = 0, 1, 2$, se obtiene el siguiente esquema triangular de diferencias divididas de Newton

300	6.81495		
		0.001475	
700	7.40495		-3.45×10^{-7}
		0.001061	
1500	8.25375		

El polinomio $P_2(T)$ que interpola a la función $c(T)$ según los datos de la tabla es

$$\begin{aligned} P_2(T) &= c[T_0] + c[T_0, T_1](T - T_0) + c[T_0, T_1, T_2](T - T_0)(T - T_1) \\ &= 6,81495 + 0,001475(T - 300) - 3,45 \times 10^{-7}(T - 300)(T - 700). \end{aligned}$$

Por lo tanto $c(T) \approx P_2(T) = 6,3 + 1,82 \times 10^{-3}T - 0,345 \times 10^{-6}T^2$.

(ii) El rango de temperaturas T donde se pueden realizar las mediciones es $T \in [300, 1500]$. Por lo tanto, utilizando la teoría de los nodos de Chebychev, las mediciones se realizarán en las temperaturas

$$T_k = \frac{300 + 1500}{2} + \frac{1500 - 300}{2} x_k ,$$

siendo los nodos de Chebychev $x_k = \cos[(2k+1)\pi/(2n+2)]$ con $k = 0, 1, \dots, n$. Realizaremos tres mediciones y por lo tanto $n = 2$. En definitiva, las temperaturas son

$$T_0 = 900 + 600 \cos\left(\frac{\pi}{6}\right) = 900 + 300\sqrt{3} \approx 1419,62 ,$$

$$T_1 = 900 + 600 \cos\left(\frac{\pi}{2}\right) = 900 ,$$

$$T_2 = 900 + 600 \cos\left(\frac{5\pi}{6}\right) = 900 - 300\sqrt{3} \approx 380,385 .$$

Problema 4.10 (i) Sea $P_n(x)$ el único polinomio de grado menor o igual que n que interpola a la función $f(x)$ en los nodos equiespaciados $x_0 < x_1 < \dots < x_n$, con longitud de paso h . Demostrar que, si $f \in C^{n+1}[x_0, x_n]$, el error absoluto $\Delta := |f(x) - P_n(x)|$ para cualquier $x \in [x_0, x_n]$ viene dado por

$$\Delta = \binom{t}{n+1} h^{n+1} |f^{(n+1)}(\xi)| , \quad \xi \in (x_0, x_n) ,$$

siendo $t = (x - x_0)/h$ y el número combinatorio $\binom{a}{b} = \frac{a!}{b!(a-b)!}$.

(ii) Supongamos que se quiere aproximar el polinomio $f(x) = x^4(5-x)/60$ mediante un polinomio interpolador $P_2(x)$ de grado menor o igual que 2 en los nodos equiespaciados $x_i = i$ con $i = 0, 1, 2$. Hallar una cota del error absoluto Δ cometido en la aproximación $f(x) \approx P_2(x)$ para cualquier $x \in [0, 2]$.

Solución. (i) Partiendo del error cometido por interpolación se sabe que existe $\xi \in (x_0, x_n)$ tal que

$$\Delta = \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} \pi(x) \right| ,$$

siendo $\pi(x) = \prod_{i=0}^n (x - x_i)$. Ya que los nodos son equiespaciados con longitud de paso h , se tiene que $x_i = x_0 + ih$ con $i = 0, 1, \dots, n$. Además, por definición se tiene que $x = x_0 + ht$, de manera que $x - x_0 = th$, $x - x_1 = x - (x_0 + h) = th - h = (t-1)h$. En general se obtiene $x - x_i = (t-i)h$ para $i = 0, 1, \dots, n$, de manera que

$$\begin{aligned} \Delta &= \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^n (t-i)h \right| = \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+1} t(t-1)(t-2) \cdots (t-n) \right| \\ &= \binom{t}{n+1} h^{n+1} |f^{(n+1)}(\xi)| . \end{aligned}$$

(ii) Podemos utilizar la fórmula hallada en el apartado anterior. En primer lugar, se observa que si $x = x_0 + th$ y $x \in [x_0, x_n]$ entonces $t \in [0, n]$ y por lo tanto $t - i \leq n$ para $i = 0, 1, \dots, n$. De este modo, se tiene que

$$\begin{aligned} \Delta &= \binom{n}{n+1} h^{n+1} |f^{(n+1)}(\xi)| = \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+1} \prod_{i=0}^n (t-i) \right| \\ &\leq \frac{|f^{(n+1)}(\xi)|}{(n+1)!} h^{n+1} n^{n+1}. \end{aligned}$$

En nuestro problema se tiene $n = 2$ y $h = 1$, por lo que

$$\Delta \leq \frac{|f'''(\xi)|}{3!} 2^3 = \frac{4}{3} |f'''(\xi)|,$$

siendo $\xi \in (0, 2)$. Calculando la tercera derivada de la función $f(x) = x^4(5 - x)/60$ se obtiene $f'''(\xi) = -(\xi - 1)^2 + 1$. La gráfica de $f'''(\xi)$ es una parábola con un máximo relativo en $\xi = 1$, de manera que podemos realizar una última acotación de la forma

$$\Delta \leq \frac{4}{3} f'''(1) = \frac{4}{3} = 1,3333.$$

NOTA: En la acotación efectuada se ha utilizado que $|\pi(t)| := |t(t-1)(t-2)| < 2^3$ para todo $t \in [0, 2]$. Existe, por supuesto, otra forma mejor de acotar $|\pi(t)|$ para todo $t \in [0, 2]$ que consiste en obtener su máximo absoluto en el intervalo compacto $[0, 2]$. Así, ya que las raíces del polinomio $\pi'(t) = 2 - 6t + 3t^2$ son $t_{\pm} = 1 \pm 1/\sqrt{3} \in [0, 2]$, es claro que $\pi(t)$ tiene extremos relativos en t_{\pm} . Además, para obtener los extremos absolutos calculamos los valores de $\pi(t)$ en los extremos del intervalo cerrado $\pi(0) = \pi(2) = 0$ y en los extremos relativos $|\pi(t_{\pm})| = 0,3849$. Se tiene en definitiva que $|\pi(t)| \leq 0,3849$ para todo $t \in [0, 2]$ y por lo tanto una mejor cota del error

$$\Delta \leq \frac{|f'''(\xi)|}{3!} 0,3849 < \frac{|f'''(1)|}{3!} 0,3849 = \frac{0,3849}{3!} = 0,06415.$$

Problema 4.11 Sean $f(x)$ y $g(x)$ dos funciones reales definidas sobre el intervalo $[a, b]$. Definamos $P_n(x)$ y $Q_n(x)$ como los polinomios de grado menor o igual que n que interpolan a f y g respectivamente en los nodos $\{x_0, x_1, \dots, x_n\} \subset [a, b]$ distintos entre sí.

- (i) ¿Es $\alpha P_n + \beta Q_n$ el polinomio de grado menor o igual que n que interpola a la función $\alpha f + \beta g$ en los nodos $\{x_0, x_1, \dots, x_n\}$, siendo $\alpha, \beta \in \mathbb{R}$?
- (ii) ¿Es $P_n Q_n$ el polinomio de grado menor o igual que n que interpola a la función fg en los nodos $\{x_0, x_1, \dots, x_n\}$?

Solución. Por definición, si $P_n(x)$ y $Q_n(x)$ son los polinomios que interpolan a f y g respectivamente en los nodos $\{x_0, x_1, \dots, x_n\}$, entonces, por ser los nodos

distintos entre sí, $P_n(x)$ y $Q_n(x)$ son polinomios de grado menor o igual que n , es decir, $P_n, Q_n \in \mathbb{R}_n[x]$. Además $P_n(x_i) = f(x_i)$ y $Q_n(x_i) = g(x_i)$ para $i = 0, 1, \dots, n$.

(i) La respuesta es positiva puesto que $\alpha P_n + \beta Q_n \in \mathbb{R}_n[x]$ y además $\alpha P_n(x_i) + \beta Q_n(x_i) = \alpha f(x_i) + \beta g(x_i)$ para $i = 0, 1, \dots, n$.

(ii) Aunque se verifique $P_n(x_i)Q_n(x_i) = f(x_i)g(x_i)$ para $i = 0, 1, \dots, n$, la respuesta es, en general, negativa puesto que $P_n Q_n \in \mathbb{R}_{2n}[x]$ no tiene grado menor o igual que n .

Problema 4.12 *Averiguar si la función*

$$S(x) = \begin{cases} 1 + x - x^3 & \text{si } x \in [0, 1] , \\ 1 - 2(x-1) - 3(x-1)^2 + 4(x-1)^3 & \text{si } x \in [1, 2] , \\ 4(x-2) + 9(x-2)^2 - 3(x-2)^3 & \text{si } x \in [2, 3] , \end{cases}$$

es el spline cúbico natural que interpola a los puntos $(0, 1)$, $(1, 1)$, $(2, 0)$ y $(3, 10)$.

Solución. Recordemos que la función

$$S(x) = \begin{cases} s_0(x) & \text{si } x \in [x_0, x_1] , \\ s_1(x) & \text{si } x \in [x_1, x_2] , \\ s_2(x) & \text{si } x \in [x_2, x_3] , \end{cases}$$

es el spline cúbico natural que interpola a los puntos (x_0, f_0) , (x_1, f_1) , (x_2, f_2) y (x_3, f_3) si se verifica lo siguiente: (i) $s_i(x) \in \mathbb{R}_3[x]$ para $i = 0, 1, 2$; (ii) $S \in C^2[x_0, x_2]$; (iii) $S(x_i) = f_i$ para $i = 0, 1, 2$; (iv) $S''(x_0) = S''(x_2) = 0$.

- En nuestro caso se tiene $s_0(x) = 1 + x - x^3$, $s_1(x) = 1 - 2(x-1) - 3(x-1)^2 + 4(x-1)^3$ y $s_2(x) = 4(x-2) + 9(x-2)^2 - 3(x-2)^3$ que verifican obviamente la condición (i).
- Los puntos de interpolación son $(x_0, f_0) = (0, 1)$, $(x_1, f_1) = (1, 1)$, $(x_2, f_2) = (2, 0)$ y $(x_3, f_3) = (3, 10)$, de modo que también se verifica la condición (iii) ya que

$$\begin{aligned} S(x_0) &= S(0) = s_0(0) = 1 = f_0 , \\ S(x_1) &= S(1) = s_0(1) = s_1(1) = 1 = f_1 , \\ S(x_2) &= S(2) = s_1(2) = s_2(2) = 0 = f_2 , \\ S(x_3) &= S(3) = s_2(3) = 10 = f_3 . \end{aligned}$$

Además, de aquí también se ve que $S \in C[0, 3]$

- Para comprobar la condición (ii) sólo es necesario demostrar que la función $S(x)$ es derivable dos veces con continuidad en los puntos donde dicha función se parte, es decir, en $x = 1$ y $x = 2$.

Calculamos los siguientes límites laterales:

$$\begin{aligned}\lim_{x \rightarrow 1^-} S'(x) &= \lim_{x \rightarrow 1^-} s'_0(x) = \lim_{x \rightarrow 1^-} 1 - 3x^2 = 2, \\ \lim_{x \rightarrow 1^+} S'(x) &= \lim_{x \rightarrow 1^+} s'_1(x) = \lim_{x \rightarrow 1^+} -2 - 6(x-1) + 12(x-1)^2 \\ &= -2.\end{aligned}$$

Como $\lim_{x \rightarrow 1^-} S'(x) = \lim_{x \rightarrow 1^+} S'(x)$ se tiene que $S(x)$ es derivable con continuidad en $x = 1$.

$$\begin{aligned}\lim_{x \rightarrow 2^-} S'(x) &= \lim_{x \rightarrow 2^-} s'_1(x) = \lim_{x \rightarrow 2^-} -2 - 6(x-1) + 12(x-1)^2 = 4, \\ \lim_{x \rightarrow 2^+} S'(x) &= \lim_{x \rightarrow 2^+} s'_2(x) = \lim_{x \rightarrow 2^+} 4 + 18(x-2) - 9(x-2)^2 \\ &= 4.\end{aligned}$$

Como $\lim_{x \rightarrow 2^-} S'(x) = \lim_{x \rightarrow 2^+} S'(x)$ se tiene que $S(x)$ es derivable con continuidad en $x = 2$.

$$\begin{aligned}\lim_{x \rightarrow 1^-} S''(x) &= \lim_{x \rightarrow 1^-} s''_0(x) = \lim_{x \rightarrow 1^-} -6x = -6, \\ \lim_{x \rightarrow 1^+} S''(x) &= \lim_{x \rightarrow 1^+} s''_1(x) = \lim_{x \rightarrow 1^+} -6 + 24(x-1) \\ &= -6.\end{aligned}$$

Como $\lim_{x \rightarrow 1^-} S''(x) = \lim_{x \rightarrow 1^+} S''(x)$ se tiene que $S(x)$ es derivable dos veces con continuidad en $x = 1$.

$$\begin{aligned}\lim_{x \rightarrow 2^-} S''(x) &= \lim_{x \rightarrow 2^-} s''_1(x) = \lim_{x \rightarrow 2^-} -6 + 24(x-1) = 18, \\ \lim_{x \rightarrow 2^+} S''(x) &= \lim_{x \rightarrow 2^+} s''_2(x) = \lim_{x \rightarrow 2^+} 18 - 18(x-2) \\ &= 18.\end{aligned}$$

Como $\lim_{x \rightarrow 2^-} S''(x) = \lim_{x \rightarrow 2^+} S''(x)$ se tiene que $S(x)$ es derivable dos veces con continuidad en $x = 2$.

Se concluye que la función S verifica la condición (ii). Finalmente, puesto que $S''(x_0) = S''(0) = s''_0(0) = 0$ y $S''(x_3) = S''(3) = s''_2(3) = 0$ también se cumple la condición (iv). En resumen, $S(x)$ es el spline cúbico natural que interpola a los puntos $(0, 1)$, $(1, 1)$, $(2, 0)$ y $(3, 10)$.

Problema 4.13 Sea $P_n(x)$ el polinomio de grado menor o igual que n que interpola a la función $f(x) = \sinh x$ en $n + 1$ puntos del intervalo $[-2, 0]$. Si n es impar, hallar una cota que sólo dependa de n del error absoluto $\Delta_n(x)$ cometido en la aproximación $f(x) \approx P_n(x)$ para cualquier $x \in [-2, 0]$. ¿Se puede afirmar a partir de la cota de error anterior que $\lim_{n \rightarrow \infty} P_n(x) = f(x)$ para todo $x \in [-2, 0]$?

Solución. Sean $\{x_0, x_1, \dots, x_n\} \subset [-2, 0]$ los nodos de interpolación. El error absoluto $\Delta_n(x)$ cometido en la aproximación $f(x) \approx P_n(x)$ para cualquier $x \in [-2, 0]$ viene dado por

$$\Delta_n(x) = |f(x) - P_n(x)| = \frac{|f^{(n+1)}(\xi)|}{(n+1)!} |\pi_n(x)| ,$$

siendo $\xi \in (-2, 0)$ y el polinomio $\pi_n(x) = \prod_{i=0}^n (x - x_i)$. Puesto que $x, x_i \in [-2, 0]$ se tiene $|x - x_i| \leq 2$ y por lo tanto la acotación $|\pi_n(x)| = \prod_{i=0}^n |x - x_i| \leq 2^{n+1}$. De esta forma

$$\Delta_n(x) \leq \frac{|f^{(n+1)}(\xi)|}{(n+1)!} 2^{n+1} .$$

Por otra parte, como n es un número impar, $f^{(n+1)}(\xi) = \sinh \xi$. Puesto que la función $\sinh \xi$ es creciente en \mathbb{R} e impar, es claro que $|f^{(n+1)}(\xi)| \leq |f^{(n+1)}(-2)| = 3,6269$. De esta forma se halla finalmente la cota del error

$$\Delta_n(x) \leq \frac{3,6269}{(n+1)!} 2^{n+1} .$$

Como por definición de error absoluto $\Delta_n(x) \geq 0$, de la anterior cota se tiene que

$$\lim_{n \rightarrow \infty} \Delta_n(x) = 0 ,$$

de modo que

$$\lim_{n \rightarrow \infty} P_n(x) = f(x) \text{ para todo } x \in [-2, 0] .$$

Problema 4.14 Considerar la función $f(x) = 1/(1+x)$ definida en el intervalo $[a, b]$ y sea $P_n(x)$ el polinomio de grado menor o igual que n que interpola a $f(x)$ en $n+1$ nodos pertenecientes al intervalo $[a, b]$.

(i) Si $[a, b] = [1, 3]$, hallar una cota del error absoluto que se comete cuando se aproxima $f(x) \approx P_n(x)$ para cualquier $x \in [1, 3]$.

(ii) Sea $[a, b] = [-2, 3]$. ¿Es posible ahora hallar una cota similar a la del apartado anterior?

Solución. (i) Considerar la función $f(x) = 1/(1+x)$ definida en el intervalo $[1, 3]$. Calculemos en primer lugar la derivada n -ésima de la función f por inducción. Observamos que las primeras derivadas de la función f son

$$\begin{aligned} f'(x) &= -\frac{1}{(1+x)^2} , \quad f''(x) = \frac{2}{(1+x)^3} , \\ f'''(x) &= -\frac{3!}{(1+x)^4} , \quad f^{(iv)}(x) = \frac{4!}{(1+x)^5} . \end{aligned}$$

Estos cálculos nos inducen a suponer cierto que

$$f^{(n)}(x) = (-1)^n \frac{n!}{(1+x)^{n+1}} . \quad (4.8)$$

Calculemos ahora $f^{(n+1)}(x)$ suponiendo cierta la ecuación (4.8). Puesto que

$$f^{(n+1)}(x) = \frac{d}{dx} f^{(n)}(x) = \frac{d}{dx} \left((-1)^n \frac{n!}{(1+x)^{n+1}} \right) = (-1)^{n+1} \frac{(n+1)!}{(1+x)^{n+2}}$$

continua obedeciendo la ley dada en (4.8), el método de inducción concluye que (4.8) es correcta.

A partir de la expresión explícita de $f^{(n+1)}(x)$ hallada anteriormenete se tiene que $f \in C^{n+1}[1, 3]$. Entonces, el error absoluto Δ cometido en la aproximación $f(x) \approx P_n(x)$ para cualquier $x \in [1, 3]$ viene dado por

$$\Delta = \left| \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{i=0}^n (x - x_i) \right| = \left| \frac{1}{(1+\xi_x)^{n+2}} \right| \prod_{i=0}^n |x - x_i|.$$

siendo x_i los nodos y $\xi_x \in (1, 3)$. Puesto que $x, x_i \in [1, 3]$ para $i = 0, 1, \dots, n$ se tiene que $|x - x_i| \leq 2$ y por lo tanto $\prod_{i=0}^n |x - x_i| \leq 2^{n+1}$. Entonces

$$\Delta \leq \left| \frac{1}{(1+\xi_x)^{n+2}} \right| 2^{n+1}.$$

Finalmente, como $\xi_x \in (1, 3)$, es claro que

$$\Delta \leq \left| \frac{1}{(1+1)^{n+2}} \right| 2^{n+1} = \frac{1}{2}.$$

(ii) Sea $[a, b] = [-2, 3]$. Ahora no es posible hallar una cota similar a la del apartado anterior. El motivo es que la función $f(x) = 1/(1+x)$ no es una función continua en $x = -1$ de modo que, en particular, $f(x)$ no es una función continua en el intervalo $[-2, 3]$ y, en consecuencia $f \notin C^{n+1}[-2, 3]$. Entonces, no podemos escribir una expresión del tipo

$$\Delta = \left| \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{i=0}^n (x - x_i) \right|$$

para el error absoluto Δ cometido en la aproximación $f(x) \approx P_n(x)$ para cualquier $x \in [-2, 3]$.

Problema 4.15 (i) Sea $P_2(x)$ el polinomio de grado menor o igual que 2 que interpola a la función

$$f(x) = 1 + \frac{1}{6}x(6 + 3x + 5x^2) + \sinh(1 - x)$$

en 3 nodos pertenecientes al intervalo $[0, 3]$. Hallar una cota del error absoluto cometido en la aproximación $f(x) \approx P_2(x)$ para cualquier $x \in [0, 3]$.

(ii) Hallar, mediante el método de las diferencias divididas de Newton, el polinomio $Q_2(x)$ que interpola a la función $f(x)$ en los nodos $x_i \in [0, 3]$ ($i = 0, 1, 2$) de modo que se minimice $\max_{0 \leq x \leq 3} \{|f(x) - Q_2(x)|\}$.

Solución. (i) Notemos que la función $f(x)$ es la suma de un polinomio y de la función $\sinh(1-x)$, por lo tanto $f \in C^\infty(\mathbb{R})$ y, en particular, $f \in C^3[1, 3]$. Entonces, para cualquier $x \in [0, 3]$, el error absoluto Δ que se comete en la aproximación $f(x) \approx P_2(x)$ viene dado por

$$\Delta = \frac{|f'''(\xi_x)|}{3!} \prod_{i=0}^2 |x - x_i| ,$$

donde $\xi_x \in (0, 3)$ y x_i son los nodos. Existe una acotación inmediata del término con el productorio: puesto que $x, x_i \in [0, 3]$ para $i = 0, 1, 2$, entonces $|x - x_i| \leq 3$ y por lo tanto $\prod_{i=0}^2 |x - x_i| \leq 3^3$. En definitiva se obtiene

$$\Delta \leq \frac{|f'''(\xi_x)|}{3!} 3^3 = \frac{9}{2} |f'''(\xi_x)| .$$

Puesto que

$$f(x) = 1 + \frac{1}{6}x(6 + 3x + 5x^2) + \sinh(1-x) ,$$

se obtiene derivando tres veces que $f'''(\xi_x) = 5 - \cosh(1 - \xi_x)$. Calculemos a continuación el máximo absoluto de la función continua $|f'''(x)|$ en el intervalo compacto $[0, 3]$. Para ello calculamos primero los extremos relativos de f''' en el intervalo $[0, 3]$, es decir, resolvemos la ecuación $f^{(iv)}(x) = \sinh(1-x) = 0$ para $x \in [0, 3]$. La solución de dicha ecuación es $x = 1$. Entonces, el máximo absoluto M de $|f'''(x)|$ en el intervalo $[0, 3]$ viene dado por

$$M = \max\{|f'''(0)|, |f'''(1)|, |f'''(3)|\} = \max\{3,45692, 4, 1,2378\} = 4 .$$

Concluimos que

$$\Delta \leq \frac{9}{2} \times 4 = 18 .$$

(ii) Los nodos $x_i \in [0, 3]$ ($i = 0, 1, 2$) de modo que se minimice $\|f(x) - Q_2(x)\|_\infty = \max_{0 \leq x \leq 3} \{|f(x) - P_2(x)|\}$ son, por definición, los nodos de Chebyshev en el intervalo $[a, b] = [0, 3]$, es decir,

$$x_i = \frac{a+b}{2} + \frac{b-a}{2} \cos\left(\frac{2i+1}{2n+2}\pi\right) , \quad i = 0, 1, \dots, n .$$

En nuestro caso $n = 2$, de modo que

$$\begin{aligned} x_0 &= \frac{3}{2} \left[1 + \cos\left(\frac{1}{6}\pi\right) \right] = 2,79904 , \\ x_1 &= \frac{3}{2} \left[1 + \cos\left(\frac{1}{2}\pi\right) \right] = 1,5 , \\ x_2 &= \frac{3}{2} \left[1 + \cos\left(\frac{5}{6}\pi\right) \right] = 0,200962 . \end{aligned}$$

Puesto que $f(x_0) = 23,0517$, $f(x_1) = 5,9164$ y $f(x_2) = 2,11474$, se obtiene el siguiente esquema triangular de diferencias divididas de Newton

2.79904	23.0517		
		13.1907	
1.5	5.9164		3.95068
		2.92652	
0.200962	2.11474		

El polinomio $Q_2(x)$ que interpola a la función $f(x)$ en los nodos de Chebyshev $\{x_0, x_1, x_2\}$ viene dado por

$$Q_2(x) = 23,0517 + 13,1907(x - 2,79904) + 3,95068(x - 2,79904)(x - 1,5) .$$

Problema 4.16 Sea $P_n(x)$ el polinomio de grado menor o igual que n que interpola a la función $f(x)$ en los nodos $\{x_0, x_1, \dots, x_n\} \subset [a, b]$ distintos dos a dos.

- (i) Si $f(x) = x^n$, calcular $P_n(x)$.
- (ii) Si $f(x) = x^{n+1}$, calcular $P_n(x)$ utilizando la fórmula del error en la interpolación.
- (iii) Si $f(x) = x^{n+2}$ y $[a, b] = [-2, -1, 5]$, acotar en función de n el error absoluto cometido en la aproximación $f(x) \approx P_n(x)$ para todo $x \in [a, b]$. ¿Qué se puede decir de $\lim_{n \rightarrow \infty} P_n(x)$?

Solución. (i) Si $f(x) = x^n$, en particular f es un polinomio de grado menor o igual que n , es decir $f \in \mathbb{R}_n[x]$. Entonces, por el Teorema de existencia y unicidad del polinomio interpolador, existe un único polinomio $P_n(x)$ de grado menor o igual que n que interpola a $f(x)$ en los nodos $\{x_0, x_1, \dots, x_n\}$ distintos dos a dos. En consecuencia se debe tener $P_n(x) = f(x) = x^n$.

(ii) Si $f \in C^{n+1}[a, b]$, entonces, la fórmula del error en la interpolación es

$$f(x) - P_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \pi_n(x) ,$$

siendo $\xi_x \in (a, b)$ y $\pi_n(x) = \prod_{i=0}^n (x - x_i)$. Como $f(x) = x^{n+1}$, se tiene $f^{(n+1)}(x) = (n+1)!$, de modo que

$$P_n(x) = f(x) - \pi_n(x) = x^{n+1} - \prod_{i=0}^n (x - x_i) .$$

(iii) Ahora $f(x) = x^{n+2}$, de modo que $f^{(n+1)}(x) = (n+2)!x$. Entonces, utilizando de nuevo la fórmula del error en la interpolación, para todo $x \in [a, b]$, el error absoluto $\Delta_n(x)$ cometido en la aproximación $f(x) \approx P_n(x)$ es

$$\Delta_n(x) = |f(x) - P_n(x)| = \left| \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \pi_n(x) \right| = (n+2) |\xi_x| |\pi_n(x)| ,$$

con $\xi_x \in (a, b)$. Como $[a, b] = [-2, -1, 5]$ y $x, x_i \in [a, b]$ es claro que

$$|\pi_n(x)| = \prod_{i=0}^n |x - x_i| < \left(\frac{1}{2}\right)^{n+1}.$$

Por otra parte, como $\xi_x \in (a, b) = (-2, -1, 5)$ se tiene $|\xi_x| < |-2| = 2$. Utilizando las dos acotaciones halladas se concluye que

$$\Delta_n(x) < 2(n+2) \left(\frac{1}{2}\right)^{n+1} = \frac{n+2}{2^n}.$$

Finalmente, puesto que $0 \leq \Delta_n(x) < \frac{n+2}{2^n}$ y $\lim_{n \rightarrow \infty} \frac{n+2}{2^n} = 0$, se tiene $\lim_{n \rightarrow \infty} \Delta_n(x) = 0$. Entonces

$$\lim_{n \rightarrow \infty} |f(x) - P_n(x)| = \lim_{n \rightarrow \infty} \Delta_n(x) = 0,$$

de modo que

$$\lim_{n \rightarrow \infty} P_n(x) = f(x).$$

Problema 4.17 Considerar la función $f(x) = (1 - \sqrt{x})/x^2$.

- (i) Calcular el polinomio $P_2(x)$ de grado menor o igual que 2 que interpola a $f(x)$ en los nodos $\{2, 3, 4\}$.
- (ii) Hallar una cota del error absoluto cometido al aproximar $f(x) \approx P_2(x)$ para cualquier $x \in [2, 4]$.

Solución. (i) Sea $P_2(x)$ el polinomio de grado menor o igual que 2 que interpola a $f(x) = (1 - \sqrt{x})/x^2$ en los nodos $\{2, 3, 4\}$. Entonces, la gráfica de $P_2(x)$ pasa por los puntos $(2, f(2)) = (2, -0,103553)$, $(3, f(3)) = (3, -0,081339)$ y $(4, f(4)) = (4, -0,0625)$. Es fácil obtener (mediante un esquema de diferencias divididas o bien con el método de Lagrange) la expresión

$$P_2(x) = -0,1581 - 0,0306x - 0,0016x^2.$$

(ii) Sea $\Delta = |f(x) - P_2(x)|$. Como $f \in C^\infty$ para todo $x > 0$, en particular, $f \in C^3([2, 4])$. Entonces, podemos utilizar la expresión del error en la interpolación

$$\Delta = \frac{|f'''(\xi_x)|}{3!} |\pi_2(x)|,$$

siendo $\pi_2(x) = (x-2)(x-3)(x-4)$ y $\xi \in (2, 4)$. Es evidente que, para todo $x \in [2, 4]$ se tiene que $|\pi_2(x)| < 2^3$. Entonces, se tiene una primera acotación

$$\Delta < \frac{|f'''(\xi_x)|}{3!} 2^3 \leq \frac{M}{3!} 2^3,$$

siendo $M = \max_{2 \leq \xi_x \leq 4} \{|f'''(\xi_x)|\}$. Derivando tres veces la función f se tiene

$$f'''(\xi_x) = \frac{3}{8} \frac{35\sqrt{\xi_x} - 64}{\xi_x^5}.$$

Calculemos a continuación los posibles extremos relativos de $f'''(\xi_x)$ en el intervalo $[2, 4]$. En concreto, la solución de la ecuación

$$f^{(iv)}(\xi_x) = -\frac{15}{16} \frac{63\sqrt{\xi_x} - 128}{\xi_x^6} = 0$$

viene dada por $\xi_x^* = 16384/3969 = 4,1279 \notin [2, 4]$. Entonces

$$M = \max\{|f'''(2)|, |f'''(4)|\} = \max\{0,169951, 0,00219727\} = 0,169951 .$$

En definitiva,

$$\Delta < \frac{0,169951}{3!} 2^3 = 0,226601 .$$

Problema 4.18 *Se desea construir una tabulación de la función $f(x) = \int_0^x \exp(-t^2)dt$ en abscisas equiespaciadas de modo que posteriormente se utilizará para aproximar los valores de $f(x)$ para cualquier $x \in \mathbb{R}$ mediante interpolación lineal entre dos abscisas consecutivas. ¿Cuál debe ser la longitud de paso en dichas abscisas si se desea que el error absoluto cometido en cualquiera de las anteriores aproximaciones sea menor que 10^{-3} ?*

Solución. Sea $h > 0$ la longitud de paso en las abscisas consecutivas que definimos como $\{x_0, x_1, \dots, x_n\}$. Se tiene pues que $x_i = x_0 + ih$ para $i = 0, 1, \dots, n$. Tomemos un punto arbitrario x entre dos abscisas consecutivas, es decir, $x \in [x_i, x_{i+1}]$. Sea $P_1^{[i]}(x)$ el polinomio de grado menor o igual que 1 que interpola a $f(x)$ en los nodos $\{x_i, x_{i+1}\}$. Puesto que la función $\exp(-x^2) \in C^\infty(\mathbb{R})$, se tiene que $f(x) \in C^\infty(\mathbb{R})$ y, en particular, $f(x) \in C^2([x_i, x_{i+1}])$. Entonces, si denotamos por $\Delta(x)$ el error absoluto cometido en la aproximación $f(x) \approx P_1^{[i]}(x)$ para cualquier $x \in [x_i, x_{i+1}]$, se tiene que existe un número $\xi_x \in (x_i, x_{i+1})$ tal que

$$\Delta(x) = \frac{|f''(\xi_x)|}{2!} |w_1(x)|, \text{ siendo } w_1(x) = (x - x_i)(x - x_{i+1}).$$

Es claro que $|w_1(x)| \leq (x_{i+1} - x_i)^2 = h^2$ para todo $x \in [x_i, x_{i+1}]$. Entonces, se tiene una primera acotación

$$\Delta(x) \leq \frac{h^2}{2} |f''(\xi_x)|.$$

Utilizando el Teorema Fundamental del Cálculo se tiene que $f'(x) = \exp(-x^2)$. Entonces, derivando de nuevo, se llega a que

$$f''(x) = -2x \exp(-x^2).$$

Hemos de hallar (si es que existe) una cota superior de $|f''(x)|$ para todo $x \in \mathbb{R}$ puesto que, aunque $x \in [x_i, x_{i+1}]$, se tiene que el intervalo $[x_i, x_{i+1}] \subset \mathbb{R}$ es arbitrario. Recordemos que \mathbb{R} no es un intervalo compacto (cerrado y acotado), de modo que en un principio no tiene ni siquiera por qué existir la cota superior que buscamos. Sin embargo, para la función que tenemos sí que existe como

veremos a continuación. En primer lugar, observemos que una condición suficiente para que dicha cota exista es que $f''(x) \in C(\mathbb{R})$ y además $f''(x)$ tenga asíntota horizontal tanto cuando $x \rightarrow \infty$ como cuando $x \rightarrow -\infty$. Como además $f''(x) \in C^1(\mathbb{R})$, denotando por $\lim_{x \rightarrow \pm\infty} f''(x) = \alpha_{\pm}$ con $\alpha_{\pm} \in \mathbb{R}$ finitos y siendo $x_i^* \in \mathbb{R}$ con $i = 1, \dots, k$ las abscisas de los extremos relativos de $f''(x)$, se tiene que existe una cota $M = \max\{|f''(x)| : \forall x \in \mathbb{R}\}$ y su valor viene dado por

$$M = \max\{|f''(\alpha_+)|, |f''(\alpha_-)|, |f''(x_1^*)|, \dots, |f''(x_k^*)|\}.$$

En nuestro caso, como además $f''(x)$ es una función impar ($f(-x) = -f(x)$), sólo es necesario estudiar el siguiente límite:

$$\lim_{x \rightarrow \infty} f''(x) = \lim_{x \rightarrow \infty} \frac{-2x}{\exp(x^2)} = \lim_{x \rightarrow \infty} \frac{-1}{x \exp(x^2)} = 0 = \alpha_+,$$

donde en la penúltima igualdad se ha utilizado la regla de l'Hopital. Por simetría, $\alpha_+ = \alpha_- = 0$. Los extremos relativos de $f''(x)$ serán solución de la ecuación $f'''(x) = -2\exp(-x^2)(1-2x^2) = 0$, es decir, $x = \pm\sqrt{1/2}$. Puesto que la derivada $f^{(iv)}(\sqrt{1/2}) = 4\sqrt{2/e} > 0$, se tiene que $x = \sqrt{1/2}$ es un mínimo de $f''(x)$ y, por simetría, $x = -\sqrt{1/2}$ un máximo. Se tiene pues que la gráfica $f''(x)$ es la dada en la Figura 4.1.

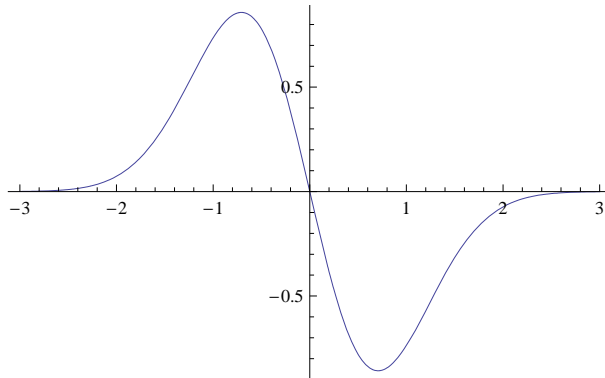


Figura 4.1: Gráfica de la función $f''(x) = -2x \exp(-x^2)$.

En definitiva, es claro que $M = |f''(\sqrt{1/2})| = \sqrt{2/e}$. De este modo se tiene que

$$\Delta(x) \leq \frac{h^2}{2} \sqrt{\frac{2}{e}} = \frac{h^2}{\sqrt{2e}}.$$

Finalmente, para hallar la longitud de paso h de modo que $\Delta(x) < 10^{-3}$, resolvemos la inecuación

$$\frac{h^2}{\sqrt{2e}} < 10^{-3},$$

y obtenemos que $h < \sqrt{\sqrt{2e} \cdot 10^{-3}} = 0,0482871$.

Problema 4.19 Consideremos la función $f(x) = \sqrt{(4+x)^5}(328 - 19x + x^2)$.

- (i) Hallar, mediante el método de las diferencias divididas de Newton, el polinomio $P_2(x)$ de grado menor o igual que 2 que interpola a la función $f(x)$ en los tres nodos equiespaciados de la partición del intervalo $[0, 2]$.
- (ii) Acotar el error absoluto $\Delta(x)$ cometido en la aproximación $f(x) \approx P_2(x)$ para todo $x \in [0, 2]$.
- (iii) Hallar el polinomio $Q_2(x)$ de grado menor o igual que 2 tal que la cota del error absoluto $\bar{\Delta}(x)$ cometido en la aproximación $f(x) \approx Q_2(x)$ para todo $x \in [0, 2]$ sea la mínima posible.

Solución. (i) Los nodos $x_i \in [0, 2]$ ($i = 0, 1, 2$) equiespaciados que definen una partición del intervalo $[0, 2]$ son $x_i = i$. Puesto que $f(0) = 10496$, $f(1) = 17329,5$ y $f(2) = 25925,4$, se obtiene el siguiente esquema triangular de diferencias divididas de Newton

0	10496		
		6833.5	
1	17329.5		881.2
		8595.9	
2	25925.4		

El polinomio $P_2(x)$ que interpola a la función $f(x)$ en los nodos $\{0, 1, 2\}$ viene dado por

$$P_2(x) = 10496 + 6833,5x + 881,2x(x-1) .$$

(ii) Notemos que la función $f(x) = \sqrt{(4+x)^5}(328 - 19x + x^2)$ es el producto de un polinomio (por lo tanto de clase $C^\infty(\mathbb{R})$) por la función $(4+x)^{5/2}$ que es de clase $C^\infty(I)$ siendo el intervalo $I = \{x \in \mathbb{R} : x > -4\}$. Como el intervalo $[0, 2] \subset I$, se concluye que $f \in C^\infty([0, 2])$ y, en particular, $f \in C^3([0, 2])$. Entonces, para cualquier $x \in [0, 2]$, el error absoluto $\Delta(x)$ que se comete en la aproximación $f(x) \approx P_2(x)$ viene dado por

$$\Delta(x) = \frac{|f'''(\xi_x)|}{3!} \prod_{i=0}^2 |x - i| ,$$

donde $\xi_x \in (0, 2)$. Puesto que $x \in [0, 2]$, entonces para $i = 0, 1, 2$, se tiene $|x - i| \leq 2$ y por lo tanto $\prod_{i=0}^2 |x - i| \leq 2^3$. En definitiva se obtiene

$$\Delta(x) \leq \frac{|f'''(\xi_x)|}{3!} 2^3 = \frac{4}{3} |f'''(\xi_x)| ,$$

con $\xi_x \in (0, 2)$. Derivando tres veces la función $f(x)$ se llega a

$$f'''(x) = \frac{315}{8} \frac{x(x-1)}{\sqrt{4+x}} .$$

Calculemos a continuación el máximo absoluto de la función continua $|f'''(x)|$ en el intervalo compacto $[0, 2]$. Para ello calculamos primero los extremos relativos de f''' en el intervalo $[0, 2]$, es decir, resolvemos la ecuación

$$f^{(iv)}(x) = \frac{315}{16} \frac{-8 + 15x + 3x^2}{\sqrt{(4+x)^3}} = 0$$

para $x \in [0, 2]$. Las soluciones de dicha ecuación son las soluciones de la ecuación $-8 + 15x + 3x^2 = 0$ que vienen dadas por

$$x = \frac{1}{6}(-15 \pm \sqrt{321}) ,$$

es decir, $x = 0,486079 \in [0, 2]$ puesto que la otra solución $-5,48608 \notin [0, 2]$. Entonces, el máximo absoluto M de $|f'''(x)|$ en el intervalo $[0, 2]$ viene dado por

$$\begin{aligned} M &= \max\{|f'''(0)|, |f'''(2)|, |f'''(0,486079)|\} = \max\{0, 32,1496, 4,64398\} \\ &= 32,1496 . \end{aligned}$$

Concluimos que

$$\Delta(x) \leq \frac{4}{3} \times 32,1496 = 42,8661 .$$

(iii) El polinomio $Q_2(x)$ deberá interpolar a la función $f(x)$ en los 3 nodos x_i ($i = 0, 1, 2$) de Chebychev en el intervalo $[a, b] = [0, 2]$. Entonces,

$$x_i = \frac{a+b}{2} + \frac{b-a}{2} \cos\left(\frac{2i+1}{2 \times 2+2}\pi\right) , \quad i = 0, 1, 2 .$$

En nuestro caso

$$x_0 = 1 + \cos\left(\frac{1}{6}\pi\right) = 1,86603 ,$$

$$x_1 = 1 + \cos\left(\frac{1}{2}\pi\right) = 1 ,$$

$$x_2 = 1 + \cos\left(\frac{5}{6}\pi\right) = 0,133975 ,$$

Puesto que $f(x_0) = 24671,3$, $f(x_1) = 17329,5$ y $f(x_2) = 11309,3$, se obtiene el siguiente esquema triangular de diferencias divididas de Newton

1.86603	24671.3		
		8477.54	
1	17329.5		881.039
		6951.53	
0.133975	11309.3		

El polinomio $Q_2(x)$ que interpola a la función $f(x)$ en los nodos de Chebychev $\{x_0, x_1, x_2\}$ viene dado por

$$Q_2(x) = 24671,3 + 8477,54(x - 1,86603) + 881,039(x - 1,86603)(x - 1) .$$

Capítulo 5

Derivación Numérica

5.1. Resumen de teoría

Las reglas de diferenciación para funciones usuales son bien conocidas. Sin embargo, no siempre es posible utilizarlas; por ejemplo si nuestro único conocimiento sobre la función es el valor que toma en unas determinadas abscisas, es decir sólo se dispone de una tabulación de la función; o bien no nos interesa utilizarlas, por ejemplo en funciones con una expresión analítica demasiado complicada. En este sentido son interesantes las técnicas de derivación numérica, ya que nos van a permitir calcular una aproximación de la derivada k -ésima de la función en un punto. Las fórmulas de derivación numérica tienen su aplicación más importante en la resolución numérica de ecuaciones diferenciales.

- Suponiendo la función f suficientemente diferenciable, mediante desarrollos de Taylor se pueden obtener fórmulas de derivación numérica. Por ejemplo:

$$f'(x) = \frac{f(x+h) - f(x)}{h} + O(h) , \quad (5.1)$$

que se conoce con el nombre de *fórmula de diferencias hacia adelante*, o bien, de forma totalmente análoga podemos obtener la llamada *fórmula de diferencias hacia atrás*

$$f'(x) = \frac{f(x) - f(x-h)}{h} + O(h) , \quad (5.2)$$

o la llamada *fórmula de diferencias centradas*

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + O(h^2) . \quad (5.3)$$

- Sea $F(h)$ una fórmula que nos proporciona una aproximación discreta (con paso de discretización h) de cierto cálculo exacto V . Supongamos además que del análisis del error de truncamiento hemos obtenido la relación

$$V = F(h) + a_1 h^{p_1} + a_2 h^{p_2} + a_3 h^{p_3} + \dots \quad \text{con } 0 < p_1 < p_2 < \dots . \quad (5.4)$$

Mediante la *Extrapolación de Richardson* se consigue

$$a_1 h^{p_1} \approx \frac{F(h) - F(2h)}{2^{p_1} - 1} . \quad (5.5)$$

- La derivación interpolatoria consiste en aproximar la derivada de f por la de un cierto polinomio interpolador P_n de f , es decir,

$$f'(a) \approx P'_n(a) .$$

5.2. Problemas resueltos

Problema 5.1 *Supongamos que conocemos el valor de una función f en dos puntos, es decir las parejas (x_0, f_0) y (x_1, f_1) .*

- (i) *Obtener (utilizando el formalismo de Lagrange) la fórmula de derivación interpolatoria que nos permite aproximar la derivada $f'(x_0)$.*
- (ii) *Demostrar que la fórmula obtenida en el apartado anterior es en realidad la fórmula de diferencias hacia adelante.*

Solución. (i) Utilizando derivación interpolatoria y el formalismo de Lagrange para el polinomio interpolador $P(x)$ podemos escribir

$$f'(x_0) \approx P'(x_0) = f_0 \ell'_0(x_0) + f_1 \ell'_1(x_0) .$$

Como los polinomios de Lagrange son

$$\ell_0(x) = \frac{x - x_1}{x_0 - x_1} , \quad \ell_1(x) = \frac{x - x_0}{x_1 - x_0} ,$$

concluimos que

$$f'(x_0) \approx \frac{f_0}{x_0 - x_1} + \frac{f_1}{x_1 - x_0} .$$

(ii) Definimos h de la forma $x_1 = x_0 + h$. Entonces, podemos reescribir la ecuación anterior como

$$f'(x_0) \approx \frac{f(x_0)}{-h} + \frac{f(x_0 + h)}{h} = \frac{f(x_0 + h) - f(x_0)}{h} ,$$

que no es más que la fórmula de diferencias hacia adelante.

Problema 5.2 *Considerar la tabulación de una función $f(x)$ siguiente*

x_i	5	-7	-6	0
$f(x_i)$	1	-23	-54	-954

- (i) *Calcular el polinomio de tercer grado que interpola a todos los puntos de la tabla mediante el método de las diferencias divididas de Newton.*
- (ii) *Aproximar el valor de la derivada $f'(-6,5)$ mediante alguna de las siguientes fórmulas: diferencias hacia adelante, diferencias hacia atrás o diferencias centradas. Explicar vuestra elección.*

Solución. (i) Realizando el esquema de diferencias divididas siguiente

5	1			
		2		
-7	-23		3	
		-31		4
-6	-54		-17	
		-150		
0	-954			

obtenemos el polinomio interpolador $P(x) = 1 + 2(x - 5) + 3(x - 5)(x + 7) + 4(x - 5)(x + 7)(x + 6)$. Finalmente, desarrollando se tiene

$$P(x) = 4x^3 + 35x^2 - 84x - 954 .$$

(ii) Según la tabulación de que disponemos, la única fórmula posible es la de diferencias centradas

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + O(h^2) .$$

En concreto, tomando la longitud de paso $h = 0,5$ se tiene

$$f'(-6,5) \approx f(-6) - f(-7) = -54 - (-23) = -31 .$$

Problema 5.3 Se está analizando un circuito eléctrico que tiene una bobina, una resistencia y un generador de corriente. Aplicando las leyes de Kirchhoff se sabe que dicho circuito está gobernado por la siguiente ecuación

$$\epsilon(t) = L \frac{dI}{dt} + RI ,$$

siendo $\epsilon(t)$ la fuerza electromotriz en función del tiempo del generador, L el coeficiente de autoinducción de la bobina, R la resistencia e I la intensidad de corriente eléctrica. Se realiza con un cronómetro y un amperímetro las siguientes mediciones

t_i (segundos)	0	1	2
$I(t_i)$ (amperios)	2	4	10

Se sabe que $L = 0,5$ henries y $R = 5$ ohms son constantes.

- (i) Utilizar el formalismo de las diferencias divididas de Newton para realizar derivación interpolatoria con todos los puntos de la tabulación con el objetivo de aproximar el valor $\epsilon(0)$.
- (ii) Suponiendo que los valores $I(t_i)$ de la tabulación dada son exactos pero que $R = 5 \pm 0,3$ ohms, aproximar el error absoluto cometido en la aproximación efectuada en el apartado (i) en función de la derivada de algún orden de $I(t)$.

Solución. (i) Utilizando el formalismo de diferencias divididas de Newton, el polinomio $P_2(t)$ de grado menor o igual que 2 que interpola a todos los datos de la tabla es

$$P(t) = I[t_0] + I[t_0, t_1](t - t_0) + I[t_0, t_1, t_2](t - t_0)(t - t_1) .$$

Realizando el esquema de diferencias divididas se obtiene

0	2		
		2	
1	4		2
		6	
2	10		

de manera que $I[t_0] = I[t_0, t_1] = I[t_0, t_1, t_2] = 2$ y por lo tanto

$$P_2(t) = 2[1 + (t - 0) + (t - 0)(t - 1)] = 2(1 + t^2) .$$

Este polinomio será utilizado para realizar la aproximación

$$\dot{I}(0) := \frac{dI(0)}{dt} \approx \frac{dP_2(0)}{dt} = 0 . \quad (5.6)$$

Podemos aproximar el valor $\epsilon(0)$ mediante la anterior aproximación y la ecuación que gobierna el circuito, es decir

$$\epsilon(0) = L \frac{dI(0)}{dt} + RI(0) \approx L \frac{dP_2(0)}{dt} + RI(0) = 0,5 \times 0 + 5 \times 2 = 10 \text{ voltios} .$$

(ii) Se ha de obtener una expresión aproximada para $\Delta_{\epsilon(0)}$, es decir el error absoluto cometido en el apartado anterior. Según el enunciado del problema, las únicas magnitudes que contienen errores son R y las aproximadas en el apartado (i), es decir $\dot{I}(0)$ y $\epsilon(0)$. Tomando la ecuación del circuito se tiene que $\epsilon(\dot{I}, R) = L\dot{I} + RI$, de manera que según la fórmula de propagación de errores absolutos

$$\Delta_{\epsilon} \approx \left| \frac{\partial \epsilon}{\partial \dot{I}} \right| \Delta_{\dot{I}} + \left| \frac{\partial \epsilon}{\partial R} \right| \Delta_R = L\Delta_{\dot{I}} + |I|\Delta_R .$$

Particularizando esta expresión para tiempo $t = 0$ se obtiene

$$\Delta_{\epsilon(0)} \approx L\Delta_{\dot{I}(0)} + |I(0)|\Delta_R = 0,5\Delta_{\dot{I}(0)} + 2 \times 0,3 . \quad (5.7)$$

Para finalizar el problema es necesario obtener una expresión de $\Delta_{\dot{I}(0)}$ en función de la derivada de algún orden de $I(t)$. Como $\dot{I}(0)$ ha sido aproximado mediante (5.6), es de allí de donde se ha de obtener tal expresión. En concreto, se sabe por la teoría de interpolación que el error absoluto cometido en la aproximación $I(t) \approx P_2(t)$ para cualquier valor $t \in [0, 2]$ viene dado por

$$\Delta_{I(t)} := |I(t) - P_2(t)| = \left| \frac{I^{(iii)}(\xi(t))}{3!} (t - 0)(t - 1)(t - 2) \right| ,$$

donde $\xi(t) \in (0, 2)$. Derivando esta ecuación respecto del tiempo t y particularizándola para $t = 0$ se obtiene

$$\Delta_{\dot{I}(0)} := |\dot{I}(0) - \dot{P}_2(0)| = \left| \frac{I^{(iii)}(\xi(0))}{3!}(-1)(-2) \right| = \frac{1}{3} \left| I^{(iii)}(\xi(0)) \right| ,$$

donde $\xi(0) \in (0, 2)$. Finalmente, introduciendo esta expresión en la aproximación (5.7) se obtiene

$$\Delta_{\epsilon(0)} \approx \frac{1}{6} \left| I^{(iii)}(\xi(0)) \right| + 0,6 .$$

Problema 5.4 *Cosideremos la fórmula de diferencias centradas*

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + O(h^2) .$$

- (i) *Aplicar la extrapolación de Richardson a la fórmula de diferencias centradas para obtener una nueva fórmula de aproximación para $f'(x)$, obteniendo también el orden del error.*
- (ii) *Sea $f \in C^4[x-2h, x+2h]$, con $h > 0$. Utilizando desarrollos de Taylor hasta cuarto orden para $f(x \pm h)$ y $f(x \pm 2h)$, obtener la misma fórmula que en el apartado anterior.*

Solución. (i) La fórmula de diferencias centradas tiene un desarrollo asintótico del error que sólo contiene potencias pares de la longitud de paso h , es decir,

$$f'(x) = F(h) + a_1 h^2 + O(h^4) ,$$

siendo a_1 una constante independiente de h y

$$F(h) = \frac{f(x+h) - f(x-h)}{2h} .$$

Realizando la extrapolación de Richardson con duplicación del paso, se obtiene el primer término del error, es decir,

$$\begin{aligned} a_1 h^2 &= \frac{F(h) - F(2h)}{2^2 - 1} + O(h^4) \\ &= \frac{-f(x+2h) + 2f(x+h) - 2f(x-h) + f(x-2h)}{12h} + O(h^4) . \end{aligned}$$

Se tiene pues que

$$f'(x) = \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h} + O(h^4) .$$

(ii) Una forma de obtener expresiones en diferencias finitas para la derivada consiste (suponiendo que la función es suficientemente derivable) en combinar diferentes desarrollos de Taylor en un entorno de los puntos donde conocemos

la función. En concreto, realizamos los desarrollos de Taylor hasta cuarto orden siguientes

$$\begin{aligned} f(x+2h) &= f(x) + 2hf'(x) + 2h^2f''(x) + \frac{4h^3}{3}f'''(x) + \frac{2h^4}{3}f^{(iv)}(\xi_1), \\ f(x+h) &= f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{6}f'''(x) + \frac{h^4}{24}f^{(iv)}(\xi_2), \\ f(x-h) &= f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{6}f'''(x) + \frac{h^4}{24}f^{(iv)}(\xi_3), \\ f(x-2h) &= f(x) - 2hf'(x) + 2h^2f''(x) - \frac{4h^3}{3}f'''(x) + \frac{2h^4}{3}f^{(iv)}(\xi_4), \end{aligned}$$

siendo $\xi_i \in [x-2h, x+2h]$ para $i = 1, 2, 3, 4$. A continuación realizamos la siguiente combinación lineal $-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)$ y obtenemos

$$12hf'(x) + \frac{1}{3}(-2f^{(iv)}(\xi_1) + f^{(iv)}(\xi_2) - f^{(iv)}(\xi_3) + 2f^{(iv)}(\xi_4))h^4.$$

Se tiene pues, despejando $f'(x)$ de la igualdad entre las dos expresiones anteriores, que

$$f'(x) = \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h} + O(h^4).$$

Problema 5.5 Considerar la siguiente tabla de valores

x	0	1	2
$f(x)$	3.121	2.714	4.128

- (i) Sabiendo que $|f'''(x)| \leq 3/4$ para todo $x \in [0, 2]$, hallar una cota para el error cometido cuando aproximamos $f(1.3)$ por interpolación polinomial con la tabla adjunta.
- (ii) Aproximar $f'(1)$ por derivación interpolatoria y, teniendo en cuenta el apartado anterior, obtener una cota del error cometido.
- (iii) ¿Se obtendría la misma aproximación que la efectuada en el apartado anterior mediante alguna fórmula de derivación numérica obtenida a través de desarrollos de Taylor?

Solución. (i) Sea $P_2(x)$ el polinomio que interpola a f mediante la tabulación adjunta. Se sabe que para todo $x \in [0, 2]$ se verifica

$$|f(x) - P_2(x)| = \left| \frac{f'''(\xi_x)}{3!} \pi(x) \right|,$$

donde $\xi_x \in (0, 2)$ y $\pi(x) = x(x-1)(x-2)$. Puesto que $|f'''(\xi_x)| \leq 3/4$, particularizando para $x = 1.3$ se tiene la siguiente acotación del error Δ cometido en

la aproximación $f(1.3) \approx P_2(1.3)$

$$\begin{aligned}\Delta &:= |f(1.3) - P_2(1.3)| = \left| \frac{f'''(\xi_{1.3})}{3!} \pi(1.3) \right| \leq \frac{3/4}{3!} |\pi(1.3)| \\ &= \frac{1}{8} |1.3 \times (1.3 - 1) \times (1.3 - 2)| = 0.034125 .\end{aligned}$$

(ii) La derivada de la función f en un nodo x_α se aproxima de forma interpolatoria por

$$f'(x_\alpha) \approx P_2(x_\alpha) ,$$

donde el error cometido Δ' vale

$$\Delta' = \left| \frac{f'''(\xi_{x_\alpha})}{3!} \pi'(x_\alpha) \right| .$$

Realizando el esquema de diferencias divididas asociado a la tabulación del enunciado se tiene

0	3.121		
		-0.407	
1	2.714		0.9105
		1.414	
2	4.128		

de manera que

$$\begin{aligned}P_2(x) &= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\ &= 3.121 - 0.407x + 0.9105x(x - 1) \\ &= 3.121 - 1.3175x + 0.9105x^2 .\end{aligned}$$

Por consiguiente, tomando el nodo $x_\alpha = 1$, la aproximación buscada es $f'(1) \approx P_2'(1) = 0.5035$.

El error absoluto Δ' cometido en la anterior aproximación es

$$\Delta' = \left| \frac{f'''(\xi_1)}{3!} \pi'(1) \right| .$$

Como $\xi_1 \in (0, 2)$ se tiene la cota $|f'''(\xi_1)| \leq 3/4$. Además, como $\pi(x) = x(x - 1)(x - 2)$, obtenemos $\pi'(1) = -1$. En definitiva la cota del error es

$$\Delta' \leq \left| -\frac{3/4}{3!} \right| = \frac{1}{8} .$$

(iii) Puesto que los nodos de la tabulación son equiespaciados, seguro que existe alguna fórmula de derivación numérica procedente de desarrollos de Taylor

que nos da la misma aproximación de $f'(1)$ que la obtenida en el apartado anterior. De hecho es natural pensar en la fórmula de diferencias centradas

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h} .$$

Como la longitud de paso de los nodos es $h = 1$, particularizando la fórmula anterior para el valor $x = 1$ se tiene

$$f'(1) \approx \frac{f(2) - f(0)}{2} = \frac{4.128 - 3.121}{2} = 0.5035 .$$

Problema 5.6 *Se han realizado las siguientes mediciones de la posición x en función del tiempo t para una partícula.*

t (s)	0	1	2
x (m)	0	1.1752	3.6268

- (i) *Aproximar la velocidad v de la partícula para tiempo $t = 1$ mediante derivación interpolatoria con todos los datos de la tabla adjunta y utilizando el formalismo de Newton. ¿Se habría obtenido la misma aproximación utilizando alguna fórmula de derivación numérica mediante desarrollos de Taylor? Razonar la respuesta.*
- (ii) *Posteriormente se sabe que la expresión explícita de la posición en función del tiempo viene dada por $x(t) = \sinh t$. Hallar una cota del error absoluto cometido en la aproximación $v(1) \approx dP_2(1)/dt$, siendo $P_2(t)$ el polinomio de grado menor o igual que dos que interpola a los datos de la tabla adjunta y posteriormente comparar dicha cota con el valor exacto del error absoluto.*

Solución. (i) Con los datos de la tabla adjunta se obtiene el siguiente esquema triangular de diferencias divididas de Newton

0	0		
		1.1752	
1	1.1752		0.6382
		2.4516	
2	3.6268		

El polinomio $P_2(t)$ que interpola a la función $x(t)$ según los datos de la tabla es

$$P_2(t) = 1,1752t + 0,6382t(t-1) .$$

La aproximación mediante derivación interpolatoria es

$$v(1) \approx dP_2(1)/dt = 1,1752 + 0,6382 = 1,8134 .$$

Notemos que los nodos de interpolación de la tabla adjunta son equiespaciados. Esto implica que existe una fórmula de derivación interpolatoria procedente

de desarrollos de Taylor que nos ofrece la misma aproximación que la efectuada mediante derivación interpolatoria. Además, la aproximación efectuada ha sido la derivada de la función $x(t)$ en el nodo central de manera que, podemos utilizar la fórmula de diferencias centradas

$$v(1) = x'(1) \approx \frac{x(1+h) - x(1-h)}{2h} = \frac{x(2) - x(0)}{2} = \frac{3,6268 - 0}{2} = 1,8134 ,$$

donde se ha tomado la longitud de paso $h = 1$ para poder utilizar la tabulación dada.

(ii) Sabemos que el error absoluto Δ cometido en la aproximación $x(t) \approx P_2(t)$ para cualquier valor de $t \in [0, 2]$ viene dado por

$$\Delta = |x(t) - P_2(t)| = \left| \frac{x'''(\xi_t)}{3!} t(t-1)(t-2) \right| ,$$

donde $\xi_t \in (0, 2)$. Entonces, derivando la expresión anterior respecto de t , el error absoluto Δ' cometido en la aproximación $v(1) \approx P'_2(1)$ viene dado por

$$\Delta' = |x'(1) - P'_2(1)| = \left| \frac{x'''(\xi_1)}{3!} 1(1-2) \right| = \frac{|x'''(\xi_1)|}{3!} ,$$

donde $\xi_1 \in (0, 2)$. Puesto que la tercera derivada de la función $x(t) = \sinh t$ viene dada por $x'''(\xi_1) = \cosh \xi_1$, siendo esta una función positiva y creciente en el intervalo $\xi_1 \in (0, 2)$, es claro que

$$\Delta' \leq \frac{\cosh 2}{3!} = \frac{3,7622}{3!} = 0,627 .$$

Por supuesto en este caso es posible calcular de manera exacta el error absoluto Δ' . En concreto

$$\Delta' = |x'(1) - P'_2(1)| = |\cosh 1 - 1,8134| = |1,54308 - 1,8134| = 0,270319 .$$

Problema 5.7 Responder a las siguientes preguntas.

(i) Suponiendo que la función $f(x)$ es suficientemente derivable, hallar mediante desarrollos de Taylor el orden del error en la aproximación

$$f'(x) \approx \frac{-3f(x) + 4f(x+h) - f(x+2h)}{2h} ,$$

cuando $h > 0$ tiende a cero.

(ii) Si $f(x) = \sin x$, acotar el error absoluto cometido en el apartado (i) en función de h utilizando adecuadamente la derivación interpolatoria.

Solución. (i) Supongamos que $f \in C^3[x, x+2h]$. Entonces, podemos efectuar los siguientes desarrollos de Taylor de f entorno de x .

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{3!} f'''(\xi_1) ,$$

$$f(x+2h) = f(x) + 2hf'(x) + 2h^2f''(x) + \frac{8}{3!}h^3f'''(\xi_2) ,$$

siendo $\xi_1 \in (x, x+h)$ y $\xi_2 \in (x, x+2h)$. Restando cuatro veces la primera ecuación menos la segunda se consigue eliminar el término que contiene $f''(x)$ y se obtiene

$$4f(x+h) - f(x+2h) = 3f(x) + 2hf'(x) + O(h^3) .$$

Despejando de aquí $f'(x)$ se llega a

$$f'(x) = \frac{-3f(x) + 4f(x+h) - f(x+2h)}{2h} + O(h^2) .$$

(ii) Observemos que la fórmula que aproxima la derivada primera del apartado anterior proviene del método de Taylor, de manera que la misma fórmula puede ser obtenida por derivación interpolatoria. Puesto que la función f en esta fórmula se evalúa en tres puntos equiespaciados con longitud de paso h , tomaremos los nodos $\Sigma = \{x_0, x_1, x_2\}$, siendo $x_i = x_0 + ih$ con $i = 0, 1, 2$. Denotando por $P_2(x)$ al polinomio de grado menor o igual que 2 que interpola a $f(x)$ en Σ , se tiene que $f(x) = P_2(x) + \Delta(x)$, siendo

$$\Delta(x) = \frac{f'''(\xi_x)}{3!}w_2(x) , \quad w_2(x) = \prod_{i=0}^2 (x - x_i) , \quad \xi_x \in (x_0, x_2) .$$

Derivando respecto de x la expresión $f(x) = P_2(x) + \Delta(x)$ y evaluando en el nodo x_0 se tiene que $f'(x_0) = P_2'(x_0) + \Delta'(x_0)$. De este modo, se tendrá que la aproximación $f'(x_0) \approx P_2'(x_0)$ coincidirá con la fórmula del apartado (i) y tendrá asociada un error absoluto $|\Delta'(x_0)|$. Tomando el conjunto de nodos equiespaciados Σ , se tiene que

$$|\Delta'(x_0)| = \frac{|f'''(\xi_{x_0})|}{3!}|x_0 - x_1||x_0 - x_2| = \frac{|f'''(\xi_{x_0})|}{3}h^2 , \quad \xi_{x_0} \in (x_0, x_0+2h) .$$

Finalmente, como $f(x) = \sin x$, es claro que $|f'''(\xi_{x_0})| \leq 1$, de modo que se obtiene la cota

$$|\Delta'(x_0)| \leq \frac{h^2}{3} .$$

Problema 5.8 La capacidad calorífica C depende de la temperatura T a la que está sometido el material. La siguiente tabla muestra unas medidas experimentales de la capacidad calorífica de 1Kg de agua a diferentes temperaturas

T (K)	273	323	373
C (KJ/K)	4.217	4.180	4.215

Suponiendo que la función $C(T)$ es suficientemente derivable, hallar mediante desarrollos de Taylor una aproximación de dC/dT en $T = 273$ utilizando los valores $C(323)$ y $C(373)$.

Solución. En primer lugar, notemos que conocemos la función $C(T)$ en un conjunto de puntos T_i equiespaciados, es decir, $T_i = T_0 + ih$ con una longitud de paso $h = 50$. Es por este motivo que podemos utilizar desarrollos de Taylor para aproximar dC/dT en $T_0 = 273$. Además, se han de utilizar los valores $C(323)$ y $C(373)$ que no son más que los valores de $C(T_0 + h)$ y $C(T_0 + 2h)$. Existen varias formas de resolver el problema. Tres de ellas son:

- Supongamos que $C \in \mathcal{C}^2[T_0, T_0 + 2h]$ y realizemos los siguientes desarrollos de Taylor de la función $C(T)$ entorno de T_0 :

$$C(T_0 + h) = C(T_0) + hC'(T_0) + \frac{1}{2}h^2C''(\xi_1) ,$$

$$C(T_0 + 2h) = C(T_0) + 2hC'(T_0) + 2h^2C''(\xi_2) ,$$

con $\xi_1 \in (T_0, T_0 + h)$ y $\xi_2 \in (T_0, T_0 + 2h)$. Aquí, la prima indica derivada respecto de T . Restando la ecuación segunda menos la primera se llega a que

$$C(T_0 + 2h) - C(T_0 + h) = hC'(T_0) + O(h^2) ,$$

de donde, despejando se obtiene

$$C'(T_0) = \frac{C(T_0 + 2h) - C(T_0 + h)}{h} + O(h) .$$

De este modo,

$$C'(273) \approx \frac{C(373) - C(323)}{50} = \frac{4,215 - 4,180}{50} = 0,0007 .$$

NOTA: Si, en los anteriores desarrollos de Taylor, en lugar de restar la ecuación segunda menos la primera se suman las dos ecuaciones se llega a que

$$C(T_0 + h) + C(T_0 + 2h) = 2C(T_0) + 3hC'(T_0) + O(h^2) ,$$

de donde, despejando se obtiene

$$C'(T_0) = \frac{C(T_0 + h) + C(T_0 + 2h) - 2C(T_0)}{3h} + O(h) .$$

De este modo,

$$\begin{aligned} C'(273) &\approx \frac{C(323) + C(373) - 2C(273)}{150} \\ &= \frac{4,180 + 4,215 - 2 \times 4,217}{150} = -0,00026 . \end{aligned}$$

- El resultado anterior puede ser mejorado si suponemos que $C \in \mathcal{C}^3[T_0, T_0 + 2h]$ y realizamos los siguientes desarrollos de Taylor de la función $C(T)$ entorno de T_0 :

$$C(T_0 + h) = C(T_0) + hC'(T_0) + \frac{1}{2}h^2C''(T_0) + \frac{1}{3!}h^3C'''(\xi_1) ,$$

$$C(T_0 + 2h) = C(T_0) + 2hC'(T_0) + 2h^2C''(T_0) + \frac{4}{3}h^3C'''(\xi_2) ,$$

con $\xi_1 \in (T_0, T_0 + h)$ y $\xi_2 \in (T_0, T_0 + 2h)$. Restando 4 veces la ecuación primera menos la segunda se llega a que

$$4C(T_0 + h) - C(T_0 + 2h) = 3C(T_0) + 2hC'(T_0) + O(h^3) ,$$

de donde, despejando se obtiene

$$C'(T_0) = \frac{4C(T_0 + h) - C(T_0 + 2h) - 3C(T_0)}{2h} + O(h^2) .$$

De este modo, tomando $h = 50$ se tiene

$$\begin{aligned} C'(273) &\approx \frac{4C(323) - C(373) - 3C(273)}{100} \\ &= \frac{4 \times 4,180 - 4,215 - 3 \times 4,217}{100} = -0,00146 . \end{aligned}$$

Problema 5.9 Sea la función $f \in C^\infty[x - h, x + 3h]$.

(i) Hallar el orden del error en la aproximación de la derivada siguiente

$$f'(x) \approx \frac{-3f(x - h) - 10f(x) + 18f(x + h) - 6f(x + 2h) + f(x + 3h)}{12h} .$$

(ii) Obtener, mediante extrapolación de Richardson aplicada a la fórmula anterior, otra aproximación de $f'(x)$ con un error menor para h suficientemente pequeño.

Solución. (i) Como $f \in C^5[x - h, x + 3h]$, podemos realizar los siguientes desarrollos de Taylor:

$$\begin{aligned} f(x \pm h) &= f(x) \pm hf'(x) + \frac{h^2}{2}f''(x) \pm \frac{h^3}{3!}f'''(x) + \frac{h^4}{4!}f^{(iv)}(x) + O(h^5), \\ f(x + 2h) &= f(x) + 2hf'(x) + \frac{(2h)^2}{2}f''(x) + \frac{(2h)^3}{3!}f'''(x) + \\ &\quad \frac{(2h)^4}{4!}f^{(iv)}(x) + O(h^5), \\ f(x + 3h) &= f(x) + 3hf'(x) + \frac{(3h)^2}{2}f''(x) + \frac{(3h)^3}{3!}f'''(x) + \\ &\quad \frac{(3h)^4}{4!}f^{(iv)}(x) + O(h^5). \end{aligned}$$

De este modo,

$$-3f(x - h) - 10f(x) + 18f(x + h) - 6f(x + 2h) + f(x + 3h) = 12hf'(x) + O(h^5) ,$$

con lo que

$$f'(x) = \frac{-3f(x - h) - 10f(x) + 18f(x + h) - 6f(x + 2h) + f(x + 3h)}{12h} + O(h^4) .$$

(ii) Del apartado anterior sabemos que

$$f'(x) = F(h) + a_1 h^4 + O(h^5) ,$$

siendo $a_1 \in \mathbb{R}$ una constante independiente de h y

$$F(h) = \frac{-3f(x-h) - 10f(x) + 18f(x+h) - 6f(x+2h) + f(x+3h)}{12h} .$$

Mediante extrapolación de Richardson se tiene

$$a_1 h^4 = \frac{F(h) - F(2h)}{2^4 - 1} + O(h^5) .$$

Entonces

$$f'(x) = F(h) + \frac{F(h) - F(2h)}{2^4 - 1} + O(h^5) = \frac{1}{15}(16F(h) - F(2h)) + O(h^5).$$

Desarrollando esta expresión se llega a

$$\begin{aligned} f'(x) = & \frac{1}{360h} [3f(x-2h) - 96f(x-h) - 310f(x) + 576f(x+h) \\ & - 210f(x+2h) + 32f(x+3h) + 6f(x+4h) - f(x+6h)] \\ & + O(h^5) . \end{aligned}$$

Problema 5.10 Considerar la función $f(x) = \sin x$ definida en el intervalo $[0, \pi/2]$. Sea $P_n(x)$ el polinomio de grado menor o igual que n que interpola a $f(x)$ en $n+1$ nodos arbitrarios pertenecientes al intervalo $[0, \pi/2]$.

- (i) Obtener una cota del error absoluto $\Delta(n)$ cometido en la aproximación $f'(x_j) \approx P'_n(x_j)$ en función únicamente de n , siendo x_j cualquier nodo.
- (ii) Teniendo en cuenta el valor de $\lim_{n \rightarrow \infty} \Delta(n)$, ¿qué se puede decir del fenómeno Runge en este problema?

Solución. (i) El error absoluto $\Delta(n)$ cometido en la aproximación $f'(x_j) \approx P'_n(x_j)$ viene dado por

$$\Delta(n) = |f'(x_j) - P'_n(x_j)| = \frac{|f^{(n+1)}(\xi)|}{(n+1)!} \prod_{\substack{i=0 \\ i \neq j}}^n |x_j - x_i| ,$$

siendo $\xi \in (0, \pi/2)$. Puesto que todos los nodos pertenecen al intervalo $[0, \pi/2]$, es claro que $|x_j - x_i| \leq \pi/2$, de modo que

$$\prod_{\substack{i=0 \\ i \neq j}}^n |x_j - x_i| \leq \left(\frac{\pi}{2}\right)^n .$$

Por otra parte, como $f(x) = \sin x$, se verifica que $|f^{(n+1)}(\xi)| \leq 1$. Entonces, obtenemos la acotación

$$\Delta(n) \leq \frac{\left(\frac{\pi}{2}\right)^n}{(n+1)!}.$$

(ii) Debido a la cota obtenida en el apartado anterior se tiene que

$$0 \leq \Delta(n) \leq \frac{\left(\frac{\pi}{2}\right)^n}{(n+1)!}.$$

Como

$$\lim_{n \rightarrow \infty} \frac{\left(\frac{\pi}{2}\right)^n}{(n+1)!} = 0,$$

se tiene que $\lim_{n \rightarrow \infty} \Delta(n) = 0$. Entonces, en el problema de aproximar $f(x)$ por $P_n(x)$ en el intervalo $[0, \pi/2]$ no se produce el fenómeno Runge.

Capítulo 6

Integración Numérica

6.1. Resumen de teoría

Dada una función $f(x)$ *integrable Riemann* en el intervalo $[a, b]$, nos planteamos el problema de calcular la integral de Riemann

$$\int_a^b f(x) \, dx .$$

Se sabe que no siempre es posible calcular la anterior integral mediante métodos analíticos. Por ejemplo si una función no admite primitiva expresable mediante funciones elementales no podremos utilizar la regla de Barrow para la determinación de la integral definida. Algunos ejemplos conocidos son

$$\int \exp(-x^2) \, dx , \quad \int \frac{\sin x}{x} \, dx ,$$

entre otros. La necesidad de disponer de técnicas numéricas para el cálculo de integrales definidas también surge cuando no conocemos la función $f(x)$ de forma explícita pero sin embargo conocemos una tabulación de la misma para un cierto conjunto de abscisas.

En primer lugar, consideramos la partición equiespaciada del intervalo de integración $[a, b]$ dada por

$$x_i = a + ih \quad \text{para } i = 0, 1, \dots, n, \quad \text{siendo } h := \frac{b-a}{n} . \quad (6.1)$$

Definimos $f(x_k) = f_k$. Si se aproxima f por el polinomio que interpola a f en los nodos x_i se obtiene la *fórmula de Newton-Cotes*

$$\int_a^b f(x) \, dx \approx h \sum_{k=0}^n \alpha_k f_k , \quad (6.2)$$

siendo

$$\alpha_k := \int_0^n \phi_k(t) \, dt , \quad (6.3)$$

con

$$\phi_k(t) = \prod_{\substack{j=0 \\ j \neq k}}^n \frac{t-j}{k-j}, \quad k = 0, \dots, n. \quad (6.4)$$

Algunos casos particulares son los siguientes.

■ *Fórmula de los Trapecios:*

$$\int_a^b f(x) dx \approx \frac{h}{2}(f_0 + f_1) = \frac{b-a}{2}(f(a) + f(b)). \quad (6.5)$$

■ *Fórmula de Simpson:*

$$\int_a^b f(x) dx \approx \frac{h}{3}(f_0 + 4f_1 + f_2) = \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]. \quad (6.6)$$

Para determinar los coeficientes α_i de la fórmula de cuadratura numérica

$$\int_a^b f(x) dx \approx \sum_{i=0}^n \alpha_i f(x_i), \quad (6.7)$$

donde f es una función real continua en el intervalo cerrado $[a, b] \subset \mathbb{R}$ y los nodos x_i son dados y no necesitan ser equiespaciados, puede utilizarse un procedimiento alternativo llamado *método de los coeficientes indeterminados*. El método es muy práctico cuando los nodos son equiespaciados o cuando el valor de n es grande. El procedimiento consiste en imponer que (6.7) tenga al menos un grado de precisión n , es decir, sea exacta para cualquier $f \in \mathbb{R}_n[x]$ polinomio de grado menor o igual que n . Imponiendo dicha condición a la base canónica $\{1, x, x^2, \dots, x^n\}$ de $\mathbb{R}_n[x]$ se obtiene el siguiente sistema lineal de $n+1$ ecuaciones para las $n+1$ incógnitas α_i

$$\begin{aligned} \sum_{i=0}^n \alpha_i &= b-a, \\ \sum_{i=0}^n \alpha_i x_i &= \frac{1}{2}(b^2 - a^2), \\ &\vdots \\ \sum_{i=0}^n \alpha_i x_i^n &= \frac{1}{n+1}(b^{n+1} - a^{n+1}). \end{aligned}$$

Observemos que el sistema es compatible y determinado puesto que el determinante de la matriz de los coeficientes del sistema es de Vandermonde

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_0 & x_1 & \cdots & x_n \\ x_0^2 & x_1^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_0^n & x_1^n & \cdots & x_n^n \end{vmatrix} = \prod_{k>i} (x_k - x_i) \neq 0,$$

ya que hemos supuesto los nodos x_i distintos entre sí.

Una estrategia muy habitual consiste en no aplicar las fórmulas de Newton-Cotes sobre todo el intervalo de integración $[a, b]$. En su lugar, dichas fórmulas se aplicarán sobre subintervalos de $[a, b]$. Tomando $h = (b - a)/n$, algunos casos particulares son los siguientes.

- *Fórmula de los Trapecios compuesta:*

$$\begin{aligned} \int_a^b f(x) dx &= T_h(f) + O(h^2) \\ &:= h \left(\frac{f_0}{2} + f_1 + \cdots + f_{n-1} + \frac{f_n}{2} \right) - \frac{h^2}{12} (b-a) f^{(2)}(\xi), \end{aligned}$$

con $\xi \in (a, b)$.

- *Fórmula de Simpson compuesta:*

$$\begin{aligned} \int_a^b f(x) dx &= S_h(f) + O(h^4) = \frac{h}{3} (f_0 + 4f_1 + 2f_2 + \cdots + 4f_{n-1} + f_n) \\ &\quad - \frac{h^4}{180} (b-a) f^{(4)}(\xi), \end{aligned}$$

con $\xi \in (a, b)$.

El *método de Romberg* consiste en una aplicación repetitiva de la extrapolación de Richardson. Si se parte de la regla trapezoidal compuesta con longitud de paso h

$$\int_a^b f(x) dx = T_h^{(1)}(f) + O(h^2),$$

se puede realizar una tabla triangular de la forma

$T_h^{(1)}(f)$		
	$T_{h/2}^{(2)}(f)$	
$T_{h/2}^{(1)}(f)$		$T_h^{(3)}(f)$
	$T_{h/4}^{(2)}(f)$	
$T_{h/4}^{(1)}(f)$		

La primera columna consta de los valores obtenidos al aplicar la regla de los trapecios con diferentes pasos $h, h/2, h/4$, etc... La segunda columna se genera de la siguiente forma recurrente

$$T_h^{(2)}(f) = \frac{1}{3} [4T_{h/2}^{(1)}(f) - T_h^{(1)}(f)], \quad T_{h/2}^{(2)}(f) := \frac{1}{3} [4T_{h/4}^{(1)}(f) - T_{h/2}^{(1)}(f)], \quad \dots$$

Se puede demostrar por inducción que

$$T_h^{(k)}(f) = \frac{4^{k-1} T_{h/2}^{(k-1)}(f) - T_h^{(k-1)}(f)}{4^{k-1} - 1}. \quad (6.8)$$

Otra técnica utilizada en la integración numérica se conoce como *integración Gaussiana*. Ya se ha visto la forma de obtener los coeficientes α_i de la fórmula de integración numérica (6.2) de manera que se obtiene un grado de precisión al menos de n cuando los $n + 1$ nodos x_0, x_1, \dots, x_n están fijados. Ahora, permitiremos una elección tanto de los coeficientes α_i como de los nodos x_i (siendo estos no equiespaciados) de forma que se obtengan fórmulas de integración numérica con el máximo grado de precisión. Notemos que, ahora se dispone de $2(n+1) = 2n+2$ grados de libertad, de manera que cabe esperar la obtención de fórmulas con una grado de precisión $2n+1$. Observemos además que si se utiliza el método de los coeficientes indeterminados, entonces el sistema de ecuaciones a resolver ya no es lineal. Se necesitará entonces una estrategia diferente que explicamos a continuación.

Definición 6.1 *Un conjunto de polinomios $\{\psi_0, \psi_1, \dots, \psi_n, \dots\}$ constituye una familia de polinomios ortogonales en el intervalo $[a, b]$ respecto a una función peso $w : (a, b) \rightarrow \mathbb{R}$ positiva y continua en (a, b) si:*

(a) *deg $\psi_i = i$ para $i = 0, 1, \dots, n, \dots$ y $\psi_0 \neq 0$.*

(b) *$\int_a^b w(x) \psi_i(x) \psi_j(x) dx = 0$ cuando $i \neq j$.*

Notemos que, debido a la condición (a) de la anterior definición, es obvio que $\{\psi_0, \psi_1, \dots, \psi_n\}$ constituye una base del espacio vectorial $\mathbb{R}_n[x]$ de los polinomios de grado menor o igual que n . En particular, esto implica que $\psi_{n+1}(x)$ es ortogonal a cualquier polinomio $P(x)$ de grado menor o igual que n puesto que

$$\begin{aligned} \int_a^b w(x) P(x) \psi_{n+1}(x) dx &= \int_a^b w(x) \left(\sum_{i=0}^n c_i \psi_i(x) \right) \psi_{n+1}(x) dx \\ &= \sum_{i=0}^n c_i \int_a^b w(x) \psi_i(x) \psi_{n+1}(x) dx = 0, \end{aligned}$$

siendo c_i las componentes de P es la anterior base de polinomios ortogonales.

El objetivo del resto de la sección es mostrar que la elección de las abscisas x_k con $k = 0, 1, \dots, n$, como los ceros de un polinomio $\psi_{n+1}(x)$ de grado $n+1$, de una familia de polinomios ortogonales, llevará siempre a fórmulas de cuadratura exactas para polinomios de grado más pequeño o igual que $2n+1$.

Proposición 6.2 *Sea $\psi_{n+1}(x)$ el polinomio ortogonal de grado $n+1$ en el intervalo $[a, b]$ respecto de la función peso $w(x)$. Entonces, el polinomio $\psi_{n+1}(x)$ tiene $n+1$ ceros reales simples x_k con $k = 0, 1, \dots, n$ en el intervalo (a, b) .*

Veamos algunos casos particulares de polinomios ortogonales.

PESO	INTERVALO	NOMBRE	EXPRESIÓN
1	$[-1, 1]$	Legendre	$P_n(x) = \frac{1}{n!2^n} \frac{d^n}{dx^n} (x^2 - 1)^n$
$\frac{1}{\sqrt{1-x^2}}$	$[-1, 1]$	Chebyshev	$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$, $T_0 = 1, T_1 = x$
$\exp(-x^2)$	$(-\infty, \infty)$	Hermite	$H_n(x) = (-1)^n \exp(x^2) \frac{d^n}{dx^n} [\exp(-x^2)]$
$\exp(-x)$	$[0, \infty)$	Laguerre	$L_n(x) = \exp(x) \frac{d^n}{dx^n} [x^n \exp(-x)]$

Consideramos la fórmula de integración numérica

$$\int_a^b w(x)f(x) dx \approx \sum_{k=0}^n \alpha_k f(x_k) . \quad (6.9)$$

Imponiendo la exactitud de la fórmula anterior para todos los polinomios de grado menor o igual que n , es decir aproximando f por el polinomio P_n que interpola a f en las abscisas x_0, x_1, \dots, x_n , se obtiene

$$\alpha_k = \int_a^b w(x)\ell_k(x) dx \quad k = 0, 1, \dots, n ,$$

siendo ℓ_k el k -ésimo polinomio básico de Lagrange.

Sea $p_{2n+1}(x)$ un polinomio de grado $n+1 \leq \text{gr } p_{2n+1} \leq 2n+1$. Sean $c_n(x)$ y $r_n(x)$ los polinomios de grado menor o igual que n que son respectivamente el cociente y el resto obtenidos de hacer la división entre $p_{2n+1}(x)$ y $\psi_{n+1}(x)$, es decir

$$p_{2n+1}(x) = c_n(x)\psi_{n+1}(x) + r_n(x) .$$

Usando la ortogonalidad entre $c_n(x)$ y $\psi_{n+1}(x)$ y teniendo en cuenta que la fórmula (6.9) es exacta para polinomios de grado menor o igual que n se tiene

$$\begin{aligned} \int_a^b w(x)p_{2n+1}(x) dx &= \int_a^b w(x)c_n(x)\psi_{n+1}(x) dx + \int_a^b w(x)r_n(x) dx \\ &= \int_a^b w(x)r_n(x) dx = \sum_{k=0}^n \alpha_k r_n(x_k) . \end{aligned}$$

Finalmente, teniendo en cuenta que $\psi_{n+1}(x_k) = 0$ para $k = 0, 1, \dots, n$ obtenemos la exactitud de (6.9) para polinomios de grado menor o igual que $2n+1$, es decir

$$\begin{aligned} \sum_{k=0}^n \alpha_k p_{2n+1}(x_k) &= \sum_{k=0}^n \alpha_k c_n(x_k)\psi_{n+1}(x_k) + \sum_{k=0}^n \alpha_k r_n(x_k) \\ &= \sum_{k=0}^n \alpha_k r_n(x_k) = \int_a^b w(x)p_{2n+1}(x) dx . \end{aligned}$$

Hemos demostrado hasta el momento que las fórmulas gaussianas tienen al menos un grado de precisión $2n+1$. Veamos que este es en realidad su

máximo grado de precisión. Para verlo basta con tomar el polinomio $\Pi^2(x) = \prod_{i=0}^n (x - x_k)^2$ de grado $2n + 2$ y comprobar que

$$0 < \int_a^b w(x) \Pi^2(x) dx \neq \sum_{k=0}^n \alpha_k \Pi^2(x_k) = 0 .$$

El uso de la fórmula de Gauss–Chebychev es la siguiente:

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx \approx \frac{\pi}{n+1} \sum_{k=0}^n f(x_k) ,$$

siendo $x_k = \cos\left(\frac{2k+1}{2n+2}\pi\right)$.

Teorema 6.3 Sea $f \in C^{2n+2}[a, b]$. Entonces existe un $\xi \in (a, b)$ tal que el error Δ cometido en la fórmula de integración gaussiana con $n+1$ puntos viene dado por

$$\Delta := \int_a^b w(x) f(x) dx - \sum_{k=0}^n \alpha_k f(x_k) = \frac{f^{(2n+2)}(\xi)}{(2n+2)!} \int_a^b w(x) \psi_{n+1}^2(x) dx ,$$

donde los nodos x_k con $k = 0, 1, \dots, n$ son las raíces del polinomio ortogonal mónico $\psi_{n+1}(x)$ de grado $n+1$ en el intervalo $[a, b]$ respecto al peso $w(x)$.

6.2. Problemas resueltos

Problema 6.1 Aproximar el valor de la integral

$$\int_0^1 \frac{\sin x}{x} dx$$

- (i) Dividiendo el intervalo de integración $[0, 1]$ en 5 subintervalos y aplicando la regla de los Trapecios compuesta. Realizar los cálculos truncando a tres decimales.
- (ii) Utilizar el algoritmo de integración de Romberg para obtener una mejora en el cálculo efectuado en el apartado (i).

Solución. (i) Si dividimos el intervalo de integración $[0, 1]$ en 5 subintervalos, en realidad estamos haciendo una partición equiespaciada con los puntos

$$x_0 = 0 , \quad x_1 = 0,2 , \quad x_2 = 0,4 , \quad x_3 = 0,6 , \quad x_4 = 0,8 , \quad x_5 = 1 ,$$

de manera que $n = 5$ o equivalentemente la longitud de paso es $h = 0,2$. Tomando la función que integramos $f(x) = \sin x/x$ y aplicando la regla de los

Trapecios compuesta tenemos que

$$\begin{aligned}
 \int_0^1 \frac{\sin x}{x} dx &\approx T_h(f) := h \left(\frac{f_0}{2} + f_1 + f_2 + f_3 + f_4 + \frac{f_5}{2} \right) \\
 &= 0,2 \left(\frac{f(0)}{2} + f(0,2) + f(0,4) + f(0,6) + f(0,8) + \frac{f(1)}{2} \right) \\
 &= 0,2 \left(\frac{1}{2} + 0,993 + 0,973 + 0,941 + 0,896 + \frac{0,841}{2} \right) \\
 &= 0,944 .
 \end{aligned}$$

NOTA: Observar que en realidad la función $f(x)$ no está definida en $x = 0$ puesto que presenta una indeterminación del tipo $0/0$. Sin embargo presenta en $x = 0$ una discontinuidad evitable ya que

$$\lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} \frac{\sin x}{x} = 0 .$$

Entonces, en la fórmula $T_h(f)$ se ha tomado $\lim_{x \rightarrow 0} f(x)$ en lugar de $f(0)$.

(ii) Para utilizar el algoritmo de Romberg a partir del cálculo efectuado en el apartado (i) hemos de calcular previamente

$$\begin{aligned}
 T_{h/2}(f) &:= \frac{h}{2} \left(\frac{f_0}{2} + f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7 + f_8 + f_9 + \frac{f_{10}}{2} \right) \\
 &= 0,1 \left(\frac{f(0)}{2} + f(0,1) + f(0,2) + f(0,3) + f(0,4) + f(0,5) + \right. \\
 &\quad \left. f(0,6) + f(0,7) + f(0,8) + f(0,9) + \frac{f(1)}{2} \right) \\
 &= 0,1 \left(\frac{1}{2} + 0,998 + 0,993 + 0,985 + 0,973 + 0,958 + \right. \\
 &\quad \left. 0,941 + 0,920 + 0,896 + 0,870 + \frac{0,841}{2} \right) = 0,945
 \end{aligned}$$

Finalmente, utilizamos el algoritmo de integración Romberg para obtener la mejora

$$\begin{aligned}
 \int_0^1 \frac{\sin x}{x} dx &\approx T_h^{(2)}(f) := \frac{1}{3} \left[4T_{h/2}^{(1)}(f) - T_h^{(1)}(f) \right] \\
 &= \frac{1}{3} [4 \times 0,945 - 0,944] = 0,945
 \end{aligned}$$

Problema 6.2 (i) Aproximar el valor de la integral $\int_1^6 2 + \sin(2\sqrt{x}) \, dx$ utilizando la regla de Simpson compuesta con 10 subintervalos.

(ii) Se sabe que el error absoluto Δ cometido al aproximar la integral $\int_a^b f(x) \, dx$ por la regla de Simpson compuesta $S_h(f)$ viene dado por

$$\Delta := \left| \int_a^b f(x) \, dx - S_h(f) \right| = \left| \frac{(b-a)f^{(iv)}(\xi)h^4}{180} \right| ,$$

donde $\xi \in (a, b)$. Calcular el número de puntos n que se han de tomar del intervalo $[2, 7]$ para obtener mediante la regla de Simpson compuesta una aproximación de la integral $\int_2^7 1/x \, dx$ con un error menor que 5×10^{-9} .

Solución. (i) La regla de Simpson compuesta con n subintervalos (siendo n un número par) viene dada por

$$\int_a^b f(x) \, dx \approx S_h(f) := \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + \cdots + 2f_{n-2} + 4f_{n-1} + f_n) ,$$

donde $h = (b-a)/n$. En este problema se tiene $n = 10$ por lo tanto $h = (6-1)/10 = 1/2$. En definitiva obtenemos

$$\begin{aligned} \int_1^6 f(x) \, dx \approx S_{1/2}(f) &:= \frac{1}{6} \left(f(1) + 4f(3/2) + 2f(2) + 4f(5/2) + 2f(3) \right. \\ &\quad \left. + 4f(7/2) + 2f(4) + 4f(9/2) + 2f(5) + 4f(11/2) + f(6) \right). \end{aligned}$$

Utilizando que la función $f(x) = 2 + \sin(2\sqrt{x})$ evaluamos dicha función en los nudos indicados en la fórmula anterior y obtenemos

$$\begin{aligned} \int_1^6 f(x) \, dx \approx \frac{1}{6} \left(2,90929 + 4 \times 2,63815 + 2 \times 2,30807 + 4 \times 1,97931 + \right. \\ \left. 2 \times 1,68305 + 4 \times 1,43530 + 2 \times 1,24319 + 4 \times 1,10831 + \right. \\ \left. 2 \times 1,02872 + 4 \times 1,00024 + 1,01735 \right) = 8,18301 . \end{aligned}$$

(ii) Sea la función $f(x) = 1/x$. La derivada cuarta de esta función viene dada por $f^{(iv)}(x) = 24/x^5$. Como $f^{(iv)}(x)$ es una función positiva y decreciente en el intervalo $[2, 7]$ podemos concluir que para cualquier $\xi \in [2, 7]$ se puede realizar la acotación

$$\left| f^{(iv)}(\xi) \right| \leq \left| f^{(iv)}(2) \right| = \frac{3}{4} .$$

De esta forma se consigue la siguiente acotación del error

$$\Delta = \left| \frac{(b-a)f^{(iv)}(\xi)h^4}{180} \right| = \left| \frac{(7-2)f^{(iv)}(\xi)h^4}{180} \right| \leq \left| \frac{(7-2)\frac{3}{4}h^4}{180} \right| = \frac{h^4}{48} .$$

Recordando la relación existente entre la longitud de paso h y el número de puntos n se tiene que $h = (7 - 2)/n$. De esta forma concluimos que

$$\Delta \leq \frac{(5/n)^4}{48} = \frac{625}{48n^4}.$$

Finalmente, imponiendo que el error sea menor que 5×10^{-9} se tiene

$$\Delta \leq \frac{625}{48n^4} < 5 \times 10^{-9},$$

de donde se puede despejar n para obtener

$$n > \sqrt[4]{\frac{125}{48} \times 10^9} = 225,9.$$

Como $n \in \mathbb{N}$ y además ha de ser par tomaremos $n = 226$.

Problema 6.3 *Se quiere construir un tejado ondulado de aluminio usando una máquina que comprime la plancha plana inicial y la transforma en una plancha cuya sección transversal tiene la forma de la función $g(x) = \sin x$. Se sabe que la longitud transversal del tejado ondulado ha de ser 20 metros.*

- (i) *Demostrar que la longitud transversal L de la plancha plana inicial viene dada por una integral elíptica de segunda especie $E(k, \phi) = \int_0^\phi \sqrt{1 - k^2 \sin^2 x} dx$.*
- (ii) *Aproximar la longitud L utilizando la regla de Simpson compuesta con $n = 4$ subintervalos.*

Solución. (i) La longitud L vendrá dada por

$$\begin{aligned} L &= \int_0^{20} \sqrt{1 + [g'(x)]^2} dx = \int_0^{20} \sqrt{1 + \cos^2 x} dx = \int_0^{20} \sqrt{2 - \sin^2 x} dx \\ &= \sqrt{2} \int_0^{20} \sqrt{1 - \frac{1}{2} \sin^2 x} dx = \sqrt{2} E(1/\sqrt{2}, 20). \end{aligned}$$

(ii) Sea $L = \int_0^{20} f(x) dx$, donde $f(x) = \sqrt{1 + \cos^2 x}$. La regla de Simpson compuesta con $n = 4$ nos permite aproximar la longitud L de la forma

$$\begin{aligned} L &\approx \frac{h}{3} (f_0 + 4f_1 + 2f_2 + 4f_3 + f_4) \\ &= \frac{20/4}{3} (f(0) + 4f(5) + 2f(10) + 4f(15) + f(20)) \\ &= \frac{5}{3} (\sqrt{2} + 4 \times 1,03945 + 2 \times 1,30539 + 4 \times 1,25584 + 1,08006) \\ &= 23,8104 \end{aligned}$$

Problema 6.4 Considerar la integral $I = \int_0^1 f(x) dx$ siendo $f(x) = \cos x / \sqrt{x}$.

- (i) ¿Qué problema existe si se pretende aproximar la integral I por alguna regla de Newton-Cotes?
- (ii) Utilizando que $f(x) = f(x) - \frac{1}{\sqrt{x}} + \frac{1}{\sqrt{x}}$, descomponer la integral I en dos integrales de la forma $I = I_1 + I_2$, siendo $I_1 = \int_0^1 \frac{1}{\sqrt{x}} dx$. Calcular la integral I mediante el cálculo exacto de I_1 y la aproximación numérica de I_2 por la fórmula compuesta de Simpson con 5 abscisas.

Solución. (i) Notemos que la integral I es impropia de segunda especie puesto que $\lim_{x \rightarrow 0^+} f(x) = \infty$, de manera que el integrando $f(x)$ no está acotado en el intervalo de integración $[0, 1]$. No es posible aplicar una regla de integración de Newton-Cotes en este caso debido a que dichas reglas utilizan siempre los extremos del intervalo de integración para evaluar la función integrando.

(ii) Realizando la descomposición indicada se tiene $I = I_1 + I_2$, siendo

$$I_1 = \int_0^1 \frac{1}{\sqrt{x}} dx, \quad I_2 = \int_0^1 \frac{\cos x - 1}{\sqrt{x}} dx.$$

La primera integral, aunque impropia, es fácilmente calculable siendo su valor

$$I_1 = \int_0^1 \frac{1}{\sqrt{x}} dx = \lim_{\epsilon \rightarrow 0^+} \int_{\epsilon}^1 \frac{1}{\sqrt{x}} dx = \lim_{\epsilon \rightarrow 0^+} [2\sqrt{x}]_{\epsilon}^1 = \lim_{\epsilon \rightarrow 0^+} (2 - 2\sqrt{\epsilon}) = 2.$$

Por lo que se refiere a la segunda integral I_2 observemos que no es impropia debido a que

$$\lim_{x \rightarrow 0^+} \frac{\cos x - 1}{\sqrt{x}} = \frac{0}{0} = \lim_{x \rightarrow 0^+} -2\sqrt{x} \sin x = 0,$$

donde se ha aplicado la regla de l'Hopital para resolver la indeterminación. Para obtener la aproximación numérica de I_2 por la fórmula compuesta de Simpson con 5 abscisas, realizamos a continuación una tabulación equiespaciada de la función $g(x) = (\cos x - 1)/\sqrt{x}$ en el intervalo $[0, 1]$ con una longitud de paso $h = 1/4$.

x_i	0	1/4	1/2	3/4	1
$g(x_i)$	0	-0.0621752	-0.173124	-0.309819	-0.459698

Aplicando la regla de Simpson compuesta se tiene

$$\begin{aligned} I_2 &= \int_0^1 g(x) dx \approx \frac{h}{3} [g_0 + 4g_1 + 2g_2 + 4g_3 + g_4] \\ &= \frac{1}{12} [0 + 4 \times (-0,0621752) + 2 \times (-0,173124) + 4 \times (-0,309819) \\ &\quad + (-0,459698)] = \frac{1}{12} (-2,29392) = -0,19116. \end{aligned}$$

Finalmente la aproximación obtenida de la integral I es

$$I \approx 2 - 0,19116 = 1,80884 .$$

Problema 6.5 Considerar la integral

$$I = \int_{-1}^1 \frac{e^x}{\sqrt{1-x^2}} dx .$$

- (i) Aproximar el valor de I utilizando el algoritmo de Gauss-Chebyshev con 3 puntos.
- (ii) Si en vez de tener la función e^x en la integral I se tiene un polinomio de grado m ¿Cuál es el máximo valor de m para el cual el algoritmo usado en el anterior apartado es exacto ?

Solución. (i) El algoritmo de Gauss-Chebyshev con $n + 1$ puntos viene dado por

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx \approx \frac{\pi}{n+1} \sum_{k=0}^n f(x_k) ,$$

donde x_k son las abscisas de Chebyshev, es decir los ceros del polinomio de Chebyshev $T_{n+1}(x) = \cos[(n+1) \arccos x]$. En concreto

$$x_k = \cos \left(\frac{2k+1}{2n+2} \pi \right) , \quad k = 0, 1, \dots, n .$$

Puesto que queremos aplicar el algoritmo con 3 puntos, se tiene $n = 2$, es decir

$$\begin{aligned} \int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx &\approx \frac{\pi}{3} \left[f \left(\cos \frac{\pi}{6} \right) + f \left(\cos \frac{\pi}{2} \right) + f \left(\cos \frac{5\pi}{6} \right) \right] \\ &= \frac{\pi}{3} \left[f \left(\frac{\sqrt{3}}{2} \right) + f(0) + f \left(-\frac{\sqrt{3}}{2} \right) \right] . \end{aligned}$$

Finalmente, como en la integral I se tiene $f(x) = e^x$, se concluye que

$$\begin{aligned} I &= \int_{-1}^1 \frac{e^x}{\sqrt{1-x^2}} dx \approx \frac{\pi}{3} \left[e^{\sqrt{3}/2} + e^0 + e^{-\sqrt{3}/2} \right] \\ &= \frac{\pi}{3} [2,37744 + 1 + 0,42062] = 3,79806 \frac{\pi}{3} = 3,97732 . \end{aligned}$$

(ii) Cualquier método de integración Gaussiana con $n + 1$ puntos es exacto cuando la función $f(x)$ es sustituida por un polinomio arbitrario de grado menor o igual que $2n + 1$. Por lo tanto el valor máximo de m es $m = 2n + 1$. En nuestro caso, como $n = 2$, el valor máximo de m es $m = 5$.

Problema 6.6 Una ecuación integral de segunda clase de Fredholm para la función $y(x)$ tiene la forma

$$y(x) = f(x) + \int_a^b K(x, t)y(t) dt ,$$

donde se conocen la función $f(x)$ y el llamado núcleo $K(x, t)$, así como las constantes a y b . Se pretende aproximar la función $y(x)$ dando su tabulación $y(x_k)$ con $k = 0, 1, \dots, n$, donde $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ es una partición equiespaciada del intervalo $[a, b]$. Para conseguirlo, se aproximarán las integrales que aparecen mediante la regla de Simpson.

- (i) Si $a = 0$, $b = 2$ y el núcleo $K(x, t) = x^2 - t$, hallar el sistema lineal de ecuaciones que deben verificar las incógnitas $y(0)$, $y(1)$ e $y(2)$.
- (ii) La tabulación $y(0)$, $y(1)$ e $y(2)$ se pretende realizar varias veces, donde en cada tabulación lo único que varía es la función $f(x)$. Razonar cual es el método más eficaz para resolver los distintos sistemas lineales que aparecen.
- (iii) Resolver por el método descrito en el apartado (ii) el sistema lineal del apartado (i) sabiendo que estamos realizando la tabulación en el caso en que $f(x) = x$.

Solución. (i) La tabulación que se pretende obtener es la siguiente

$$y(x_k) = f(x_k) + \int_a^b K(x_k, t)y(t) dt , \quad k = 0, 1, \dots, n .$$

Puesto que $a = 0$, $b = 2$, la partición del intervalo $[0, 2]$ dada por $0 = x_0 < x_1 < \dots < x_{n-1} < x_n = 2$ es equiespaciada y se quiere obtener $y(0)$, $y(1)$ e $y(2)$, de aquí se deduce que $n = 2$, es decir

$$x_0 = a < x_1 = \frac{a+b}{2} < x_2 = b .$$

Utilizando la regla de Simpson para calcular las integrales que aparecen cuando se realiza la tabulación se obtiene

$$\begin{aligned} y(a) &= f(a) + \int_a^b K(a, t)y(t) dt \approx f(a) + \frac{b-a}{6}[K(a, a)y(a) \\ &\quad + 4K(a, (a+b)/2)y((a+b)/2) + K(a, b)y(b)] , \\ y((a+b)/2) &= f((a+b)/2) + \int_a^b K((a+b)/2, t)y(t) dt \approx f((a+b)/2) \\ &\quad + \frac{b-a}{6}[K((a+b)/2, a)y(a) + 4K((a+b)/2, (a+b)/2) \\ &\quad y((a+b)/2) + K((a+b)/2, b)y(b)] , \end{aligned}$$

$$y(b) = f(b) + \int_a^b K(b, t)y(t) dt \approx f(b) + \frac{b-a}{6}[K(b, a)y(a) + 4K(b, (a+b)/2)y((a+b)/2) + K(b, b)y(b)] ,$$

que particularizado para los valores $a = 0$ y $b = 2$ reduce a

$$\begin{aligned} y(0) &= f(0) + \int_0^2 K(0, t)y(t) dt \\ &\approx f(0) + \frac{1}{3}[K(0, 0)y(0) + 4K(0, 1)y(1) + K(0, 2)y(2)] , \\ y(1) &= f(1) + \int_0^2 K(1, t)y(t) dt \\ &\approx f(1) + \frac{1}{3}[K(1, 0)y(0) + 4K(1, 1)y(1) + K(1, 2)y(2)] , \\ y(2) &= f(2) + \int_0^2 K(2, t)y(t) dt \\ &\approx f(2) + \frac{1}{3}[K(2, 0)y(0) + 4K(2, 1)y(1) + K(2, 2)y(2)] . \end{aligned}$$

Escrito en forma matricial, el anterior sistema lineal de ecuaciones para las incógnitas $y(0)$, $y(1)$ y $y(2)$ es

$$\begin{pmatrix} \frac{1}{3}K(0, 0) - 1 & \frac{4}{3}K(0, 1) & \frac{1}{3}K(0, 2) \\ \frac{1}{3}K(1, 0) & \frac{4}{3}K(1, 1) - 1 & \frac{1}{3}K(1, 2) \\ \frac{1}{3}K(2, 0) & \frac{4}{3}K(2, 1) & \frac{1}{3}K(2, 2) - 1 \end{pmatrix} \begin{pmatrix} y(0) \\ y(1) \\ y(2) \end{pmatrix} = \begin{pmatrix} -f(0) \\ -f(1) \\ -f(2) \end{pmatrix} .$$

Finalmente, teniendo en cuenta que $K(x, t) = x^2 - t$, el sistema lineal a resolver es

$$\begin{pmatrix} -1 & -\frac{4}{3} & -\frac{2}{3} \\ \frac{1}{3} & -1 & -\frac{1}{3} \\ \frac{4}{3} & 4 & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} y(0) \\ y(1) \\ y(2) \end{pmatrix} = \begin{pmatrix} -f(0) \\ -f(1) \\ -f(2) \end{pmatrix} . \quad (6.10)$$

(ii) Si se desean realizar distintas tabulaciones de la función $y(x)$ donde en cada tabulación sólo cambia la función $f(x)$, se ha de resolver para cada tabulación el sistema (6.10) donde lo único que varía son los términos independientes. Entonces, el método más eficaz para resolver los distintos sistemas lineales que aparecen es el método de la descomposición LU de la matriz A de coeficientes del sistema (6.10) definida por

$$A = \begin{pmatrix} -1 & -\frac{4}{3} & -\frac{2}{3} \\ \frac{1}{3} & -1 & -\frac{1}{3} \\ \frac{4}{3} & 4 & -\frac{1}{3} \end{pmatrix} .$$

(iii) Si $f(x) = x$, el sistema (6.10) es

$$\begin{pmatrix} -1 & -\frac{4}{3} & -\frac{2}{3} \\ \frac{1}{3} & -1 & -\frac{1}{3} \\ \frac{4}{3} & 4 & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} y(0) \\ y(1) \\ y(2) \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ -2 \end{pmatrix} .$$

Para resolver el anterior sistema lineal en primer lugar hallaremos la descomposición LU de la matriz A de coeficientes del sistema. El método que se utilizará es el de eliminación gaussiana mediante transformaciones elementales por filas sin pivotación. En concreto se obtendrá una sucesión de matrices equivalentes $A \sim A^{(1)} \sim A^{(2)}$, de manera que $U = A^{(2)}$ es una matriz triangular superior. Los multiplicadores utilizados en las transformaciones elementales formarán parte de la matriz L triangular inferior con unos en la diagonal principal. Mostramos a continuación el cálculo detallado.

$$\begin{aligned} A &= \begin{pmatrix} -1 & -\frac{4}{3} & -\frac{2}{3} \\ \frac{1}{3} & -1 & -\frac{1}{3} \\ \frac{4}{3} & 4 & -\frac{1}{3} \end{pmatrix} \sim \begin{pmatrix} -1 & -\frac{4}{3} & -\frac{2}{3} \\ 0 & -\frac{13}{9} & -\frac{5}{9} \\ 0 & \frac{20}{9} & -\frac{11}{9} \end{pmatrix} \\ &\sim \begin{pmatrix} -1 & -\frac{4}{3} & -\frac{2}{3} \\ 0 & -\frac{13}{9} & -\frac{5}{9} \\ 0 & 0 & -\frac{27}{13} \end{pmatrix} = U, \end{aligned}$$

donde las transformaciones elementales por filas efectuadas han sido

- (1) $f_i \rightarrow f_i - m_{i1}f_1$, $i = 2, 3$, $m_{21} = -1/3$, $m_{31} = -4/3$,
- (2) $f_3 \rightarrow f_3 - m_{32}f_2$, $m_{32} = -20/13$.

De esta forma, se tiene que

$$A = \begin{pmatrix} -1 & -\frac{4}{3} & -\frac{2}{3} \\ \frac{1}{3} & -1 & -\frac{1}{3} \\ \frac{4}{3} & 4 & -\frac{1}{3} \end{pmatrix} = LU = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{3} & 1 & 0 \\ -\frac{4}{3} & -\frac{20}{13} & 1 \end{pmatrix} \begin{pmatrix} -1 & -\frac{4}{3} & -\frac{2}{3} \\ 0 & -\frac{13}{9} & -\frac{5}{9} \\ 0 & 0 & -\frac{27}{13} \end{pmatrix}$$

Resolveremos el sistema lineal mediante la resolución de los dos sistemas triangulares siguientes

$$L \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ -2 \end{pmatrix}, \quad U \begin{pmatrix} y(0) \\ y(1) \\ y(2) \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}.$$

La solución del primer sistema mediante el algoritmo de sustitución hacia adelante es $(\alpha, \beta, \gamma) = (0, -1, -46/13)$. Finalmente, la solución del segundo sistema mediante el algoritmo de sustitución hacia atrás es

$$(y(0), y(1), y(2)) = (-32/27, 1/27, 46/27).$$

Problema 6.7 La función gamma se define mediante la integral impropia

$$\Gamma(t) = \int_0^{\infty} x^{t-1} e^{-x} dx ,$$

convergente para todo $t > 0$.

- (i) Una propiedad notable de dicha función es que $\Gamma(n) = (n-1)!$ para todo $n \in \mathbb{N}$. Utilizar este hecho para aproximar $\Gamma(5/2)$ mediante una interpolación polinomial cúbica por el método de las diferencias divididas de Newton.
- (ii) Aproximar el valor de $\Gamma(5/2)$ utilizando el método de integración Gauss-Laguerre con 2 puntos.

INDICACIÓN 1: Los polinomios de Laguerre son ortogonales en el intervalo $[0, \infty)$ respecto de la función peso $w(x) = e^{-x}$. Las raíces del polinomio de Laguerre de segundo grado $L_2(x)$ son $x_0 = 0,58578$, $x_1 = 3,41421$.

INDICACIÓN 2: $\int_0^{\infty} e^{-x} dx = \int_0^{\infty} x e^{-x} dx = 1$.

Solución. (i) La propiedad $\Gamma(n) = (n-1)!$ para todo $n \in \mathbb{N}$ nos permite realizar una tabulación de la función $\Gamma(x)$ en los números naturales. Como queremos aproximar $\Gamma(5/2)$ mediante una interpolación polinomial cúbica, realizamos la siguiente tabulación

x_i	1	2	3	4
$\Gamma(x_i)$	1	1	2	6

Utilizando el método de las diferencias divididas de Newton, el polinomio interpolador $P_3(x)$ se escribe como

$$\begin{aligned} P_3(x) = & f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\ & + f[x_0, x_1, x_2, x_3](x - x_0)(x - x_1)(x - x_2) . \end{aligned}$$

Realizando el esquema triangular de las diferencias divididas se obtiene

1	1			
		0		
2	1		1/2	
		1		1/3
3	2		3/2	
		4		
4	6			

de donde se sigue que

$$P_3(x) = 1 + \frac{1}{2}(x-1)(x-2) + \frac{1}{3}(x-1)(x-2)(x-3) = \frac{1}{6}x(13-9x+2x^2) .$$

La aproximación a efectuar es

$$\Gamma(5/2) \approx P_3(5/2) = \frac{5}{4} = 1,25 .$$

(ii) Los métodos gaussianos de integración con $n + 1$ puntos aproximan de la forma

$$\int_a^b w(x)f(x) dx \approx \sum_{k=0}^n \alpha_k f(x_k) ,$$

donde x_k con $k = 0, 1, \dots, n$ son las raíces del polinomio ortogonal en el intervalo $[a, b]$ respecto de la función peso $w(x)$ y α_k con $k = 0, 1, \dots, n$, son los coeficientes a determinar imponiendo un grado de precisión al menos n aunque se sabe que al final la fórmula tendrá un grado de precisión $2n + 2$.

En nuestro problema se tiene

$$\int_0^\infty e^{-x} f(x) dx \approx \alpha_0 f(0,58578) + \alpha_1 f(3,41421) . \quad (6.11)$$

Los coeficientes α_0 y α_1 se han de calcular imponiendo que la fórmula (6.11) tenga un grado de precisión al menos 1 (aunque ya sabemos que al final tendrá un grado de precisión 4). Esto es equivalente a imponer que (6.11) sea exacta para $f(x) = 1$ y $f(x) = x$, es decir

$$\begin{aligned} 1 &= \int_0^\infty e^{-x} dx = \alpha_0 + \alpha_1 , \\ 1 &= \int_0^\infty x e^{-x} dx = 0,58578\alpha_0 + 3,41421\alpha_1 . \end{aligned}$$

La solución de este sistema lineal es $\alpha_0 = 0,853551$ y $\alpha_1 = 0,146449$. En definitiva, la fórmula de Gauss-Laguerre con 2 puntos es

$$\int_0^\infty e^{-x} f(x) dx \approx 0,853551 f(0,58578) + 0,146449 f(3,41421) .$$

Si queremos aproximar $\Gamma(5/2)$ mediante la fórmula anterior se ha de tomar $f(x) = x^{5/2-1} = x^{3/2}$. De esta forma se tiene

$$\begin{aligned} \Gamma(5/2) &= \int_0^\infty x^{3/2} e^{-x} dx \approx 0,853551 \times 0,58578^{3/2} + \\ &0,146449 \times 3,41421^{3/2} = 1,30657 . \end{aligned}$$

Problema 6.8 La función de Bessel de primera especie y de orden n , con $n \in \mathbb{N}$, está definida como

$$J_n(x) = \frac{1}{\pi} \int_0^\pi \cos[x \sin(t) - nt] dt .$$

Se conoce la siguiente tabulación de $J_2(x)$

x_k	5	7	9
$J_2(x_k)$	0.0465	-0.3014	0.1448

- (i) Aproximar el valor de $J_2(6)$ por interpolación polinomial con todos los datos de la tabla adjunta y utilizando el formalismo de las diferencias divididas de Newton y acotar el error cometido.
- (ii) Si se pudiera tabular la función $J_2(x)$ en tres nodos cualesquiera del intervalo $[5, 9]$, explicar detalladamente en qué nodos se tabularía sabiendo que posteriormente dicha tabulación se va a usar para realizar una interpolación polinomial.
- (iii) Aproximar el valor de $J_2(6)$ mediante una extrapolación de Richardson según el método de Romberg.
- (iv) Aproximar el valor de $J_2'(7)$ mediante una derivación interpolatoria a partir de la tabulación dada y acotar el error absoluto asociado. Sin realizar ningún cálculo ¿se conoce alguna fórmula de aproximación de la derivada $J_2'(x)$ por un truncamiento de series de Taylor de la función $J_2(x)$ alrededor de x que proporcione la misma aproximación de $J_2'(7)$ que hemos obtenido en el cálculo anterior?
- (v) Aproximar, a partir de la expresión integral de $J_2(x)$, el valor de $J_2'(7)$ mediante el método de Simpson.

Solución. (i) Utilizando la tabla del enunciado, construimos las diferencias divididas de Newton mediante el esquema triangular siguiente

5	0.0465		
		-0.1739	
7	-0.3014		0.0992
		0.2231	
9	0.1448		

El polinomio $P_2(x)$ que interpola a la función de Bessel $J_2(x)$ en los nodos $\{x_0 = 5, x_1 = 7, x_2 = 9\}$ viene dado por

$$\begin{aligned}
 P_2(x) &= J_2[x_0] + J_2[x_0, x_1](x - x_0) + J_2[x_0, x_1, x_2](x - x_0)(x - x_1) \\
 &= 0,0465 - 0,1739(x - 5) + 0,0992(x - 5)(x - 7) \\
 &= 4,388 - 1,3643x + 0,0992x^2 .
 \end{aligned}$$

Por consiguiente, la aproximación interpolatoria es $J_2(6) \approx P_2(6) = -0,2266$.

El error absoluto Δ cometido en la aproximación $J_2(x) \approx P_2(x)$ con $x \in [5, 9]$ se puede expresar como

$$\Delta = \left| \frac{J_2'''(\xi_x)}{3!} (x-5)(x-7)(x-9) \right| ,$$

donde $\xi_x \in (5, 9)$. En el apartado (i) se ha realizado la aproximación $J_2(6) \approx P_2(6)$, por lo tanto el error absoluto asociado es

$$\Delta = \left| \frac{J_2'''(\xi)}{3!} (6-5)(6-7)(6-9) \right| = \frac{|J_2'''(\xi)|}{2} , \quad (6.12)$$

con $\xi \in (5, 9)$. A continuación obtendremos una cota de $|J_2'''(\xi)|$. En primer lugar, a partir de la expresión integral de $J_2(x)$, calculamos $J_2'''(x)$. Derivando sucesivamente respecto de x se tiene

$$\begin{aligned} J_2'(x) &= \frac{-1}{\pi} \int_0^\pi \sin t \sin[x \sin t - 2t] dt , \\ J_2''(x) &= \frac{-1}{\pi} \int_0^\pi \sin^2 t \cos[x \sin t - 2t] dt , \\ J_2'''(x) &= \frac{1}{\pi} \int_0^\pi \sin^3 t \sin[x \sin t - 2t] dt . \end{aligned}$$

Utilizando que $\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt$ para cualquier función f integrable Riemann en el intervalo $[a, b]$ y que $|\sin \theta| \leq 1$ para todo $\theta \in [0, 2\pi)$ se obtiene la siguiente cadena de acotaciones

$$\begin{aligned} |J_2'''(x)| &\leq \frac{1}{\pi} \int_0^\pi |\sin^3 t| |\sin[x \sin t - 2t]| dt \\ &\leq \frac{1}{\pi} \int_0^\pi |\sin^3 t| dt . \end{aligned}$$

Puesto que $\sin t \geq 0$ para $0 \leq t \leq \pi$, se puede quitar el valor absoluto de la última integral, de manera que

$$|J_2'''(x)| \leq \frac{1}{\pi} \int_0^\pi \sin^3 t dt .$$

Esta integral se calcula de manera exacta utilizando la identidad

$$\sin^3 t = \sin^2 t \sin t = (1 - \cos^2 t) \sin t$$

y realizando el cambio de variable $z = \cos t$. En definitiva

$$|J_2'''(x)| \leq \frac{1}{\pi} \int_0^\pi \sin^3 t dt = \frac{1}{\pi} \int_{-1}^1 (1 - z^2) dz = \frac{4}{3\pi} = 4,18879 .$$

Se tiene pues que, a partir de la ecuación (6.12), el error absoluto Δ está acotado de la forma

$$\Delta = \frac{|J_2'''(\xi)|}{2} \leq \frac{2}{3\pi} = 0,212207 .$$

(ii) Se sabe que, si la interpolación se efectúa en el intervalo $-1 \leq x \leq 1$, entonces las abscisas de Chebychev definidas de la forma

$$x_k = \cos \left(\frac{2k+1}{2n+2} \right) \in (-1, 1) , \quad k = 0, 1, \dots, n ,$$

son las que minimizan el error absoluto Δ cometido en la interpolación si se utiliza la norma del máximo o del infinito. Puesto que hemos de tomar tres nodos se tiene $n = 2$, de manera que

$$\begin{aligned} x_0 &= \cos \left(\frac{\pi}{6} \right) = \frac{\sqrt{3}}{2} = 0,866025 , \\ x_1 &= \cos \left(\frac{\pi}{2} \right) = 0 , \\ x_2 &= \cos \left(\frac{5\pi}{6} \right) = \frac{-\sqrt{3}}{2} = -0,866025 . \end{aligned}$$

Puesto que en realidad la interpolación es realizada en el intervalo $[5, 9]$, realizamos el cambio de variables $y = 7 + 2x$ que transforma el intervalo $-1 \leq x \leq 1$ en el intervalo $-5 \leq y \leq 9$. Se tiene pues que los nodos donde se tabularía la función de Bessel $J_2(x)$ son

$$\begin{aligned} y_0 &= 7 + 2x_0 = 7 + \sqrt{3} = 8,73205 , \\ y_1 &= 7 + 2x_1 = 7 + 2 \times 0 = 7 , \\ y_2 &= 7 + 2x_2 = 7 - \sqrt{3} = 5,26795 . \end{aligned}$$

(iii) A partir de su expresión integral se tiene que

$$J_2(6) = \frac{1}{\pi} \int_0^\pi f(t) dt , \quad f(t) = \cos[6 \sin t - 2t] .$$

Vamos a aproximar la integral $\int_0^\pi f(t) dt$ mediante una extrapolación de Richardson según el método de Romberg, es decir, partiendo de las aproximaciones proporcionadas por la fórmula de los Trapecios $T_h^{(1)}(f) = \frac{h}{2}[f(0) + f(\pi)]$ y $T_{h/2}^{(1)}(f) = h \left[\frac{f(0)}{2} + f\left(\frac{\pi}{2}\right) + \frac{f(\pi)}{2} \right]$, siendo $h = \pi$. Se tiene pues que

$$\begin{aligned} T_\pi^{(1)}(f) &= \frac{\pi}{2} [f(0) + f(\pi)] = \frac{\pi}{2} [1 + 1] = \pi = 3,14159 , \\ T_{\pi/2}^{(1)}(f) &= \frac{\pi}{2} \left[\frac{f(0)}{2} + f\left(\frac{\pi}{2}\right) + \frac{f(\pi)}{2} \right] = \frac{\pi}{2} \left[\frac{1}{2} - 0,96017 + \frac{1}{2} \right] \\ &= 0,06256 . \end{aligned}$$

Finalmente, mediante la extrapolación de Richardson con los dos valores anteriores se obtiene

$$\begin{aligned} T_{\pi}^{(2)}(f) &= \frac{1}{3} [4T_{\pi/2}^{(1)}(f) - T_{\pi}^{(1)}(f)] = \frac{1}{3} [4 \times 0,06256 - 3,14159] \\ &= -0,963783 . \end{aligned}$$

Concluimos que la aproximación obtenida es

$$J_2(6) = \frac{1}{\pi} \int_0^{\pi} f(t) dt \approx \frac{1}{\pi} T_{\pi}^{(2)}(f) = \frac{1}{\pi} \times (-0,963783) = -0,30678 .$$

(iv) Una aproximación de $J_2'(7)$ mediante derivación interpolatoria con todos los datos de la tabla adjunta consiste en realizar la aproximación $J_2'(7) \approx P_2'(7)$ siendo $P_2(x)$ el polinomio de interpolación calculado en el apartado (i). Puesto que $P_2(x) = 0,0465 - 0,1739(x-5) + 0,0992(x-5)(x-7)$, se tiene

$$J_2'(7) \approx P_2'(7) = -0,1739 + 0,0992(7-5) = 0,0245 .$$

Si definimos $\Delta' = |J_2'(7) - P_2'(7)|$ como el error absoluto cometido en la anterior aproximación de la derivada, se sabe que

$$\Delta' = \left| \frac{J_2'''(\xi)}{3!} (7-5)(7-9) \right| = \frac{2}{3} |J_2'''(\xi)| , \quad \xi \in (5, 9) .$$

En el apartado (i) hemos obtenido una cota de $|J_2'''(x)|$ para cualquier valor de $x \in \mathbb{R}$. En concreto se tiene $|J_2'''(x)| \leq \frac{4}{3\pi}$. Por lo tanto tenemos la siguiente cota del error Δ' cometido en la derivación interpolatoria efectuada anteriormente

$$\Delta' \leq \frac{2}{3} \frac{4}{3\pi} = \frac{8}{9\pi} = 0,282942 .$$

Por último, recordemos que las fórmulas de derivación interpolatoria con nodos equiespaciados $x_k = x_0 + kh$, siendo h la longitud de paso, son totalmente equivalentes a las fórmulas que aproximan la derivada obtenidas a partir de un truncamiento de series de Taylor. Puesto que los nodos $\{x_0 = 5, x_1 = 7, x_2 = 9\}$ donde se ha realizado la interpolación del apartado (i) son equiespaciados con paso $h = 2$ y además, en este apartado se ha aproximado la derivada de la función $J_2(x)$ en el nodo central x_1 se tiene que la fórmula de derivación aproximada equivalente obtenida por un método de Taylor debe ser la fórmula de diferencias centradas

$$J_2'(x) \approx \frac{J_2(x+h) - J_2(x-h)}{2h} .$$

La comprobación de este resultado es la siguiente. Aplicando la fórmula de diferencias centradas se obtiene

$$J_2'(7) \approx \frac{J_2(9) - J_2(5)}{4} = 0,0245 ,$$

que coincide con la aproximación $J'_2(7) \approx P'_2(7)$ realizada anteriormente.

(v) En este apartado se pretende aproximar el valor de $J'_2(7)$ por el método de Simpson. En consecuencia se debe de expresar $J'_2(7)$ mediante una integral de Riemann. Esta expresión integral la hemos obtenido en el apartado (i). En concreto

$$J'_2(7) = \frac{-1}{\pi} \int_0^\pi g(t) dt, \quad g(t) = \sin t \sin[7 \sin t - 2t].$$

Realizando la aproximación de la anterior integral por el método de Simpson se obtiene

$$\begin{aligned} J'_2(7) &\approx \frac{-1}{\pi} S_{\pi/2}(g) := \frac{-1}{\pi} \frac{\pi/2}{3} \left[g(0) + 4g\left(\frac{\pi}{2}\right) + g(\pi) \right] \\ &= \frac{-1}{6} [0 - 4 \times 0,656987 + 0] = 0,437991. \end{aligned}$$

Existe una forma alternativa de resolver el problema como se explica a continuación. En primer lugar se aproxima, por el método de Simpson, el valor de $J_2(x)$. Puesto que

$$J_2(x) = \frac{1}{\pi} \int_0^\pi f(t) dt, \quad f(t) = \cos[x \sin t - 2t],$$

se tiene

$$\begin{aligned} J_2(x) &\approx \frac{1}{\pi} S_{\pi/2}(f) := \frac{1}{\pi} \frac{\pi/2}{3} \left[f(0) + 4f\left(\frac{\pi}{2}\right) + f(\pi) \right] \\ &= \frac{1}{6} [1 + 4 \cos(x - \pi) + 1] = \frac{1}{3} [1 + 2 \cos(x - \pi)]. \end{aligned}$$

Finalmente, derivamos la expresión anterior

$$J'_2(x) \approx \frac{d}{dx} \left(\frac{1}{3} [1 + 2 \cos(x - \pi)] \right) = \frac{-2}{3} \sin(x - \pi),$$

y la evaluamos en $x = 7$ para obtener

$$J'_2(7) \approx \frac{-2}{3} \sin(7 - \pi) = 0,437991.$$

Problema 6.9 Aproximar el valor de $\ln 2$ a partir de la expresión integral

$$\ln 2 = \int_1^2 \frac{1}{x} dx,$$

utilizando el algoritmo de integración Romberg con una única extrapolación.

Solución. Definimos la función f de manera que

$$\ln 2 = \int_1^2 f(x) dx = \int_1^2 \frac{1}{x} dx.$$

El primer paso del método de Romberg consiste en aproximar $\ln 2$ utilizando la regla trapezoidal simple que denotaremos por $T_h^{(1)}(f)$, siendo el paso $h = 2 - 1 = 1$. De esta forma

$$\ln 2 \approx T_1^{(1)}(f) = \frac{1}{2} [f(1) + f(2)] = \frac{1}{2} \left[1 + \frac{1}{2} \right] = \frac{3}{4}.$$

A continuación disminuimos a la mitad el paso y aplicamos la regla trapezoidal compuesta para obtener

$$\begin{aligned} \ln 2 \approx T_{1/2}^{(1)}(f) &= \frac{1}{2} \left[\frac{f(1)}{2} + f(3/2) + \frac{f(2)}{2} \right] \\ &= \frac{1}{2} \left[\frac{1}{2} + \frac{2}{3} + \frac{1/2}{2} \right] = \frac{17}{24}. \end{aligned}$$

A partir de los dos cálculos efectuados se realiza una extrapolación de Richardson para obtener

$$\begin{aligned} \ln 2 \approx T_1^{(2)}(f) &= \frac{1}{3} [4T_{1/2}^{(1)}(f) - T_1^{(1)}(f)] = \frac{1}{3} \left[4 \times \frac{17}{24} - \frac{3}{4} \right] \\ &= \frac{25}{36}. \end{aligned}$$

Problema 6.10 Se ha medido la capacidad calorífica molar a presión constante c_p del nitrógeno en función de la temperatura T obteniéndose la siguiente tabla.

T (K)	300	1500	2100
c_p (cal/mol K)	6.81495	8.25375	8.60055

- (i) Hallar una aproximación de la función $c_p(T)$ mediante una expresión polinomial de segundo grado, realizando interpolación con los datos de la tabla adjunta.
- (ii) Utilizar el resultado del apartado anterior para aproximar el calor total Q absorbido por 23 moles de nitrógeno en un proceso isobárico cuasi-estático sabiendo que su temperatura inicial es 350 K y la final 800 K.
INDICACIÓN: El calor Q absorbido por n moles en un proceso isobárico cuasi-estático con temperatura inicial T_i y final T_f viene dado por la expresión $Q = n \int_{T_i}^{T_f} c_p dT$.
- (iii) Sabiendo que $|c_p'''(T)| \leq 10^{-13}$ para toda $T \in [350, 800]$, hallar una cota del error absoluto cometido en el apartado anterior.

Solución. (i) Sea $P_2(T)$ el polinomio que interpola a c_p mediante la tabulación adjunta. Utilizando el formalismo de las diferencias divididas de Newton, se tiene que

$$P_2(T) = c_p[T_0] + c_p[T_0, T_1](T - T_0) + c_p[T_0, T_1, T_2](T - T_0)(T - T_1).$$

Realizando el esquema de diferencias divididas asociado a la tabulación del enunciado obtenemos

300	6.81495		
		0.001199	
1500	8.25375		-3.45×10^{-7}
		0.000578	
2100	8.60055		

de manera que

$$\begin{aligned} P_2(T) &= 6,81495 + 0,001199(T - 300) - 3,45 \times 10^{-7}(T - 300)(T - 1500) \\ &= 6,3 + 0,00182T - 3,45 \times 10^{-7}T^2. \end{aligned}$$

(ii) El calor total Q absorbido por n moles de nitrógeno en el proceso isobárico con temperatura inicial T_i y final T_f es $Q = n \int_{T_i}^{T_f} c_p(T) dT$. Puesto que desconocemos la función $c_p(T)$, utilizaremos el resultado del apartado anterior para aproximar $c_p(T) \approx P_2(T)$ de lo que deduciremos una aproximación $Q \approx Q^*$. Más concretamente, se tiene

$$\begin{aligned} Q &\approx Q^* := n \int_{T_i}^{T_f} P_2(T) dT \\ &= 23 \int_{350}^{800} (6,3 + 0,00182T - 3,45 \times 10^{-7}T^2) dT = 74795,4 \text{ cal} \end{aligned}$$

(iii) Al realizar la aproximación $Q \approx Q^*$, el error absoluto Δ cometido es

$$\Delta = |Q - Q^*| = n \left| \int_{T_i}^{T_f} (c_p(T) - P_2(T)) dT \right| \leq n \int_{T_i}^{T_f} |c_p(T) - P_2(T)| dT,$$

donde en el último paso se han tenido en cuenta las propiedades del valor absoluto y de la integral de Riemann. Puesto que $P_2(T)$ es el polinomio que interpola a c_p en los nodos T_0, T_1 y T_2 , se tiene que existe un valor $\xi_T \in (T_i, T_f)$ tal que $c_p(T) - P_2(T) = c_p'''(T)\pi(T)/3!$, siendo $\pi(T) = (T - T_0)(T - T_1)(T - T_2)$. En definitiva obtenemos que

$$\Delta \leq \frac{n}{3!} \int_{T_i}^{T_f} |c_p'''(T)| |(T - T_0)(T - T_1)(T - T_2)| dT,$$

de manera que, teniendo en cuenta que $|c_p'''(T)| \leq 10^{-13}$ para toda $T \in [350, 800]$ y substituyendo los datos del enunciado obtenemos

$$\begin{aligned} \Delta &\leq 10^{-13} \frac{23}{3!} \int_{350}^{800} |(T - 300)(T - 1500)(T - 2100)| dT \\ &= 10^{-13} \frac{23}{6} \int_{350}^{800} (T - 300)(T - 1500)(T - 2100) dT \\ &= 10^{-13} \frac{23}{6} \int_{350}^{800} (-945000000 + 4230000T - 3900T^2 + T^3) dT \\ &= 10^{-13} \frac{23}{6} 158048437500 = 0,0605852. \end{aligned}$$

donde se ha tenido en cuenta que el polinomio $(T - 300)(T - 1500)(T - 2100)$ es positivo en el intervalo de integración.

Problema 6.11 *La distribución normal es una distribución de probabilidad para una variable aleatoria continua Z frecuentemente utilizada en las aplicaciones estadísticas. La densidad de probabilidad $f(x)$ asociada a dicha distribución es la llamada campana de Gauss y viene dada por*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

siendo μ y σ la media y desviación típica respectivamente. De este modo, la probabilidad $P(a \leq Z \leq b)$ de que en un experimento aleatorio Z tome un valor entre a y b viene dada por $P(a \leq Z \leq b) = \int_a^b f(x) dx$. Sabiendo que $\mu = 0$ y $\sigma = 1$, aproximar la probabilidad de que Z tome un valor entre 10 y 20 mediante la fórmula de Simpson compuesta con 4 subintervalos.

Solución. La probabilidad de que Z tome un valor entre 10 y 20 es

$$P(10 \leq Z \leq 20) = \int_{10}^{20} f(x) dx,$$

siendo

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Observemos que la función $f(x)$ no admite primitiva expresable mediante funciones elementales, de modo que, su aproximación numérica está perfectamente justificada. Utilizando ahora la fórmula de Simpson compuesta $S_h(f)$ con $n = 4$ subintervalos se tiene que la longitud de paso vale $h = (b - a)/n = 5/2$, de modo que

$$\begin{aligned} P(10 \leq Z \leq 20) &\approx S_{5/2}(f) = \frac{5/2}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + f_4) \\ &= \frac{5}{6}[f(10) + 4f_1(25/2) + 2f(15) + 4f(35/2) + f(20)] \\ &= \frac{5}{6} \frac{1}{\sqrt{2\pi}} \left[e^{-10^2/2} + 4e^{-(25/2)^2/2} + 2e^{-15^2/2} \right. \\ &\quad \left. + 4e^{-(35/2)^2/2} + e^{-20^2/2} \right] \\ &= 0,332452[1,92875 \times 10^{-22}] = 6,41217 \times 10^{-23}. \end{aligned}$$

Problema 6.12 *Utilizar el método de los coeficientes indeterminados para hallar una fórmula de integración numérica del tipo*

$$\int_1^3 f(x) dx \approx af(0) + bf(2) + cf(4),$$

con $a, b, c \in \mathbb{R}$, que sea exacta para polinomios del máximo grado posible. Dar dicho grado máximo.

Solución. Se han de hallar $a, b, c \in \mathbb{R}$ de modo que

$$\int_1^3 f(x) dx = af(0) + bf(2) + cf(4) ,$$

para todo $f \in \mathbb{R}_n[x]$ polinomio de grado n , siendo n lo más grande posible. Utilizaremos la base canónica de $\mathbb{R}_n[x]$. Dicho de otro modo, como

$$\int_1^3 x^m dx = \frac{3^{m+1} - 1}{m + 1} , \quad m \in \mathbb{N} \cup \{0\} ,$$

imponiendo la exactitud de la fórmula numérica para $f(x) = x^m$ con $m = 0, 1, 2$, se obtiene el sistema lineal

$$\begin{aligned} 2 &= a + b + c , \\ 4 &= 2b + 4c , \\ \frac{26}{3} &= 4b + 16c , \end{aligned}$$

siendo su única solución

$$a = c = \frac{1}{12} , \quad b = \frac{11}{6} .$$

De momento hemos demostrado que la fórmula de aproximación numérica

$$\int_1^3 f(x) dx \approx \frac{1}{12} [f(0) + f(4)] + \frac{11}{6} f(2) ,$$

es exacta para todo $f \in \mathbb{R}_2[x]$. Sin embargo, si tomamos en dicha fórmula $f(x) = x^3$ y $f(x) = x^4$ obtenemos

$$\begin{aligned} \int_1^3 x^3 dx &= 20 = \frac{1}{12} [0^3 + 4^3] + \frac{11}{6} 2^3 , \\ \int_1^3 x^4 dx &= \frac{242}{5} \neq \frac{1}{12} [0^4 + 4^4] + \frac{11}{6} 2^4 = \frac{152}{3} , \end{aligned}$$

respectivamente. Se concluye que la fórmula numérica

$$\int_1^3 f(x) dx \approx \frac{1}{12} [f(0) + f(4)] + \frac{11}{6} f(2) ,$$

es exacta para todo $f \in \mathbb{R}_3[x]$, siendo 3 el máximo grado polinomial para el cual se verifica dicha exactitud.

Problema 6.13 Considerar la integral impropia de primera especie

$$I = \int_0^\infty \frac{e^{-x}}{1 + xe^{-x}} dx .$$

- (i) Realizar el cambio de variable $z = e^{-x}$ en la integral I . ¿Se ha convertido I en una integral propia?
- (ii) Aproximar el valor de I utilizando la regla de Simpson compuesta dividiendo el intervalo de integración en 4 subintervalos.

Solución. (i) Realizamos el cambio de variable $z = e^{-x}$. Diferenciando el cambio se tiene $dz = -e^{-x}dx$. Además, si $x = 0$ se tiene $z = 1$ y si x tiende a ∞ entonces z tiende a cero. De este modo la integral I se escribe como

$$I = \int_0^1 f(z) dz, \text{ siendo } f(z) = \frac{1}{1 - z \ln z}. \quad (6.13)$$

Notar que se han permutado los límites de integración y se ha cambiado el signo de la integral. Observar que la función $f(z)$ no está bien definida en $z = 0$ debido al término $z \ln z$ que produce una indeterminación del tipo $0 \cdot \infty$. Sin embargo, aplicando la regla de L'Hôpital se tiene

$$\lim_{z \rightarrow 0^+} z \ln z = \lim_{z \rightarrow 0^+} \frac{\ln z}{1/z} = \lim_{z \rightarrow 0^+} \frac{1/z}{-1/z^2} = - \lim_{z \rightarrow 0^+} z = 0,$$

de modo que

$$\lim_{z \rightarrow 0^+} f(z) = 1. \quad (6.14)$$

Finalmente, observamos que la función $f(z)$ está acotada en el intervalo de integración $[0, 1]$ puesto que su denominador $1 - z \ln z \neq 0$ para todo $z \in [0, 1]$. De este modo, la integral I se ha transformado en la integral propia dada por (6.13).

(ii) Utilizamos la regla de Simpson compuesta con $n = 4$ subintervalos, es decir,

$$I = \int_0^1 f(z) dz \approx S_h(f) = \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + f_4),$$

siendo $h = 1/4$ y $f_i = f(z_i)$ con $z_i = ih$ para $i = 1, 2, 3, 4$. Se tiene pues

$$\begin{aligned} I &\approx \frac{1/4}{3}[f(0) + 4f(1/4) + 2f(1/2) + 4f(3/4) + f(1)] \\ &= \frac{1}{12}[1 + 4 \times 0,742626 + 2 \times 0,742626 + 4 \times 0,82253 + 1] = 0,812156, \end{aligned}$$

donde se ha tenido en cuenta (6.14) para calcular la indeterminación $f(0)$.

Problema 6.14 Utilizar la técnica de la extrapolación de Richardson aplicada a la fórmula de Simpson compuesta para obtener una fórmula que aproxime integrales con un error menor que la fórmula de Simpson compuesta.

Solución. La fórmula de Simpson compuesta viene dada por

$$\int_a^b f(x) dx = S_h(f) + a_1 h^4 + O(h^5), \quad (6.15)$$

siendo h la longitud de paso, a_1 una constante independiente de h . Tomando la mitad de la longitud de paso se tiene

$$\int_a^b f(x) dx = S_{h/2}(f) + a_1 \left(\frac{h}{2}\right)^4 + O(h^5).$$

Restando las dos expresiones de la misma integral se obtiene

$$0 = S_h(f) - S_{h/2}(f) + a_1 h^4 \left(1 - \frac{1}{2^4}\right) + O(h^5) ,$$

de donde

$$a_1 h^4 = \frac{16}{15} [S_{h/2}(f) - S_h(f)] + O(h^5) .$$

Sustituyendo este término en (6.15) se obtiene la siguiente fórmula mejorada

$$\int_a^b f(x) dx = \tilde{F}(h) + O(h^5) ,$$

siendo

$$\tilde{F}(h) = S_h(f) + \frac{16}{15} [S_{h/2}(f) - S_h(f)] = \frac{1}{15} [16S_{h/2}(f) - S_h(f)] .$$

Problema 6.15 Considerar un anillo circular de radio R con una distribución de carga eléctrica uniforme de carga total Q . Tomando el origen de coordenadas en el centro del anillo y el eje z perpendicular al plano del anillo se puede demostrar, utilizando coordenadas esféricas (r, φ, θ) , que el potencial eléctrico $V(r, \theta)$ que crea el anillo en el punto de coordenadas (r, φ, θ) viene dado por

$$V(r, \theta) = \frac{2\kappa\lambda}{\sqrt{Rr \sin \theta}} K(\kappa) ,$$

siendo $\lambda = Q/(2\pi R)$ la densidad lineal de carga en el anillo,

$$\kappa^2 = \frac{4Rr \sin \theta}{R^2 + r^2 + 2Rr \sin \theta} ,$$

y $K(\kappa)$ la integral elíptica de primera especie

$$K(\kappa) = \int_0^{\pi/2} \frac{d\phi}{\sqrt{1 - \kappa^2 \sin^2 \phi}} .$$

(i) Utilizando el método de los trapecios compuesto con una división de 3 subintervalos, aproximar el potencial eléctrico en cualquier punto de la circunferencia $(r, \varphi, \theta) = (1, \varphi, \pi/2)$ con $0 \leq \varphi < 2\pi$ concéntrica al anillo.

DATOS: $Q = 10C$, $R = 5m$.

(ii) ¿Es necesario utilizar métodos numéricos para calcular el potencial eléctrico en cualquier punto del eje z ?

Solución. (i) Con los datos del problema $Q = 10C$, $R = 5m$ se tiene $\lambda = Q/(2\pi R) = 1/\pi$ y

$$\kappa^2 = \frac{20r \sin \theta}{25 + r^2 + 10r \sin \theta} .$$

Sobre los puntos de la circunferencia $(r, \varphi, \theta) = (1, \varphi, \pi/2)$ se obtiene $\kappa^2 = 5/9$ con lo cual el potencial eléctrico en cualquier punto de dicha circunferencia viene dado por

$$V(1, \pi/2) = \frac{2\sqrt{\frac{5}{9}}}{\sqrt{5}} K\left(\sqrt{\frac{5}{9}}\right) = \frac{2}{3\pi} K(\sqrt{5/9}) ,$$

siendo

$$K(\sqrt{5/9}) = \int_0^{\pi/2} f(\phi) d\phi ,$$

con

$$f(\phi) = \frac{1}{\sqrt{1 - \frac{5}{9} \sin^2 \phi}} .$$

Utilizamos a continuación el método de los trapecios compuesto con una división de $n = 3$ subintervalos para aproximar $K(\sqrt{5/9})$. Dicho método consiste en realizar la aproximación

$$\int_a^b f(\phi) d\phi \approx T_h(f) := h \left(\frac{f_0}{2} + f_1 + \cdots + f_{n-1} + \frac{f_n}{2} \right) ,$$

con $h = (b - a)/n$ y $f_i = f(\phi_i)$, siendo $\phi_i = a + ih$. En nuestro caso se tiene $h = \pi/6$ y $\phi_i = i\pi/6$ para $i = 0, 1, 2, 3$, de modo que

$$\begin{aligned} K(\sqrt{5/9}) &= \int_0^{\pi/2} f(\phi) d\phi \approx T_{\pi/6}(f) := \frac{\pi}{6} \left(\frac{f_0}{2} + f_1 + f_2 + \frac{f_3}{2} \right) \\ &= \frac{\pi}{6} \left(\frac{f(0)}{2} + f(\pi/6) + f(\pi/3) + \frac{f(\pi/2)}{2} \right) \\ &= \frac{\pi}{6} \left(\frac{1}{2} + 1,07763 + 1,30931 + \frac{1,5}{2} \right) = 1,9043 . \end{aligned}$$

Finalmente

$$V(1, \pi/2) = \frac{2}{3\pi} K(\sqrt{5/9}) \approx 0,4041 .$$

(ii) Nótese que los puntos del eje z verifican $\theta = 0$. De este modo, en dichos puntos se tiene $\kappa = 0$ con lo cual el potencial eléctrico en cualquier punto del eje z será

$$V(r, 0) = \frac{4\lambda}{\sqrt{r^2 + R^2}} K(0) ,$$

donde la integral elíptica de primera especie que se debe calcular reduce a la integral inmediata

$$K(0) = \int_0^{\pi/2} d\phi = \frac{\pi}{2} .$$

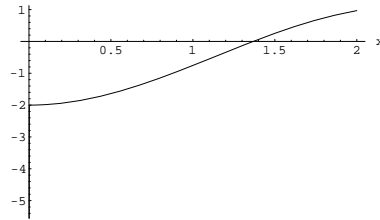
En consecuencia, no es necesario utilizar métodos numéricos para calcular el potencial eléctrico en cualquier punto del eje z .

Problema 6.16 Según la teoría de Debye, la capacidad calorífica a volumen constante C_v de los sólidos se calcula de la forma

$$C_v = \frac{9Nk}{x_D^3} \int_0^{x_D} \frac{x^4 e^x}{(e^x - 1)^2} dx ,$$

siendo N el número de átomos del cristal y $k = 1,381 \times 10^{-23} \text{ J/}^\circ\text{K}$ la constante de Boltzmann. La variable auxiliar x_D está definida como $x_D = \Theta_D/T$, siendo Θ_D la temperatura de Debye característica del material y T la temperatura.

- (i) Suponiendo que el material es cobre ($\Theta_D = 310^\circ\text{K}$) y se tiene un mol (con lo que $N = 6,023 \times 10^{23}$ átomos) y que está a una temperatura $T = 300^\circ\text{K}$, utilizar la regla de Simpson compuesta con una división de 4 subintervalos para obtener una aproximación de la capacidad calorífica C_v . NOTA: Realizar los cálculos con 4 decimales de precisión.
- (ii) Acotar el error absoluto cometido en el apartado (i), sabiendo que la gráfica de derivada cuarta de la función integrando en la expresión de C_v es:



Solución. (i) Tomando los datos $\Theta_D = 310$ y $T = 300$ se tiene $x_D = \Theta_D/T = 1,0333$. De este modo

$$C_v = 67,8472 \int_0^{1,0333} f(x) dx , \quad f(x) = \frac{x^4 e^x}{(e^x - 1)^2} .$$

Aproximando la anterior integral mediante la regla de Simpson compuesta con una división de $n = 4$ subintervalos se tiene una longitud de paso $h = 1,0333/4 = 0,2583$. Entonces

$$\begin{aligned} C_v &\approx 67,8472 S_h(f) = 67,8472 \times \frac{h}{3} (f_0 + 4f_1 + 2f_2 + 4f_3 + f_4) \\ &= 67,8472 \times \frac{0,2583}{3} (f(0) + 4f(0,2583) + 2f(0,5166) + 4f(0,7749) \\ &\quad + f(1,0333)) \end{aligned}$$

Antes de continuar notamos que la función $f(x)$ no está bien definida en $x = 0$. En particular, se obtiene la indeterminación $f(0) = 0/0$. Sin embargo, la función $f(x)$ presenta en $x = 0$ una discontinuidad evitable ya que $\lim_{x \rightarrow 0} f(x)$

existe. En concreto¹ teniendo en cuenta que $e^x - 1$ es un infinitésimo equivalente a x cuando $x \rightarrow 0$ se llega a que

$$\lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} \frac{x^4 e^x}{(e^x - 1)^2} = \lim_{x \rightarrow 0} x^2 e^x = 0 .$$

Entonces,

$$\begin{aligned} C_v &\approx 67,8472 \times \frac{0,2583}{3} (0 + 4 \times 0,0663 + 2 \times 0,2610 + 4 \times 0,5713 \\ &\quad + 0,9775) = 67,8472 \times 0,348709 = 23,659 . \end{aligned}$$

(ii) Se sabe que, si $f \in C^4[a, b]$, el error absoluto Δ cometido en la aproximación

$$\int_a^b f(x) dx \approx S_h(f) ,$$

siendo $S_h(f)$ la fórmula de Simpson compuesta con una longitud de paso h aplicada a la función f viene dado por

$$\Delta = \left| \frac{h^4}{180} (b-a) f^{(iv)}(\xi) \right| , \quad \xi \in (a, b) .$$

En nuestro caso, el error absoluto cometido en la aproximación

$$\int_0^{1,0333} f(x) dx \approx S_{0,2583}(f) ,$$

realizada en el apartado anterior es

$$\Delta = \left| \frac{0,2583^4}{180} (1,0333 - 0) f^{(iv)}(\xi) \right| = 0,000025 |f^{(iv)}(\xi)| , \quad \xi \in (0, 1,0333) .$$

Teniendo en cuenta la gráfica de la función $f^{(iv)}$, es claro que $|f^{(iv)}(\xi)| < 2$ para todo $\xi \in (0, 1,0333)$. Entonces

$$\Delta < 0,000025 \times 2 = 0,000051 .$$

Finalmente, como la aproximación efectuada en el apartado (i) ha sido

$$C_v = 67,8472 \int_0^{1,0333} f(x) dx \approx 67,8472 S_{0,2583}(f) ,$$

el error absoluto cometido es

$$67,8472 \left| \int_0^{1,0333} f(x) dx - S_{0,2583}(f) \right| = 67,8472 \Delta ,$$

de modo que una cota es

$$67,8472 \times 0,000051 = 0,00346 .$$

¹También se puede calcular el límite $\lim_{x \rightarrow 0} f(x)$ aplicando varias veces la regla de l'Hopital.

Problema 6.17 Hallar el valor que se obtiene al aproximar la integral

$$\int_0^n |x \cos(\pi x)| \, dx ,$$

siendo $n \in \mathbb{N}$ un número par, mediante la fórmula de Simpson compuesta con n subintervalos.

INDICACIÓN: Si $\{a_m\}_{m \in \mathbb{N}} \subset \mathbb{R}$ es una sucesión aritmética, entonces, la suma de los k primeros términos vale $\sum_{i=1}^k a_i = \frac{1}{2}k(a_1 + a_k)$.

Solución. Definimos $f(x) = |x \cos(\pi x)|$. Entonces, aplicando la fórmula de Simpson compuesta con n subintervalos se tiene

$$\int_0^n f(x) \, dx \approx S_h(f) = \frac{h}{3}(f_0 + 4f_1 + 2f_2 + \cdots + 4f_{n-1} + f_n) ,$$

siendo $h = (n - 0)/n = 1$ y $f_k = f(x_k) = f(0 + kh) = f(k)$ para $k = 0, 1, \dots, n$. Entonces

$$\int_0^n f(x) \, dx \approx S_1(f) = \frac{1}{3}(f(0) + 4f(1) + 2f(2) + \cdots + 4f(n-1) + f(n)) .$$

En el caso que nos ocupa, se tiene $f(k) = |k \cos(k\pi)| = |k(-1)^k| = |k| = k$, de modo que

$$\int_0^n f(x) \, dx \approx S_1(f) = \frac{1}{3}[n + 4(1 + 3 + 5 + \cdots + n-1) + 2(2 + 4 + 6 + \cdots + n-2)] .$$

Aplicando la fórmula de la suma finita de términos consecutivos de una sucesión aritmética se tiene que

$$\begin{aligned} 1 + 3 + 5 + \cdots + n-1 &= \frac{1}{2} \frac{n}{2} (1 + n-1) = \frac{n^2}{4} , \\ 2 + 4 + 6 + \cdots + n-2 &= \frac{1}{2} \left(\frac{n}{2} - 1 \right) (2 + n-2) = \frac{n}{2} \left(\frac{n}{2} - 1 \right) . \end{aligned}$$

Finalmente, sustituyendo obtenemos

$$\int_0^n f(x) \, dx \approx S_1(f) = \frac{1}{3} \left[n + 4 \frac{n^2}{4} + 2 \frac{n}{2} \left(\frac{n}{2} - 1 \right) \right] = \frac{n^2}{2} .$$

Problema 6.18 Considerar la integral impropia de segunda especie

$$I = \int_0^1 \frac{\sin x}{\sqrt{1-x^2}} \, dx .$$

(i) Aproximar el valor de I mediante la regla de los trapecios compuesta con 4 subintervalos. INDICACIÓN: Realizar el cambio de variable $x = \sin t$ para evitar la singularidad en $x = 1$.

(ii) Acotar el error absoluto cometido en el apartado anterior.

Solución. (i) Notemos que la integral del enunciado

$$I = \int_0^1 f(x) \, dx, \quad f(x) = \frac{\sin x}{\sqrt{1-x^2}},$$

es impropia de segunda especie puesto que la función integrando $f(x)$ no está acotada en el intervalo de integración $[0, 1]$. En particular, la gráfica de $f(x)$ tiene una asíntota vertical en $x = 1$. Realizando el cambio de variable $x = \sin t$ en la integral I se obtiene

$$I = \int_0^{\pi/2} \frac{\sin(\sin t)}{\sqrt{1-\sin^2 t}} \cos t \, dt = \int_0^{\pi/2} \sin(\sin t) \, dt.$$

Notemos que, tras el cambio de variable, la integral I deja de ser impropia puesto que la función integrando $g(t) = \sin(\sin t)$ está acotada en el intervalo de integración $[0, \pi/2]$.

A continuación aproximamos el valor de I mediante la regla de los trapecios compuesta con $n = 4$ subintervalos, es decir, con una longitud de paso $h = (\pi/2 - 0)/4 = \pi/8$.

$$\begin{aligned} I &\approx T_h(g) = h \left(\frac{g_0}{2} + g_1 + g_2 + g_3 + \frac{g_4}{2} \right) \\ &= \frac{\pi}{8} \left(\frac{g(0)}{2} + g(\pi/8) + g(\pi/4) + g(3\pi/8) + \frac{g(\pi/2)}{2} \right) \\ &= \frac{\pi}{8} \left(\frac{0}{2} + 0,373411 + 0,649637 + 0,797946 + \frac{0,841471}{2} \right) = 0,880325. \end{aligned}$$

(ii) El error absoluto Δ en la aproximación $\int_a^b g(t) \, dt \approx T_h(g)$ viene dado por

$$\Delta = \left| \frac{h^2}{12} (b-a) g''(\xi) \right|, \quad \xi \in (a, b).$$

En nuestro caso se tiene

$$\Delta = \frac{(\pi/8)^2}{12} (\pi/2 - 0) |g''(\xi)| = \frac{\pi^3}{1536} |g''(\xi)|,$$

siendo $\xi \in (0, \pi/2)$. Derivando 2 veces la función $g(t) = \sin(\sin t)$ se tiene $g''(t) = -\cos(\sin x) \sin x - \cos^2 x \sin(\sin x)$. Para hallar el valor máximo (es decir, la menor de las cotas superiores) de Δ se ha de obtener el valor

$$M = \max_{0 \leq t \leq \pi/2} \{|g''(t)|\}.$$

Sin embargo, el problema de extremos absolutos que se plantea para hallar M , pasa por obtener los extremos relativos de $f''(t)$, es decir, por la ecuación $f'''(t) = 0$ que resulta ser una ecuación trascendente. Entonces, optamos por hallar una cota superior de $|g''(t)|$ en el intervalo $[0, \pi/2]$ por simple inspección.

En concreto, utilizando las propiedades del valor absoluto y las acotaciones de las funciones trigonométricas, se tiene

$$\begin{aligned} |g''(t)| &= |-\cos(\sin x) \sin x - \cos^2 x \sin(\sin x)| \\ &\leq |\cos(\sin x) \sin x| + |\cos^2 x \sin(\sin x)| \\ &= |\cos(\sin x)| |\sin x| + \cos^2 x |\sin(\sin x)| \\ &\leq 1 + 1 = 2, \end{aligned}$$

para todo $t \in \mathbb{R}$. Entonces

$$\Delta = \frac{\pi^3}{1536} |g''(\xi)| \leq \frac{2\pi^3}{1536} = 0,0403728.$$

Problema 6.19 *En la descripción del fenómeno óptico conocido como difracción de Fresnel, aparece la llamada integral de Fresnel $S(z) = \int_0^z \sin(x^2) dx$ que no puede expresarse mediante funciones elementales.*

(i) *Utilizando el cambio de variable $x = 1/t$, expresar la integral impropia*

$$I = \int_1^\infty \frac{1}{t^2} \sin\left(\frac{1}{t^2}\right) dt$$

como una integral de Fresnel.

(ii) *Aproximar el valor de I mediante la regla de Simpson compuesta con 4 subintervalos.*

(iii) *Acotar el error absoluto cometido en el apartado anterior.*

Solución. (i) Realizando el cambio de variable $x = 1/t$ en la integral impropia

$$I = \int_1^\infty \frac{1}{t^2} \sin\left(\frac{1}{t^2}\right) dt,$$

se obtiene

$$I = \int_0^1 \sin(x^2) dx = S(1).$$

Como $S(1)$ es una integral propia, es claro que $I < \infty$.

(ii) Aplicamos la regla de Simpson para aproximar el valor de I . Obviamente, utilizaremos este método numérico aplicado a la expresión equivalente dada por la integral propia $S(1)$. De este modo

$$I = S(1) = \int_0^1 f(x) dx \text{ con } f(x) = \sin(x^2).$$

Aproximando I con la regla de Simpson compuesta con $n = 4$ subintervalos, se tiene una longitud de paso $h = 1/4$, de modo que

$$I \approx S_{1/4}(f) = \frac{h}{3} (f_0 + 4f_1 + 2f_2 + 4f_3 + f_4)$$

$$\begin{aligned}
&= \frac{1}{12} (f(0) + 4f(1/4) + 2f(1/2) + 4f(3/4) + f(1)) \\
&= \frac{1}{12} (0 + 4 \times 0,0624593 + 2 \times 0,247404 + 4 \times 0,533303 + 0,841471) \\
&= 0,309944 .
\end{aligned}$$

(iii) Sea $\Delta = |I - S_{1/4}(f)|$ el error absoluto cometido en el apartado anterior. Se sabe que

$$\Delta = \frac{h^4}{180} (b-a) |f^{(iv)}(\xi)| ,$$

siendo $[a, b] = [0, 1]$ el intervalo de integración y $\xi \in (0, 1)$. En resumen,

$$\Delta = \frac{1/4^4}{180} |f^{(iv)}(\xi)| ,$$

siendo $f^{(iv)}(\xi) = 4[(4\xi^4 - 3)\sin(\xi^2) - 12\xi^2\cos(\xi^2)]$. Aplicando las propiedades del valor absoluto y teniendo en cuenta que las funciones seno y coseno están acotadas entre 1 y -1 se tiene la siguiente acotación

$$\begin{aligned}
|f^{(iv)}(\xi)| &\leq 4[|4\xi^4 - 3|\sin(\xi^2)| + 12\xi^2|\cos(\xi^2)|] < 4[|4\xi^4 - 3| + 12\xi^2] \\
&< 4[3 + 12] = 60,
\end{aligned}$$

donde en la última desigualdad se ha utilizado que $\xi \in (0, 1)$. En definitiva,

$$\Delta < \frac{1/4^4}{180} 60 = 0,00130208 .$$

Problema 6.20 La teoría de estática de fluidos dice que la presión P que ejerce un fluido de densidad ρ a una profundidad x de su superficie viene dada por $P(x) = \rho gx$, siendo g la intensidad gravitatoria. Entonces, la fuerza total F que ejerce un fluido estancado una profundidad H sobre la compuerta de una presa es $F = \rho g \int_0^H xL(x) dx$ donde $L(x)$ es la anchura de la compuerta a una profundidad x .

(i) Sabiendo que el fluido es agua ($\rho = 10^3$), $H = 8$ y que se han realizado las siguientes mediciones de la función $L(x)$

x	0	2	4	6	8
$L(x)$	150	20	2.5	0.5	0

aproximar el valor de F mediante la fórmula de Simpson compuesta.

(ii) Aproximar, mediante interpolación polinómica de Newton con todos los puntos de la tabulación dada, el valor de $L(1)$.

(iii) Supongamos ahora que se dispone de más información. En concreto $L(x) = (150 - x)/e^x$. Acotar el error absoluto cometido al aproximar F mediante la fórmula de Simpson compuesta con una longitud de paso $h = 0,1$.

Solución. (i) La fuerza total F que ejerce el agua sobre la compuerta de la presa es

$$F = \rho g \int_0^8 f(x) dx, \quad f(x) = xL(x).$$

A partir de la tabulación del enunciado para la función $L(x)$ se obtiene la siguiente tabulación de la función $f(x)$.

x	0	2	4	6	8
$f(x)$	0	40	10	3	0

Notemos que las abscisas de la tabulación son equiespaciadas con una longitud de paso $h = 2$. Entonces, podemos aproximar mediante la fórmula de Simpson compuesta la integral

$$\begin{aligned} \int_0^8 f(x) dx &\approx \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + f_4) \\ &= \frac{2}{3}(0 + 4 \times 40 + 2 \times 10 + 4 \times 3 + 0) = 128. \end{aligned}$$

En definitiva, tomando la gravedad $g = 9,8$, se tiene

$$F \approx 10^3 \times 9,8 \times 128 = 1,25 \times 10^6.$$

(ii) Realizamos un esquema de diferencias divididas de Newton con todos los puntos de la tabulación dada de la función $L(x)$.

0	150				
		-65			
2	20		14.062		
		-8.75		-2.027	
4	2.5		1.9		0.2177
		-1		-0.285	
6	0.5		0.1875		
		-0.25			
8	0				

Entonces, el polinomio de interpolación es

$$\begin{aligned} P_4(x) &= 150 - 65(x-0) + 14,062(x-0)(x-2) \\ &\quad - 2,027(x-0)(x-2)(x-4) \\ &\quad + 0,2177(x-0)(x-2)(x-4)(x-6). \end{aligned}$$

Finalmente, $L(1) \approx P_4(1) = 61,5915$.

(iii) Recordemos que, al realizar la aproximación de una integral por el método de Simpson $I = \int_a^b f(x) dx \approx S_h(f)$, el error absoluto cometido es

$\Delta_I = h^4(b-a)/180|f^{(4)}(\xi)|$, siendo $\xi \in (a, b)$. En nuestro caso $a = 0$, $b = 8$ y $h = 0,1$. Entonces, una cota de Δ_I se halla de la forma

$$\Delta_I < \frac{(0,1)^4(8-0)}{180} M ,$$

siendo $M = \max_{0 \leq \xi \leq 8} \{|f^{(4)}(\xi)|\}$. Como $L(x) = (150-x)/e^x$, es decir, $f(x) = xL(x) = x(150-x)/e^x$ se obtiene su derivada cuarta

$$f^{(4)}(\xi) = \frac{-612 + 158\xi - \xi^2}{e^\xi} .$$

Para hallar M , calculamos en primer lugar los extremos relativos de la función $f^{(4)}(\xi)$ en el intervalo $(0, 8)$. Puesto que

$$f^{(5)}(\xi) = \frac{770 - 160\xi + \xi^2}{e^\xi} ,$$

las únicas soluciones de la ecuación $f^{(5)}(\xi) = 0$ son $80 \pm \sqrt{5630}$, es decir, $\xi = 4,96$ y $\xi = 155,03$. Entonces

$$\begin{aligned} M &= \max_{0 \leq \xi \leq 8} \{|f^{(4)}(\xi)|\} = \max\{|f^{(4)}(0)|, |f^{(4)}(4,96)|, |f^{(4)}(8)|\} \\ &= \max\{612, 1,03, 0,19\} = 612 . \end{aligned}$$

En definitiva

$$\Delta_I < \frac{(0,1)^4(8-0)}{180} 612 = 0,00272 .$$

Puesto que $F = \rho g I = 10^3 \times 9,8 \times I$, el error absoluto Δ_F cometido en la aproximación de F está acotado por

$$\Delta_F = 10^3 \times 9,8 \times \Delta_I < 10^3 \times 9,8 \times 0,00272 = 26,656 .$$

Problema 6.21 Aproximar el valor de la integral impropia

$$\int_1^\infty x^2 e^{-x^2} dx$$

mediante la fórmula de los trapecios compuesta utilizando 4 subintervalos.

INDICACIÓN 1: Utilizar el cambio de variable $x = 1/\sqrt{t}$ en la integral.

INDICACIÓN 2: El cálculo de $\lim_{z \rightarrow 0^+} z^{-5/2} e^{-1/z}$ se simplifica con el cambio $z = 1/k$.

Solución. Definamos la integral impropia de primera especie

$$I = \int_1^\infty x^2 e^{-x^2} dx .$$

Realizamos el cambio de variable $x = 1/\sqrt{t}$, con lo cual, diferenciando se obtiene que $dx = -\frac{1}{2}t^{-3/2}dt$. En la nueva variable t , la integral I se escribe como

$$I = \frac{1}{2} \int_0^1 f(t) dt , \quad f(t) = \frac{e^{-1/t}}{t^{5/2}} .$$

La integral I dada por la anterior expresión se ha convertido en una integral propia si y sólo si el límite siguiente $\lim_{t \rightarrow 0^+} f(t)$ es finito. Calculamos dicho límite realizando el cambio $t = 1/k$. Entonces, se tiene que

$$\lim_{t \rightarrow 0^+} f(t) = \lim_{t \rightarrow 0^+} \frac{e^{-1/t}}{t^{5/2}} = \lim_{k \rightarrow \infty} \frac{k^{5/2}}{e^k} = \frac{\infty}{\infty} = 0 ,$$

donde la indeterminación ∞/∞ se ha resuelto teniendo que el numerador es un infinito de orden inferior al del denominador. También se puede ver aplicando tres veces la regla de l'Hopital. En definitiva, la integral I se ha transformado en una integral propia y por lo tanto convergente. Ahora podemos utilizar para su aproximación la fórmula de los trapecios compuesta con 4 subintervalos:

$$\int_0^1 f(t) dt \approx h \left(\frac{f_0}{2} + f_1 + f_2 + f_3 + \frac{f_4}{2} \right) ,$$

siendo $h = 1/4$ y $f_i = f(ih)$ para $i = 0, 1, 2, 3, 4$. Se tiene pues que

$$\int_0^1 f(t) dt \approx \frac{1}{4} \left(\frac{0}{2} + 0,5861 + 0,765572 + 0,541112 + \frac{0,367879}{2} \right) = 0,519181 .$$

Se obtiene entonces la aproximación

$$I \approx \frac{1}{2} \times 0,519181 = 0,25959 .$$

Problema 6.22 Consideremos la siguiente integral impropia

$$I = \int_{-\infty}^0 \frac{e^x}{1 + e^{2x}} dx .$$

(i) Hallar el valor exacto de I .

INDICACIÓN: Usar el cambio de variable $z = e^x$.

(ii) Mediante el método de coeficientes indeterminados, hallar los valores de los pesos w_i para que la fórmula de integración numérica

$$\int_0^1 f(z) dz \approx w_0 f(0) + w_1 f(1/3) + w_2 f(1) ,$$

sea exacta para los polinomios de grado menor o igual que 2.

(iii) Aproxima el valor de I usando la expresión hallada en la fórmula del apartado anterior y calcula el error absoluto cometido.

Solución. (i) Diferenciando el cambio de variable $z = e^x$ se tiene $dz = e^x dx$. Entonces, la integral I escrita en la nueva variable z adopta la forma

$$I = \int_0^1 \frac{1}{1 + z^2} dz .$$

Observamos que I se ha convertido en una integral propia. Además, puesto que una primitiva de la función $\frac{1}{1+z^2}$ es $\arctan z$, aplicando la regla de Barrow se obtiene que

$$I = \int_0^1 \frac{1}{1+z^2} dz = \left[\arctan z \right]_0^1 = \arctan 1 = \frac{\pi}{4} .$$

(ii) Utilizando el método de coeficientes indeterminados, impondremos la igualdad

$$\int_0^1 f(z) dz = w_0 f(0) + w_1 f(1/3) + w_2 f(1) ,$$

para cualquier polinomio $f \in \mathbb{R}_2[z]$ de grado menor o igual que 2. Esto es equivalente a imponer la anterior igualdad para $f(z)$ igual a todos los monomios $\{1, z, z^2\}$ de la base canónica del espacio vectorial $\mathbb{R}_2[z]$. De este modo se obtiene

$$\begin{aligned} \int_0^1 dz &= 1 = w_0 + w_1 + w_2 , \\ \int_0^1 z dz &= \frac{1}{2} = w_0 \times 0 + w_1 \times \frac{1}{3} + w_2 \times 1 , \\ \int_0^1 z^2 dz &= \frac{1}{3} = w_0 \times 0^2 + w_1 \times \left(\frac{1}{3}\right)^2 + w_2 \times 1^2 . \end{aligned}$$

En definitiva, los pesos w_i son solución del sistema lineal

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1/3 & 1 \\ 0 & 1/9 & 1 \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1/2 \\ 1/3 \end{pmatrix} ,$$

es decir, $(w_0, w_1, w_2) = (0, 3/4, 1/4)$.

(iii) Para aproximar el valor de I usando la expresión hallada en la fórmula del apartado anterior, recordemos primero el resultado del apartado (i), es decir, la expresión

$$I = \int_0^1 f(z) dz , \quad f(z) = \frac{1}{1+z^2} .$$

Entonces, se tiene que

$$\int_0^1 f(z) dz \approx \frac{3}{4} f\left(\frac{1}{3}\right) + \frac{1}{4} f(1) = \frac{3}{4} \frac{1}{1 + (1/3)^2} + \frac{1}{4} \frac{1}{2} = \frac{4}{5} .$$

El error absoluto Δ cometido en esta aproximación es

$$\Delta = |I - 4/5| = |\pi/4 - 4/5| = 0,0146018 .$$

Problema 6.23 Sea $F : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ una función de clase $C^1([a, b])$. Al rotar la curva plana $y = F(x)$ alrededor del eje de abscisas desde $x = a$ hasta $x = b$ se obtiene una superficie de revolución en \mathbb{R}^3 cuya área A vale

$$A = \int_a^b |F(x)| \sqrt{1 + (F'(x))^2} dx .$$

Aproximar el área que genera la curva $y = \tan x$ al rotar alrededor del eje de abscisas desde $x = 0$ hasta $x = 1$ mediante una única extrapolación de Richardson según el método de Romberg.

Solución. La derivada de la función $F(x) = \tan x$ es $F'(x) = 1/\cos^2 x$, de modo que el área A que genera la curva $y = \tan x$ al rotar alrededor del eje de abscisas desde $x = 0$ hasta $x = 1$ será

$$A = \int_0^1 f(x) dx \quad \text{con} \quad f(x) = \tan x \sqrt{1 + \frac{1}{\cos^2 x}} .$$

Notar que se ha quitado el valor absoluto puesto que $\tan x \geq 0$ para todo $x \in [0, 1]$. Vamos a aproximar la integral A mediante una extrapolación de Richardson según el método de Romberg. Entonces, partimos de las aproximaciones proporcionadas por la fórmula de los Trapecios $T_h^{(1)}(f) = \frac{h}{2}[f(0) + f(1)]$ y $T_{h/2}^{(1)}(f) = h \left[\frac{f(0)}{2} + f\left(\frac{1}{2}\right) + \frac{f(1)}{2} \right]$ siendo $h = 1$. Se tiene pues que

$$\begin{aligned} T_1^{(1)}(f) &= \frac{1}{2} [f(0) + f(1)] = \frac{1}{2} [0 + 5,55761] = 2,77881 , \\ T_{1/2}^{(1)}(f) &= \frac{1}{2} \left[\frac{f(0)}{2} + f\left(\frac{1}{2}\right) + \frac{f(1)}{2} \right] = \frac{1}{2} \left[\frac{0}{2} + 0,89533 + \frac{5,55761}{2} \right] \\ &= 1,83707 . \end{aligned}$$

Finalmente, mediante la extrapolación de Richardson con los dos valores anteriores se obtiene

$$\begin{aligned} T_1^{(2)}(f) &= \frac{1}{3} [4T_{1/2}^{(1)}(f) - T_1^{(1)}(f)] = \frac{1}{3} [4 \times 1,83707 - 2,77881] \\ &= 1,52316 . \end{aligned}$$

Concluimos que la aproximación obtenida es

$$A \approx T_1^{(2)}(f) = 1,52316 .$$

Capítulo 7

Ecuaciones Diferenciales Ordinarias

7.1. Resumen de teoría

Muchos problemas importantes en ciencia y ingeniería son modelados mediante ecuaciones diferenciales. Sin embargo, la probabilidad de saber resolver una ecuación diferencial tomada al azar es prácticamente nula. Este hecho implica la necesidad de utilizar métodos numéricos para su solución. Por otra parte, aunque algunas ecuaciones diferenciales pueden resolverse por métodos analíticos, dicha solución es demasiado complicada y no es de mucha utilidad. Existen también ecuaciones diferenciales cuya solución no puede expresarse en términos de funciones elementales.

Definición 7.1 Sea $D \subset \mathbb{R}^2$ un dominio y $f : D \rightarrow \mathbb{R}$ una función continua. Decimos que la función $y : [a, b] \rightarrow \mathbb{R}$ de clase $C^1[a, b]$ es una solución de la ecuación diferencial $y' = f(x, y)$ si $(x, y(x)) \in D$ y además $y'(x) = f(x, y(x))$ para toda $x \in [a, b]$.

Teorema 7.2 Si f y $\partial f / \partial y$ son continuas en un rectángulo $R = \{(x, y) \in \mathbb{R}^2 : |x - x_0| \leq \alpha, |y - y_0| \leq \beta\}$ centrado en (x_0, y_0) , entonces el problema de Cauchy

$$y' = f(x, y), \quad y(x_0) = y_0,$$

tiene una única solución $y(x)$ definida para $|x - x_0| \leq \min\{\alpha, \beta/M\}$, siendo M el máximo de $|f(x, y)|$ en el rectángulo R .

Veamos algunos de los métodos numéricos para la aproximación de la solución del problema de Cauchy para una ecuación diferencial de primer orden

$$y' = f(x, y), \quad y(x_0) = y_0. \quad (7.1)$$

Resolver numéricamente el problema de Cauchy (7.1) no significa obtener una fórmula explícita para la dependencia $y(x)$ de la solución. Se entiende por solución numérica una tabulación del tipo

x_0	x_1	x_2	\cdots	x_n
y_0	y_1	y_2	\cdots	y_n

siendo y_i una aproximación numérica del valor exacto $y(x_i)$. Además, si no decimos lo contrario, siempre consideraremos las abscisas equiespaciadas $x_i = x_0 + ih$ para $i = 0, 1, \dots, n$, con longitud de paso h .

- Euler: $y_{i+1} = y_i + hf(x_i, y_i)$.
- Heun: $y_{i+1}^* = y_i + hf(x_i, y_i)$, $y_{i+1} = y_i + \frac{h}{2} \left[f(x_i, y_i) + f(x_{i+1}, y_{i+1}^*) \right]$.
- Taylor de 2° orden: $y_{i+1} = y_i + hf(x_i, y_i) + \frac{h^2}{2} \left[\frac{\partial f}{\partial x}(x_i, y_i) + \frac{\partial f}{\partial y}(x_i, y_i) f(x_i, y_i) \right]$.
- RK4: $y_{i+1} = y_i + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$, con $k_1 = f(x_i, y_i)$, $k_2 = f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}k_1\right)$, $k_3 = f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}k_2\right)$, $k_4 = f(x_i + h, y_i + hk_3)$.

Los métodos numéricos presentados pueden ser generalizados utilizando notación vectorial a ecuaciones diferenciales ordinarias definidas en \mathbb{R}^n . De esta forma, consideremos el problema de Cauchy

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}), \quad \mathbf{x}(t_0) = \mathbf{x}_0. \quad (7.2)$$

Aquí, $t \in \mathbb{R}$ es la variable independiente y $\mathbf{x} = (x_1, \dots, x_n)^t \in \mathbb{R}^n$ es la variable vectorial dependiente y $\mathbf{f}(t, \mathbf{x}) = (f_1(t, \mathbf{x}), \dots, f_n(t, \mathbf{x}))^t \in \mathbb{R}^n$. Se tienen en definitiva las funciones incógnita $x_i(t)$, con $i = 1, \dots, n$. La condición inicial es el valor vectorial $\mathbf{x}_0 = (x_1^0, \dots, x_n^0)^t \in \mathbb{R}^n$.

- La iteración de Euler con longitud de paso h consiste en: $y_{i+1} = y_i + hf(x_i, y_i)$.

$$\mathbf{x}_{i+1} = \mathbf{x}_i + h\mathbf{f}(t_i, \mathbf{x}_i). \quad (7.3)$$

- La iteración RK4 con longitud de paso h consiste en

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \frac{h}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4), \quad (7.4)$$

siendo

$$\begin{aligned} \mathbf{k}_1 &= \mathbf{f}(t_i, \mathbf{x}_i), \\ \mathbf{k}_2 &= \mathbf{f}\left(t_i + \frac{h}{2}, \mathbf{x}_i + \frac{h}{2}\mathbf{k}_1\right), \\ \mathbf{k}_3 &= \mathbf{f}\left(t_i + \frac{h}{2}, \mathbf{x}_i + \frac{h}{2}\mathbf{k}_2\right), \\ \mathbf{k}_4 &= \mathbf{f}(t_i + h, \mathbf{x}_i + h\mathbf{k}_3). \end{aligned}$$

Problema Lineal de Contorno: Si $p(x), q(x), r(x) \in C[a, b]$ con $q(x) > 0$ en $[a, b]$, entonces el problema lineal de frontera $y''(x) = p(x)y' + q(x)y + r(x)$ con $y(a) = \alpha$, $y(b) = \beta$ tiene una única solución $y(x)$ con $x \in [a, b]$. Además dicha solución viene dada por $y(x) = u(x) + Cv(x)$, siendo $C = (\beta - u(b))/v(b)$, donde $u(x)$ y $v(x)$ son la única solución de los problemas de Cauchy (1) $u''(x) = p(x)u' + q(x)u + r(x)$ con $u(a) = \alpha$, $u'(a) = 0$; (2) $v''(x) = p(x)v' + q(x)v$ con $v(a) = 0$, $v'(a) = 1$.

7.2. Problemas resueltos

Problema 7.1 Se define la función $y(x)$ de la forma

$$y(x) = \int_0^x \sqrt{1+t^3} dt .$$

Obtener una tabulación de la función $y(x)$ en el intervalo $0 \leq x \leq 2$ utilizando el método de Taylor de segundo orden con paso $h = 1/2$ para resolver de forma numérica un problema de valor inicial adecuado.

NOTA: La integral $\int \sqrt{1+t^3} dt$ es elíptica y por lo tanto no es posible obtener una primitiva expresable en términos de funciones elementales.

Solución. Aplicando el Teorema fundamental del cálculo obtenemos $y'(x)$, es decir, la derivada de la función $y(x)$. Además, por definición es claro que $y(0) = 0$. De esta forma se tiene que la función $y(x)$ es la única solución del problema de valor inicial

$$y' = \sqrt{1+x^3}, \quad y(0) = 0 .$$

Definiendo la función $f(x, y) = \sqrt{1+x^3}$, el método de Taylor de segundo orden para aproximar la solución del problema de valor inicial viene dado por

$$y_{i+1} = y_i + hf(x_i, y_i) + \frac{h^2}{2} \left[\frac{\partial f(x_i, y_i)}{\partial x} + \frac{\partial f(x_i, y_i)}{\partial y} f(x_i, y_i) \right] .$$

En nuestro problema, las derivadas parciales de primer orden de la función f son

$$\frac{\partial f(x, y)}{\partial x} = \frac{3x^2}{2\sqrt{1+x^3}}, \quad \frac{\partial f(x, y)}{\partial y} = 0 ,$$

de manera que el método de Taylor adopta la forma

$$y_{i+1} = y_i + \frac{1}{2} \left[\sqrt{1+x_i^3} + \frac{3x_i^2}{8\sqrt{1+x_i^3}} \right] .$$

Realizando esta iteración con el valor inicial $(x_0, y_0) = (0, 0)$ obtenemos la tabla siguiente

i	x_i	y_i
0	0	0
1	0.5	0.5
2	1	1.07452
3	1.5	1.91421
4	2	3.16173

donde la tercera columna es una tabulación aproximada de la función $y(x)$.

Problema 7.2 La función de error $\operatorname{erf}(x)$, utilizada en estadística y en probabilidad, está definida de la forma

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt .$$

- (i) Aproximar el valor de $\operatorname{erf}(2)$ mediante la fórmula de Simpson.
- (ii) Aproximar el valor de $\operatorname{erf}(2)$ mediante la transformación en un problema de valor inicial y la posterior utilización del método de Euler con una longitud de paso $h = 1/2$.

Solución. (i) Utilizando la fórmula de Simpson

$$\int_a^b f(t) dt \approx \frac{h}{3} (f_0 + 4f_1 + f_2) , \text{ siendo } h = (b - a)/2 ,$$

se tiene que

$$\operatorname{erf}(2) = \frac{2}{\sqrt{\pi}} \int_0^2 \exp(-t^2) dt \approx \frac{2}{\sqrt{\pi}} \frac{1}{3} \left(1 + \frac{4}{e} + \frac{1}{e^4} \right) \approx 0,936492 .$$

(ii) Aplicando el Teorema fundamental del cálculo, la derivada de la función error es

$$\frac{d \operatorname{erf}(x)}{dx} = \frac{2}{\sqrt{\pi}} \exp(-x^2) .$$

Además, de la definición de función error se desprende que $\operatorname{erf}(0) = 0$. En definitiva, la función $y = \operatorname{erf}(x)$ es la solución del problema de Cauchy

$$y' = f(x, y) = \frac{2}{\sqrt{\pi}} \exp(-x^2) , \quad y(0) = 0 .$$

Aplicando el método de Euler

$$y_{i+1} = y_i + hf(x_i, y_i) ,$$

con paso $h = 1/2$ para aproximar la solución del anterior problema de Cauchy se tiene la siguiente iteración

$$y_{i+1} = y_i + \frac{1}{\sqrt{\pi}} \exp(-x_i^2) , \quad \text{con } y_0 = 0 ,$$

que da lugar a la tabla siguiente.

i	x_i	y_i
0	0	0
1	0.5	0.56419
2	1	1.00358
3	1.5	1.21113
4	2	1.2706

Finalmente, se concluye con la aproximación $\text{erf}(2) \approx y_4 = 1,2706$.

Problema 7.3 Considerar el problema de Cauchy $y' = f(x, y)$, $y(x_0) = y_0$. Se pretende aproximar la solución $y(x)$ en el intervalo $[x_0, x_n]$ mediante el método de Runge-Kutta

$$y_{i+1} = y_i + h \left[\frac{2}{9}k_1 + \frac{1}{3}k_2 + \frac{4}{9}k_3 \right],$$

donde $k_1 = f(x_i, y_i)$, $k_2 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2}k_1)$, $k_3 = f(x_i + \frac{3h}{4}, y_i + \frac{3h}{4}k_2)$.

- (i) Demostrar que dicho método es consistente.
- (ii) Demostrar que si $f \in C^2[x_0, x_n]$ entonces el método tiene al menos orden de consistencia 2.

Solución. (i) El método numérico es de un paso $y_{i+1} = y_i + h\phi(x_i, y_i; h)$. Por lo tanto será consistente si y sólo si

$$\lim_{h \rightarrow 0} \phi(x_i, y_i; h) = f(x_i, y_i). \quad (7.5)$$

En este método RK se tiene la función

$$\phi(x_i, y_i; h) = \frac{2}{9}k_1 + \frac{1}{3}k_2 + \frac{4}{9}k_3.$$

Además, de la definición de k_j se verifica que $\lim_{h \rightarrow 0} k_j = f(x_i, y_i)$ para $j = 1, 2, 3$. Se tiene por lo tanto que

$$\lim_{h \rightarrow 0} \phi(x_i, y_i; h) = \frac{2}{9}f(x_i, y_i) + \frac{1}{3}f(x_i, y_i) + \frac{4}{9}f(x_i, y_i) = f(x_i, y_i),$$

de manera que el método es consistente.

- (ii) El error de discretización local $\Delta(x_i, y_i; h)$ viene dado por

$$\begin{aligned} \Delta(x_i, y_i; h) &= f(x_i, y_i) + \frac{y''(x_i)}{2!}h + O(h^2) - \phi(x_i, y_i; h) \\ &= k_1 + \frac{y''(x_i)}{2!}h + O(h^2) - \left[\frac{2}{9}k_1 + \frac{1}{3}k_2 + \frac{4}{9}k_3 \right] \\ &= \frac{7}{9}k_1 + \frac{y''(x_i)}{2!}h + O(h^2) - \left[\frac{1}{3}k_2 + \frac{4}{9}k_3 \right]. \end{aligned}$$

Como $y(x)$ es solución exacta de la ecuación diferencial $y' = f(x, y)$, aplicando la regla de la cadena se tiene

$$y''(x_i) = \frac{df}{dx}(x_i, y_i) = \frac{\partial f(x_i, y_i)}{\partial x} + \frac{\partial f(x_i, y_i)}{\partial y} f(x_i, y_i) .$$

Por otra parte, como $f \in C^2[x_0, x_n]$, podemos realizar los siguientes desarrollos de Taylor

$$\begin{aligned} k_2 &= f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}k_1\right) \\ &= f(x_i, y_i) + \frac{\partial f(x_i, y_i)}{\partial x} \frac{h}{2} + \frac{\partial f(x_i, y_i)}{\partial y} \frac{h}{2}k_1 + O(h^2) , \\ k_3 &= f\left(x_i + \frac{3h}{4}, y_i + \frac{3h}{4}k_2\right) \\ &= f(x_i, y_i) + \frac{\partial f(x_i, y_i)}{\partial x} \frac{3h}{4} + \frac{\partial f(x_i, y_i)}{\partial y} \frac{3h}{4}k_2 + O(h^2) \\ &= f(x_i, y_i) + \frac{3}{4} \left[\frac{\partial f(x_i, y_i)}{\partial x} + f(x_i, y_i) \frac{\partial f(x_i, y_i)}{\partial y} \right] h + O(h^2) . \end{aligned}$$

Introduciendo estos resultados y agrupando por potencias de h , el error de discretización local adopta la forma

$$\begin{aligned} \Delta(x_i, y_i; h) &= \left[\frac{7}{9} - \frac{1}{3} - \frac{4}{9} \right] f(x_i, y_i) + \left[\frac{1}{2} - \frac{1}{6} - \frac{3}{9} \right] \left[\frac{\partial f(x_i, y_i)}{\partial x} \right. \\ &\quad \left. + f(x_i, y_i) \frac{\partial f(x_i, y_i)}{\partial y} \right] h + O(h^2) = O(h^2) , \end{aligned}$$

de manera que el orden de consistencia es al menos 2.

Problema 7.4 La variación de la velocidad v con el tiempo t de una partícula verifica la ecuación

$$\frac{dv}{dt} = \frac{5000 - 0,1v^2}{300 - 10t} - g ,$$

siendo $g = 9,8 \text{ms}^{-2}$ la aceleración de la gravedad. Sabiendo que inicialmente la partícula está en reposo, aproximar la velocidad al cabo de 1 segundo de haber iniciado el movimiento $v(1)$ mediante el método de Euler con incrementos de tiempo de 0,5 segundos.

Solución. Dada la ecuación diferencial $\dot{v} = f(t, v)$ con la condición inicial $v(t_0) = v_0$, el método de Euler aproxima su solución exacta $v(t)$ mediante una tabulación (t_i, v_i) para $i = 0, 1, \dots, n$. Los tiempos son equiespaciados $t_i = t_0 + ih$ siendo h el paso constante y los valores de v_i son generados a partir de la recurrencia $v_{i+1} = v_i + hf(t_i, v_i)$.

En el problema planteado se tiene $h = 0,5$ y como la partícula está inicialmente en reposo $t_0 = 0$ y $v_0 = 0$. Además

$$f(t, v) = \frac{5000 - 0,1v^2}{300 - 10t} - 9,8 .$$

De esta forma, la iteración que se ha de realizar es

$$v_{i+1} = v_i + hf(t_i, v_i) = v_i + 0,5 \left[\frac{5000 - 0,1v_i^2}{300 - 10t_i} - 9,8 \right] ,$$

con valor inicial $(t_0, v_0) = (0, 0)$. Se obtiene de esta forma la tabla siguiente

i	t_i	v_i
0	0	0
1	0.5	3.43333
2	1	7.00591

Finalmente la aproximación pedida es $v(1) \approx v_2 = 7,00591m/s$.

Problema 7.5 Consideremos el problema de valor inicial $y' = f(x, y)$ con $y(x_0) = y_0$. Calcular las constantes A y B de manera que se obtenga el siguiente método multipaso implícito de Adams-Moulton

$$y_{i+1} = y_i + h \left[Af(x_i, y_i) + Bf(x_{i+1}, y_{i+1}) \right] ,$$

siendo h la longitud de paso, es decir, $x_{i+1} = x_i + h$ para $i = 0, 1, \dots$

Solución. Sea $y(x)$ la única solución del problema de valor inicial $y' = f(x, y)$ con $y(x_0) = y_0$. Como la fórmula de Adams-Moulton que hemos de demostrar es de la forma $y_{i+1} = y_i + h(TERMINO)$, entonces separamos variables e integramos la ecuación diferencial de la forma

$$\int_{y(x_i)}^{y(x_{i+1})} dy = \int_{x_i}^{x_{i+1}} f(x, y(x)) dx ,$$

es decir

$$y(x_{i+1}) = y(x_i) + \int_{x_i}^{x_{i+1}} f(x, y(x)) dx .$$

Como $TERMINO$ depende de las abscisas x_i, x_{i+1} , la integral anterior será aproximada por la integral de un polinomio $P(x)$ que interpole a la función $f(x, y(x))$ en dichas abscisas. Pero ahora sucede un comportamiento inusual en las fórmulas de Adams-Moulton que consiste en observar que la aproximación numérica de la integral que pretendemos efectuar no es más que una regla de Newton-Cotes (en concreto es la regla de los Trapecios). Aprovechando esta información obtenemos

$$y_{i+1} = y_i + \int_{x_i}^{x_{i+1}} P(x) dx = y_i + \frac{h}{2} [f(x_i, y_i) + f(x_{i+1}, y_{i+1})] ,$$

de lo que se deduce $A = B = 1/2$.

Problema 7.6 *Demostrar que el método de Heun aplicado a la resolución del problema de valor inicial $y' = f(x)$ con $y(x_0) = y_0$ es equivalente a la aplicación de la regla de los trapecios al reformular el problema de Cauchy en forma integral.*

Solución. El método de Heun para la resolución del problema de valor inicial $y' = f(x, y)$ con $y(x_0) = y_0$ se basa en el esquema predictor-corrector

$$y_{i+1}^* = y_i + hf(x_i, y_i) \ , \quad y_{i+1} = y_i + \frac{h}{2}[f(x_i, y_i) + f(x_{i+1}, y_{i+1}^*)] \ .$$

En el problema que abordamos $f(x, y) = f(x)$, de manera que el método de Heun se simplifica y queda

$$y_{i+1} = y_i + \frac{h}{2}[f(x_i) + f(x_{i+1})] \ .$$

Por otra parte la ecuación diferencial $y' = f(x)$ es de variables separables, de manera que se tiene

$$\int_{y(x)}^{y(x+h)} dy = \int_x^{x+h} f(x) dx \ .$$

Resolviendo la primera integral y despejando se tiene la formulación integral del problema de Cauchy

$$y(x+h) = y(x) + \int_x^{x+h} f(x) dx \ .$$

Tomando en la anterior igualdad $x = x_i$ y recordando que las abscisas x_i son equiespaciadas con paso h , es decir $x_{i+1} = x_i + h$, se tiene

$$y(x_{i+1}) = y(x_i) + \int_{x_i}^{x_i+h} f(x) dx \ .$$

Finalmente, utilizando la regla de los trapecios para aproximar la anterior integral y denotando $y(x_i) \approx y_i$, concluimos que

$$y_{i+1} = y_i + \frac{h}{2}[f(x_i) + f(x_{i+1})] \ ,$$

es decir, obtenemos el mismo tipo de aproximación que con el método de Heun.

Problema 7.7 *Una masa m oscila entorno de su punto de equilibrio debido a la acción de un muelle de constante elástica k . La masa está sumergida en un medio viscoso que le provoca una fuerza de rozamiento proporcional a su velocidad al cubo. En estas condiciones, la ecuación diferencial que rige el movimiento de la partícula es*

$$m\ddot{x} + c\dot{x}^3 + kx = 0 \ .$$

Teniendo en cuenta que la partícula se encuentra inicialmente en $x(0) = 1/2$ y que tiene una velocidad inicial $\dot{x}(0) = 1/4$, utilizar el método de Euler con una longitud de paso $h = 1$ para evidenciar de forma numérica que dicha partícula realiza un movimiento periódico y aproximar el valor del periodo T . DATOS: $m = c = k = 1$.

Solución. Teniendo en cuenta los datos del problema, la ecuación diferencial del movimiento del oscilador viene dada por la siguiente ecuación diferencial no lineal de segundo orden $\ddot{x} + \dot{x}^3 + x = 0$. Esta ecuación puede ser reescrita mediante el cambio de variables $\dot{x} = y$ en un sistema de ecuaciones diferenciales de primer orden. Además, teniendo en cuenta las condiciones iniciales del movimiento del oscilador se obtiene el siguiente problema de Cauchy

$$\dot{x} = y, \quad \dot{y} = -x - y^3, \quad (x(0), y(0)) = (1/2, 1/4).$$

Utilizaremos el método de Euler con una longitud de paso $h = 1$ para aproximar la solución de este problema de valor inicial. En concreto se tiene $(x_{i+1}, y_{i+1}) = (x_i, y_i) + hf(x_i, y_i)$, donde $h = 1$ y $f(x, y) = (y, -x - y^3)$. En resumen, se ha de realizar la iteración

$$\begin{pmatrix} x_{i+1} \\ y_{i+1} \end{pmatrix} = \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} y_i \\ -x_i - y_i^3 \end{pmatrix}, \quad \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix}.$$

Con la notación $x(t_i) \approx x_i$, $y(t_i) \approx y_i$, $t_i = t_0 + ih = i$, la anterior iteración de Euler da lugar a la tabulación

i	x_i	y_i	t_i
0	1/2	1/4	0
1	0.75	-0.2656	1
2	0.484	-0.9968	2
3	-0.512	-0.4905	3
4	-1.003	0.1399	4
5	-0.863	1.1403	5
6	0.2772	0.5205	6

Hemos detenido la iteración puesto que ya se observa la ejecución de un ciclo completo en el plano de las fases xy . La aproximación del periodo T dada por el método es $T \approx 6$.

Problema 7.8 *La solución exacta del problema de Cauchy*

$$\begin{cases} \dot{x} &= -x + (t+1)y \\ \dot{y} &= x - ty \end{cases} \quad (7.6)$$

con condiciones iniciales

$$x(0) = \sqrt{\frac{2}{e\pi}} = 0.483941, \quad y(0) = \operatorname{erf}\left(\frac{1}{\sqrt{2}}\right) = 0.682689, \quad (7.7)$$

viene dada por

$$x(t) = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{(t+1)^2}{2}\right) + t \operatorname{erf}\left(\frac{t+1}{\sqrt{2}}\right), \quad y(t) = \operatorname{erf}\left(\frac{t+1}{\sqrt{2}}\right), \quad (7.8)$$

donde la función error $\operatorname{erf}(z)$ está definida por

$$\operatorname{erf}(z) := \frac{2}{\sqrt{\pi}} \int_0^z \exp(-s^2) ds. \quad (7.9)$$

- (i) Aproximar a partir de (7.9) el valor de $\operatorname{erf}(2)$ utilizando una vez la extrapolación de Richardson según el método de Romberg.
- (ii) Utilizar un paso del método clásico de Runge-Kutta de cuarto orden para aproximar el valor de $x(t)$ e $y(t)$ solución de (7.6) y (7.7) para $t = 2\sqrt{2} - 1 = 1.82843$.
- (iii) ¿Es posible comparar los resultados obtenidos en los anteriores apartados?

Solución. (i) Definamos la función integrando $f(s) = \exp(-s^2)$. Sea $T_h^{(1)}(f)$ la regla de los trapecios aplicada a la función f con longitud de paso h . Entonces, aplicando la regla de los trapecios simple se tiene

$$\begin{aligned} \operatorname{erf}(2) &= \frac{2}{\sqrt{\pi}} \int_0^2 f(s) ds \approx \frac{2}{\sqrt{\pi}} T_2^{(1)}(f) \\ &= \frac{2}{\sqrt{\pi}} \frac{2}{2} [f(0) + f(2)] = \frac{2}{\sqrt{\pi}} [1 + 0.0183156] = 1.14905. \end{aligned}$$

Utilizando de nuevo la regla de los trapecios, pero ahora con longitud de paso la mitad que la anterior, se obtiene

$$\begin{aligned} \operatorname{erf}(2) &= \frac{2}{\sqrt{\pi}} \int_0^2 f(s) ds \approx \frac{2}{\sqrt{\pi}} T_1^{(1)}(f) \\ &= \frac{2}{\sqrt{\pi}} \left[\frac{f(0)}{2} + f(1) + \frac{f(2)}{2} \right] \\ &= \frac{2}{\sqrt{\pi}} \left[\frac{1}{2} + 0.367879 + \frac{0.0183156}{2} \right] = 0.989631 \end{aligned}$$

A partir de las dos estimaciones de $\text{erf}(2)$ realizadas se obtiene una mejor aproximación por extrapolación de la forma

$$\text{erf}(2) \approx \frac{2}{\sqrt{\pi}} T_2^{(2)}(f) = \frac{4 \times 0.989631 - 1.14905}{4 - 1} = 0.936491 .$$

(ii) El método RK4 con una longitud de paso h aplicado a la solución numérica del problema de Cauchy

$$\begin{cases} \dot{x} &= f(t, x, y) \\ \dot{y} &= g(t, x, y) , \end{cases}$$

con la condición inicial $(x(0), y(0)) = (x_0, y_0)$ consiste en realizar la iteración

$$\begin{pmatrix} x_{i+1} \\ y_{i+1} \end{pmatrix} = \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \frac{h}{6} \begin{pmatrix} r_1 + 2r_2 + 2r_3 + r_4 \\ k_1 + 2k_2 + 2k_3 + k_4 \end{pmatrix} ,$$

donde

$$\begin{aligned} r_1 &= f(t_i, x_i, y_i) , \\ k_1 &= g(t_i, x_i, y_i) , \\ r_2 &= f\left(t_i + \frac{h}{2}, x_i + \frac{h}{2}r_1, y_i + \frac{h}{2}k_1\right) , \\ k_2 &= g\left(t_i + \frac{h}{2}, x_i + \frac{h}{2}r_1, y_i + \frac{h}{2}k_1\right) , \\ r_3 &= f\left(t_i + \frac{h}{2}, x_i + \frac{h}{2}r_2, y_i + \frac{h}{2}k_2\right) , \\ k_3 &= g\left(t_i + \frac{h}{2}, x_i + \frac{h}{2}r_2, y_i + \frac{h}{2}k_2\right) , \\ r_4 &= f(t_i + h, x_i + hr_3, y_i + hk_3) , \\ k_4 &= g(t_i + h, x_i + hr_3, y_i + hk_3) . \end{aligned}$$

En nuestro caso $f(t, x, y) = -x + (t + 1)y$, $g(t, x, y) = x - ty$. Además, puesto que se ha de aproximar el valor de $x(2\sqrt{2} - 1)$ e $y(2\sqrt{2} - 1)$ con una sola iteración y las condiciones iniciales son dadas para tiempo $t_0 = 0$, es obvio que $h = 2\sqrt{2} - 1 = 1.82843$. En definitiva, sustituyendo los datos se tiene que

$$\begin{aligned} r_1 &= f(t_0, x_0, y_0) = f(0, 0.483941, 0.682689) = 0.198748 , \\ k_1 &= g(t_0, x_0, y_0) = g(0, 0.483941, 0.682689) = 0.483941 , \end{aligned}$$

de modo que

$$\begin{aligned} r_2 &= f\left(\frac{1.82843}{2}, 0.483941 + \frac{1.82843}{2}0.198748, 0.682689 \right. \\ &\quad \left. + \frac{1.82843}{2}0.483941\right) = 1.48807 , \end{aligned}$$

$$k_2 = g\left(\frac{1.82843}{2}, 0.483941 + \frac{1.82843}{2} \cdot 0.198748, 0.682689 + \frac{1.82843}{2} \cdot 0.483941\right) = -0.362958,$$

y en consecuencia

$$r_3 = f\left(\frac{1.82843}{2}, 0.483941 + \frac{1.82843}{2} \cdot 1.48807, 0.682689 - \frac{1.82843}{2} \cdot 0.362958\right) = -1.17272,$$

$$k_3 = g\left(\frac{1.82843}{2}, 0.483941 + \frac{1.82843}{2} \cdot 1.48807, 0.682689 - \frac{1.82843}{2} \cdot 0.362958\right) = 1.52359,$$

de manera que

$$r_4 = f(1.82843, 0.483941 - 1.82843 \times 1.17272, 0.682689 + 1.82843 \times 1.52359) = 11.4706,$$

$$k_4 = g(1.82843, 0.483941 - 1.82843 \times 1.17272, 0.682689 + 1.82843 \times 1.52359) = -8.00215.$$

En definitiva se obtiene, según la recurrencia del método RK4

$$x(2\sqrt{2} - 1) \approx x_1 = 4.23224, \quad y(2\sqrt{2} - 1) \approx y_1 = -0.90102.$$

(iii) Notar que, a partir de la solución exacta (7.8), se tiene que

$$y(2\sqrt{2} - 1) = \operatorname{erf}(2),$$

de modo que es posible realizar una comparación de los resultados obtenidos en los apartados (i) y (ii).

Problema 7.9 Considerar el problema de valor inicial de primer orden $y' = x^2 + y$, $y(0) = 1$.

- (i) Averiguar si el problema de Cauchy tiene solución única.
- (ii) Utilizar el método de Taylor de segundo orden con una longitud de paso $h = 0,1$ para aproximar el valor de $y(0,3)$ y calcular el error absoluto cometido sabiendo que la solución exacta es $y(x) = 3e^x - x^2 - 2x - 2$.

Solución. (i) Según un corolario del Teorema de Picard-Lindelöf, si f y $\partial f/\partial y$ son continuas en un entorno del punto (x_0, y_0) , entonces el problema de Cauchy $y' = f(x, y)$ con $y(x_0) = y_0$ tiene una única solución $y(x)$ definida en un entorno de x_0 . Puesto que en nuestro problema $f(x, y) = x^2 + y$ es un polinomio en x e

y está claro que se verifica que tanto f como $\partial f / \partial y$ son continuas en un entorno de cualquier punto $(x_0, y_0) \in \mathbb{R}^2$ y en particular en un entorno del punto $(0, 1)$. Se concluye pues que el problema de Cauchy $y' = x^2 + y$, $y(0) = 1$ tiene solución única.

(ii) Para aplicar el método de Taylor de segundo orden se ha de realizar un desarrollo de Taylor de la solución hasta segundo orden, es decir,

$$y(x+h) = y(x) + hy'(x) + \frac{h^2}{2}y''(x) + O(h^3) .$$

Calculando las derivadas

$$y'(x) = f(x, y(x)) , \quad y''(x) = \frac{\partial f}{\partial x}(x, y(x)) + \frac{\partial f}{\partial y}(x, y(x))f(x, y(x)) ,$$

se concluye que el método de Taylor para el problema de Cauchy del enunciado consiste en la aplicación de la iteración siguiente

$$y_{i+1} = y_i + hf(x_i, y_i) + \frac{h^2}{2} \left[\frac{\partial f}{\partial x}(x_i, y_i) + \frac{\partial f}{\partial y}(x_i, y_i)f(x_i, y_i) \right] , \quad i = 0, 1, \dots$$

Como en nuestro problema se tiene $f(x, y) = x^2 + y$, las derivadas parciales de primer orden de la función f son

$$\frac{\partial f}{\partial x}(x, y) = 2x , \quad \frac{\partial f}{\partial y}(x, y) = 1 ,$$

y el método de Taylor toma la forma

$$y_{i+1} = y_i + h(x_i^2 + y_i) + \frac{h^2}{2} [2x_i + x_i^2 + y_i] , \quad i = 0, 1, \dots$$

La aplicación de esta iteración permite realizar la tabla siguiente

n	x_n	y_n
0	0.0	1.0
1	0.1	1.105
2	0.2	1.22307
3	0.3	1.35769

de manera que se obtiene la estimación $y(0,3) \approx y_3 = 1,35769$. Finalmente, el valor exacto es

$$y(0,3) = 3e^{0,3} - (0,3)^2 - 2(0,3) - 2 = 1,35958 ,$$

de manera que el error absoluto cometido es

$$\Delta = |y(0,3) - y_3| = |1,35769 - 1,35958| = 0,00189 .$$

Problema 7.10 Consideremos la ecuación diferencial $dy/dx = y \cos x$ con la condición inicial $y(0) = 1$.

- (i) Demostrar que este problema de Cauchy tiene una única solución $y(x)$.
- (ii) Aproximar el valor de $y(1)$ mediante el método de Heun con una longitud de paso $h = 1/2$ y calcular, mediante el cálculo de la solución exacta, el error absoluto cometido.

Solución. (i) El problema de Cauchy se puede escribir como $y' = f(x, y)$ con $y(0) = 1$ si definimos la función $f(x, y) = y \cos x$. Puesto que tanto f como $\partial f / \partial y$ son funciones de clase $C^\infty(\mathbb{R}^2)$, en particular son ambas funciones continuas en un entorno del punto $(0, 1) \in \mathbb{R}^2$. Por lo tanto el problema de valor inicial admite una única solución $y(x)$.

(ii) Utilizando el método predictor-corrector de Heun (también llamado de Euler mejorado), hallaremos una aproximación de la función $y(x)$ solución del problema de Cauchy para los valores $x_i = x_0 + ih$, es decir, $x_i = i/2$ con $i = 0, 1, 2$. Dicha tabulación (x_i, y_i) para $i = 0, 1, 2$, donde $y(x_i) \approx y_i$ se obtiene a partir de los tres primeros términos de la sucesión generada de forma recurrente mediante la iteración

$$\begin{aligned} y_{i+1}^* &= y_i + hf(x_i, y_i) , \\ y_{i+1} &= y_i + \frac{h}{2} [f(x_i, y_i) + f(x_{i+1}, y_{i+1}^*)] . \end{aligned}$$

En concreto se tiene la iteración

$$\begin{aligned} y_{i+1}^* &= y_i + \frac{1}{2} y_i \cos x_i , \quad i = 0, 1, 2. \\ y_{i+1} &= y_i + \frac{1}{4} [y_i \cos x_i + y_{i+1}^* \cos x_{i+1}] , \end{aligned}$$

que da lugar a la tabulación

i	x_i	y_i
0	0	1
1	1/2	1.57909
2	1	2.23242

Por lo tanto, la aproximación efectuada es $y(1) \approx y_2 = 2,23242$.

Por otra parte, es fácil resolver de manera exacta el problema de Cauchy planteado. Esto es debido a que la ecuación diferencial es de variables separables. Separando las variables e integrando se tiene

$$\int_1^y \frac{dy}{y} = \int_0^x \cos x \, dx ,$$

donde se han incorporado las condiciones iniciales del problema de Cauchy en los límites inferiores de las integrales. En definitiva, la solución general del problema de Cauchy viene dada por $y(x) = e^{\sin x}$.

Finalmente, el error absoluto Δ cometido en la aproximación $y(1) \approx y_2$ es

$$\Delta = |y(1) - y_2| = |e^{\sin 1} - 2,23242| = |2,31978 - 2,23242| = 0,08736 .$$

Problema 7.11 *Obtener el error absoluto que comete el método de Euler cuando aproxima $y(1/2)$ mediante 2 iteraciones siendo $y(x)$ la solución del problema de Cauchy $y' = (x-1)y^2$ con la condición inicial $y(0) = 2$.*

Solución. El método de Euler para aproximar la solución del problema de Cauchy $y' = f(x, y)$, $y(x_0) = y_0$ consiste en realizar la iteración $y_{i+1} = y_i + hf(x_i, y_i)$, donde $y(x_i) \approx y_i$ y la variable independiente toma valores equiespaciados $x_i = x_0 + ih$.

En nuestro caso se tiene $(x_0, y_0) = (0, 2)$. Como además se quiere aproximar $y(1/2)$ mediante 2 iteraciones, es decir, $y(1/2) \approx y_2$, se concluye que la condición $x_2 = x_0 + 2h$ debe ser $1/2 = 0 + 2h$ de modo que la longitud de paso vale $h = 1/4$. De este modo, la iteración de Euler que se debe realizar es

$$y_{i+1} = y_i + hf(x_i, y_i) = y_i + \frac{1}{4}(x_i - 1)y_i^2 .$$

La tabla que se obtiene es la siguiente.

i	x_i	y_i
0	0	2
1	1/4	1
2	1/2	0.8125

Se tiene pues que el método de Euler con 2 iteraciones realiza la aproximación $y(1/2) \approx y_2 = 0,8125$. Por definición, el error absoluto Δ cometido en esta aproximación es $\Delta = |y(1/2) - 0,8125|$, de modo que se necesita calcular el valor exacto de $y(1/2)$. Esto es fácil en este caso puesto que la ecuación diferencial es de variables separables y se puede hallar la solución exacta del problema de Cauchy $y' = (x-1)y^2$ con $y(0) = 2$. En concreto, separando variables se obtiene

$$\int_2^{y(x)} \frac{dy}{y^2} = \int_0^x (x-1) dx .$$

Realizando las integrales y despejando $y(x)$ se obtiene

$$y(x) = \frac{2}{-x^2 + 2x + 1} ,$$

con lo cual $y(1/2) = 8/7 = 1,1428$. Se tiene finalmente

$$\Delta = |y(1/2) - y_2| = |1,1428 - 0,8125| = 0,3303 .$$

Problema 7.12 Dos masas m_1 y m_2 están conectadas mediante un hilo inextensible que pasa por dos poleas fijas, pequeñas y sin fricción colocadas en los puntos de coordenadas $(0,0)$ y $(L,0)$. La localización de las masas m_1 y m_2 en función del tiempo t viene dada por las coordenadas $(x(t), H)$ con H una altura constante y $(L, y(t))$ respectivamente, ver Figura 7.1. Es fácil ver que la ecuación del movimiento para la masa m_1 es

$$(H^2 + x^2)[H^2 m_1 + (m_1 + m_2)x^2]\ddot{x} = -m_2 x[H^2 \dot{x}^2 + g\sqrt{(H^2 + x^2)^3}] ,$$

siendo g la aceleración de la gravedad. Además, se puede demostrar que el movimiento de la masa m_1 es periódico en un entorno del origen del eje x . Utilizar el método de Euler con una longitud de paso $h = 0,5$ para aproximar el periodo T de dicho movimiento suponiendo las condiciones iniciales $x(0) = 1$, $\dot{x}(0) = 0$. DATOS: $m_1 = 1$, $m_2 = 0,9$, $H = 0,1$, $g = 9,8$.

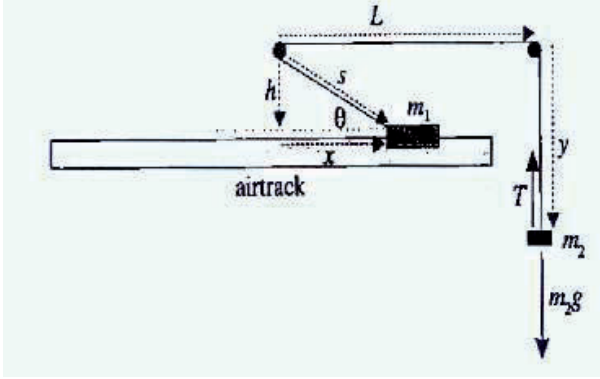


Figura 7.1: Oscilador mecánico.

Solución. Pasando la ecuación diferencial del movimiento no lineal de segundo orden a sistema de primer orden mediante el cambio $\dot{x} = v$ se obtiene

$$\dot{x} = f(t, x, v) = v , \quad \dot{v} = g(t, x, v) = - \frac{m_2 x [H^2 v^2 + g \sqrt{(H^2 + x^2)^3}]}{(H^2 + x^2)[H^2 m_1 + (m_1 + m_2)x^2]} .$$

Utilizando ahora el método de Euler para aproximar la solución $(x(t), v(t))$ del anterior sistema de ecuaciones diferenciales con la condición inicial $(x(0), v(0)) = (1, 0)$ se tiene la siguiente iteración

$$\begin{pmatrix} x_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ v_n \end{pmatrix} + h \begin{pmatrix} f(t_n, x_n, v_n) \\ g(t_n, x_n, v_n) \end{pmatrix} ,$$

siendo $t_n = nh$ con h la longitud de paso. En nuestro caso se tiene

$$\begin{pmatrix} x_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ v_n \end{pmatrix} + 0,5 \begin{pmatrix} v_n \\ - \frac{0,9x_n[9,8\sqrt{(0,01+x_n^2)^3} + 0,01v_n^2]}{(0,01+x_n^2)(0,01+1,9x_n^2)} \end{pmatrix} ,$$

con la condición inicial $(x_0, v_0) = (1, 0)$. Se genera de esta forma la siguiente sucesión

n	t_n	x_n	v_n
0	0	1	0
1	0.5	1	-2.32042
2	1	-0.160208	-4.65339
3	1.5	-2.4869	5.06518
4	2	0.0456857	7.39007

Observamos que la sucesión de puntos (x_n, v_n) con $n = 0, 1, 2, 3, 4$ realiza en el plano de fases $x - v$ un ciclo completo pasando de un cuadrante a otro entre puntos consecutivos. Por supuesto, debido principalmente a los errores de truncamiento o discretización asociados al método de Euler y a la longitud de paso utilizada, la sucesión de puntos obtenida no es periódica. De cualquier modo, una aproximación del periodo T del movimiento de la masa m_1 viene dado por

$$T \approx t_4 = 4h = 4 \times 0,5 = 2 .$$

Problema 7.13 *Un cohete tiene una velocidad $v(t)$ respecto de un cierto sistema de referencia inercial y sus gases de escape son expulsados a una velocidad constante V respecto del cohete. Supóngase que $0 \ll v/c < 1$, siendo $c = 3 \times 10^8 \text{m/s}$ la velocidad de la luz en el vacío, es decir para describir el movimiento del cohete es necesaria la teoría de la relatividad especial de Einstein. En estas condiciones se puede demostrar que la ecuación del movimiento del cohete es*

$$m_0 \frac{dv}{dt} + V \frac{dm_0}{dt} \left[1 - \frac{v^2}{c^2} \right]^{3/2} = 0 ,$$

donde $m_0(t)$ es la masa variable del cohete en su sistema de referencia en reposo.

- (i) *Supongamos que se conoce la siguiente tabulación en el sistema internacional de la función $m_0(t)$*

t_i	0	10	20	30	40
$m_0(t_i)$	1200	1100	1005	910	900

Calcular una tabulación aproximada de dm_0/dt para los tiempos $t_i = 10i$, con $i = 0, 1, 2, 3, 4$, utilizando de entre las fórmulas de diferencias hacia adelante, hacia atrás y centradas, la más adecuada en cada caso.

- (ii) *Utilizando el resultado del apartado anterior, calcular mediante el método de Euler con una longitud de paso adecuada, una aproximación de $v(20)$ sabiendo que $V = 10^3 \text{m/s}$ y $v(0) = 10^4$.*

Solución. (i) Las fórmulas de diferencias hacia adelante, hacia atrás y centradas son

$$\begin{aligned}\frac{dm_0}{dt}(t_i) &= \frac{m_0(t_i + h) - m_0(t_i)}{h} + O(h) , \\ \frac{dm_0}{dt}(t_i) &= \frac{m_0(t_i) - m_0(t_i - h)}{h} + O(h) , \\ \frac{dm_0}{dt}(t_i) &= \frac{m_0(t_i + h) - m_0(t_i - h)}{2h} + O(h^2) ,\end{aligned}$$

respectivamente, siendo h la longitud de paso. Siempre que sea posible (es decir, en todos los nodos interiores) utilizaremos la fórmula de diferencias centradas puesto que es la que contiene un menor error de truncamiento. Para el primer y último nodo sólo se pueden usar las fórmulas de diferencias hacia adelante y hacia atrás respectivamente. En definitiva, tomando la longitud de paso $h = 10$, se tiene

$$\begin{aligned}\frac{dm_0}{dt}(0) &\approx \frac{m_0(10) - m_0(0)}{10} = \frac{1100 - 1200}{10} = -10 , \\ \frac{dm_0}{dt}(10) &\approx \frac{m_0(20) - m_0(0)}{20} = \frac{1005 - 1200}{20} = -\frac{39}{4} , \\ \frac{dm_0}{dt}(20) &\approx \frac{m_0(30) - m_0(10)}{20} = \frac{910 - 1100}{20} = -\frac{19}{2} , \\ \frac{dm_0}{dt}(30) &\approx \frac{m_0(40) - m_0(20)}{20} = \frac{900 - 1005}{20} = -\frac{21}{4} , \\ \frac{dm_0}{dt}(40) &\approx \frac{m_0(40) - m_0(30)}{10} = \frac{900 - 910}{10} = -1 .\end{aligned}$$

(ii) La función $v(t)$ es la solución del problema de Cauchy

$$\frac{dv}{dt} = f(t, v) , \quad v(0) = 10^4 ,$$

siendo

$$f(t, v) = -\frac{V}{m_0(t)} \frac{dm_0}{dt}(t) \left[1 - \frac{v^2}{c^2} \right]^{3/2} .$$

Teniendo en cuenta que $c = 3 \times 10^8$, $V = 10^3$ y los valores $m_0(t_i)$ y $\frac{dm_0}{dt}(t_i)$ para los tiempos $t_i = 10i$, con $i = 0, 1, 2, 3, 4$, se pueden obtener del apartado anterior, podemos aplicar el método de Euler

$$v_{i+1} = v_i + hf(t_i, v_i) , \quad i = 0, 1, 2, 3 ,$$

donde $v(t_i) \approx v_i$. Tomaremos $h = 10$ de modo que los tiempos discretizados de forma equiespaciada sean justo $t_i = 10i$, es decir, los de la tabulación de que disponemos en el apartado anterior. De esta forma se tiene

$$v_1 = v_0 + 10f(t_0, v_0) = 10^4 + 10f(0, 10^4)$$

$$\begin{aligned}
&= 10^4 - 10 \times \frac{10^3}{m_0(0)} \frac{dm_0}{dt}(0) \left[1 - \frac{(10^4)^2}{(3 \times 10^8)^2} \right]^{3/2} \\
&= 10^4 - 10 \times \frac{10^3}{1200} \times (-10) \left[1 - \frac{(10^4)^2}{(3 \times 10^8)^2} \right]^{3/2} = 10083,3, \\
v_2 &= v_1 + 10f(t_1, v_1) = 10083,3 + 10f(10, 10083,3) \\
&= 10083,3 - 10 \times \frac{10^3}{m_0(10)} \frac{dm_0}{dt}(10) \left[1 - \frac{(10083,3)^2}{(3 \times 10^8)^2} \right]^{3/2} \\
&= 10083,3 - 10 \times \frac{10^3}{1100} \left(-\frac{39}{4} \right) \left[1 - \frac{(10083,3)^2}{(3 \times 10^8)^2} \right]^{3/2} = 10171,9
\end{aligned}$$

Concluimos que $v(20) \approx v_2 = 10171,9 \text{ m/s}$.

Problema 7.14 *Un proyectil de masa m es lanzado desde la superficie terrestre verticalmente hacia arriba con una velocidad inicial v_0 . El proyectil es frenado debido a la acción de la gravedad y mediante una fuerza de fricción con el aire proporcional a la velocidad v del proyectil.*

- (i) *Suponiendo la gravedad constante, aplicar las leyes de la dinámica clásica y hallar la ecuación diferencial que verifica la función $v(t)$, siendo t el tiempo.*
- (ii) *Sabiendo que la masa es $m = 2$, la constante de proporcionalidad en la fuerza de fricción $k = 4$ y la velocidad inicial $v_0 = 20$, utilizar el método de Heun con una longitud de paso $h = 1/2$ para aproximar el tiempo τ que tarda el proyectil en llegar al punto de altura máximo.*
- (iii) *Calcular el error absoluto cometido en el apartado anterior.*

Solución. (i) Tomando el sistema de referencia en la dirección vertical y positivo en el sentido contrario a la gravedad y aplicando la segunda ley de Newton se obtiene que la ecuación diferencial que verifica la función $v(t)$, siendo t el tiempo, es

$$m\dot{v} = -mg - kv,$$

siendo $k > 0$ la constante de proporcionalidad en la fuerza de fricción.

(ii) Reescribimos la ecuación diferencial del apartado anterior despejando la derivada y obtenemos

$$\dot{v} = f(t, v) = -g - \frac{k}{m}v = -9,8 - 2v$$

El problema de Cauchy está formado por la anterior ecuación diferencial acompañada de la condición inicial $v(0) = v_0 = 20$. Aplicaremos el método de Heun para aproximar la solución $v(t)$. Dicho método consiste en realizar la iteración

predictor–corrector

$$\begin{aligned} v_{i+1}^* &= v_i + hf(t_i, v_i) = v_i + \frac{1}{2}(-9,8 - 2v_i) , \\ v_{i+1} &= v_i + \frac{h}{2} [f(t_i, v_i) + f(t_{i+1}, v_{i+1}^*)] \\ &= v_i + \frac{1}{4} [-19,6 - 2(v_i + v_{i+1}^*)] , \end{aligned}$$

donde los tiempos t_i son equiespaciados con longitud de paso h , es decir, $t_i = ih = i/2$ con $i = 0, 1, 2, \dots$. De este modo se obtienen las aproximaciones $v(t_i) \approx v_i$ dadas en la siguiente tabla

i	t_i	v_i
0	0	20
1	0.5	7.55
2	1	1.325
3	1.5	-1.7875

Hemos parado las iteraciones puesto que el tiempo τ verificará $v(\tau) = 0$. Entonces, el método de Heun predice que $\tau \in (1, 1,5)$. Una aproximación es, por ejemplo, el punto medio del intervalo $\tau \approx 1,25$.

(iii) Para calcular el error absoluto cometido en el apartado anterior, hallaremos τ de manera exacta. En primer lugar, resolvemos el problema de valor inicial

$$\dot{v} = -g - \frac{k}{m}v , \quad v(0) = v_0 .$$

Separando las variables en la ecuación diferencial e integrando se tiene

$$\int_{v_0}^v \frac{dv}{g + \frac{k}{m}v} = - \int_0^t dt ,$$

de modo que

$$v(t) = \frac{m}{k} \left[-g + \left(g + \frac{k}{m}v_0 \right) \exp \left(-\frac{k}{m}t \right) \right] .$$

Con los datos del problema se tiene

$$v(t) = \frac{1}{2} [-9,8 + 49,8 \exp(-2t)] ,$$

de modo que, la ecuación $v(\tau) = 0$ tiene por solución

$$\tau = \frac{m}{k} \ln \left(1 + \frac{v_0 k}{mg} \right) = 0,8128 .$$

El error absoluto cometido en el apartado anterior es

$$\Delta = |0,8128 - 1,25| = 0,4372 .$$

Problema 7.15 Una viga horizontal, colocada a lo largo del eje x se pandea bajo la influencia de diversas cargas verticales. Según la teoría de la elasticidad, la curva de flexión $y(x)$ se puede determinar, para pequeñas flexiones, mediante la ecuación diferencial lineal de segundo orden

$$EI \frac{d^2 y}{dx^2} = M(x) ,$$

siendo $M(x)$ el momento flector en x , E el módulo de Young del material e I el momento de inercia de una sección transversal de la viga respecto a su eje central.

- (i) Se tiene una viga con un extremo (situado en $x = 0$) empotrado en una pared, es decir, $y(0) = \frac{dy}{dx}(0) = 0$ y tal que $M(x) = 10(1 - x)$ con $EI = 2$. Aproximar la flexión de la viga en $x = 2$ mediante el método de Runge-Kutta de orden 4 con una longitud de paso $h = 1$.
- (ii) Hallar la flexión exacta $y(x)$ de la viga del apartado anterior. ¿Se calculó una buena aproximación en el apartado (i)?

Solución. (i) Con los datos del problema, la flexión $y(x)$ de la viga verifica la ecuación diferencial

$$\frac{d^2 y}{dx^2} = 5(1 - x) ,$$

con las condiciones iniciales $y(0) = \frac{dy}{dx}(0) = 0$. Realizando el cambio de variable usual $\frac{dy}{dx} = z$, reescribimos el anterior problema de Cauchy de segundo orden como un problema de Cauchy de primer orden de la forma

$$y' = z , \quad z' = 5(1 - x) , \quad (y(0), z(0)) = (0, 0) ,$$

donde la prima indica derivada respecto de x . Escrito en forma vectorial, se tiene

$$\mathbf{w}' = \mathbf{f}(x, \mathbf{w}) , \quad \mathbf{w}(0) = (0, 0) .$$

Aquí, $x \in \mathbb{R}$ es la variable independiente, $\mathbf{w} = (y, z)^t \in \mathbb{R}^2$ es la variable vectorial dependiente y $\mathbf{f}(x, \mathbf{w}) = (z, 5(1 - x))^t \in \mathbb{R}^2$. La iteración RK4 con longitud de paso h consiste en

$$\mathbf{w}_{i+1} = \mathbf{w}_i + \frac{h}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4) ,$$

siendo

$$\begin{aligned} \mathbf{k}_1 &= \mathbf{f}(x_i, \mathbf{w}_i) , \\ \mathbf{k}_2 &= \mathbf{f}\left(x_i + \frac{h}{2}, \mathbf{w}_i + \frac{h}{2}\mathbf{k}_1\right) , \\ \mathbf{k}_3 &= \mathbf{f}\left(x_i + \frac{h}{2}, \mathbf{w}_i + \frac{h}{2}\mathbf{k}_2\right) , \\ \mathbf{k}_4 &= \mathbf{f}(x_i + h, \mathbf{w}_i + h\mathbf{k}_3) . \end{aligned}$$

Nos piden aproximar la flexión de la viga en $x = 2$, es decir, $y(2)$. Para ello calcularemos $\mathbf{w}(2) = (y(2), z(2))^t$ mediante el método de Runge-Kutta de orden 4 con una longitud de paso $h = 1$. La tabla que se genera es la siguiente:

i	x_i	$\mathbf{w}_i = (y_i, z_i)$
0	0	(0,0)
1	1	(5/3 , 5/2)
2	2	(10/3 , 0)

Se tiene pues que $y(2) \approx y_2 = 10/3$.

NOTA: En este problema se puede también utilizar el método RK4 para la ecuación escalar $y' = f(x)$ con $f(x) = 5 \left(x - \frac{x^2}{2} \right)$ como se puede ver en el siguiente apartado.

(ii) La flexión $y(x)$ de la viga verifica la ecuación diferencial

$$\frac{d^2 y}{dx^2} = 5(1 - x) ,$$

con las condiciones iniciales $y(0) = \frac{dy}{dx}(0) = 0$. Podemos hallar la solución $y(x)$ de manera exacta integrando dos veces la ecuación diferencial respecto de x . Integrando una vez se tiene que

$$\frac{dy}{dx} = 5 \int (1 - x) dx + C_1 = 5 \left(x - \frac{x^2}{2} \right) + C_1 ,$$

siendo C_1 una constante arbitraria que calculamos con la condición inicial $\frac{dy}{dx}(0) = 0$, es decir, $C_1 = 0$. Integrando de nuevo obtenemos

$$y(x) = 5 \int \left(x - \frac{x^2}{2} \right) dx + C_2 = 5 \left(\frac{x^2}{2} - \frac{x^3}{6} \right) + C_2 ,$$

siendo C_2 una constante arbitraria que calculamos con la condición inicial $y(0) = 0$, es decir, $C_2 = 0$. En resumen, la flexión exacta de la viga es

$$y(x) = \frac{5}{6}(3 - x)x^2 ,$$

de modo que $y(2) = 10/3$. Entonces, el método RK4 del apartado (i) calculó $y(2)$ de manera exacta, es decir, con un error absoluto cero.

Problema 7.16 Consideremos un circuito eléctrico LRC que consta de una bobina de inductancia L , una resistencia R y un condensador de capacidad C . Conectamos los dispositivos en serie con un generador de corriente que proporciona un voltaje dependiente del tiempo $V(t)$. Entonces, se sabe que la carga eléctrica $Q(t)$ almacenada en el condensador satisface la ecuación diferencial

$$L\ddot{Q} + R\dot{Q} + \frac{Q}{C} = V(t) .$$

Si $L = R = 2$, $C = 1$ y $V(t) = 3 \sin(t)$, aproximar mediante el método de Euler con una longitud de paso $h = 1$ la carga del condensador para tiempo $t = 2$, sabiendo que inicialmente tenía una carga $Q(0) = 3$. Recordar que la intensidad de corriente $I(t)$ que circula es, por definición, $I = \dot{Q}$ siendo $I(0) = 0$.

Solución. Con los datos del enunciado, la carga eléctrica $Q(t)$ almacenada en el condensador satisface la ecuación diferencial de segundo orden $2\ddot{Q} + 2\dot{Q} + Q = 3 \sin(t)$. Esta ecuación se escribe como un sistema de primer orden de la forma

$$\dot{Q} = I , \quad \dot{I} = \frac{1}{2} [3 \sin(t) - 2I - Q] .$$

Se satisface la condición inicial $(Q(0), I(0)) = (3, 0)$. En definitiva, definiendo el vector $x = (Q, I) \in \mathbb{R}^2$ y la función $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ de la forma

$$f(t, x) = \left(I, \frac{1}{2} [3 \sin(t) - 2I - Q] \right) ,$$

la función vectorial $x(t)$ es la solución del problema de Cauchy

$$\dot{x} = f(t, x) , \quad x(0) = (3, 0) .$$

Utilizando el método de Euler con una longitud de paso $h = 1$, se tiene que las aproximaciones $x(t_n) = (Q(t_n), I(t_n)) \approx x_n = (Q_n, I_n) \in \mathbb{R}^2$ para los tiempos equiespaciados $t_n = 0 + nh = n$ con $n \in \mathbb{N}$ vienen dadas de forma recurrente por la iteración $x_{n+1} = x_n + hf(t_n, x_n)$. Se tiene pues

$$\begin{pmatrix} Q_{n+1} \\ I_{n+1} \end{pmatrix} = \begin{pmatrix} Q_n \\ I_n \end{pmatrix} + \begin{pmatrix} I_n \\ \frac{1}{2} [3 \sin(t_n) - 2I_n - Q_n] \end{pmatrix} , \quad \begin{pmatrix} Q_0 \\ I_0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix} .$$

Realizando la iteración se llega a la tabla siguiente.

n	t_n	(Q_n, I_n)
0	0	(3, 0)
1	1	(3, -3/2)
2	2	(1.5, -0.237794)

Se tiene entonces la aproximación buscada $Q(2) \approx Q_2 = 1.5$.

Problema 7.17 Considerar el problema de Cauchy

$$\frac{dy}{dx} = \frac{xy}{3x^2 - y^4}, \quad y(1) = 1.$$

- (i) Estudiar la existencia y unicidad de la solución.
 (ii) Averiguar si alguna de las siguientes funciones es la solución

$$(a) \ y(x) = \sqrt{x}, \quad (b) \ y(x) = x^2.$$

- (iii) Aproximar el valor de la solución en $x = 2$ mediante el método de Heun con una longitud de paso 0,5. Calcular también el error absoluto cometido.

Solución. (i) El problema de Cauchy $y' = f(x, y)$, $y(1) = 1$ con

$$f(x, y) = \frac{xy}{3x^2 - y^4},$$

tiene una única solución puesto que la función $f(x, y)$ es racional y el denominador no se anula en el punto $(x_0, y_0) = (1, 1)$. En particular esto implica que las funciones f y $\partial f / \partial y$ son continuas en un entorno del punto $(1, 1)$ y por lo tanto se puede utilizar el Teorema de Picard–Lindelöf concluyendo la existencia y unicidad de la solución del problema de Cauchy.

(ii) Una función $y(x)$ de clase C^1 en un intervalo $I \subset \mathbb{R}$ que contenga a $x = 1$ será solución del problema de Cauchy si verifica $y'(x) = f(x, y(x))$ para todo $x \in I$ y además satisface la condición inicial $y(1) = 1$. Es obvio que esta condición inicial es satisfecha por las dos funciones del enunciado del problema. Veamos pues si satisfacen la ecuación diferencial:

(a) Sea $y(x) = \sqrt{x}$. Entonces se tiene que

$$y'(x) - f(x, y(x)) = \frac{1}{2\sqrt{x}} - \frac{x\sqrt{x}}{3x^2 - (\sqrt{x})^4} = \frac{1}{2\sqrt{x}} - \frac{\sqrt{x}}{2x} = 0,$$

de modo que la función $y(x) = \sqrt{x}$ satisface la ecuación diferencial y por lo tanto es la solución del problema de Cauchy.

(b) Sea $y(x) = x^2$. Es obvio que esta función no puede ser solución del problema de Cauchy debido a que \sqrt{x} lo es y la unicidad de la solución. De todos modos veamos que la función $y(x) = x^2$ no satisface la ecuación diferencial.

$$y'(x) - f(x, y(x)) = 2x - \frac{xx^2}{3x^2 - (x^2)^4} = 2x - \frac{x}{3 - x^6} \neq 0.$$

(iii) Sea $y(x)$ la solución del problema de Cauchy $y' = f(x, y)$, $y(x_0) = y_0$. Discretizando equiespaciadamente la variable independiente con una longitud de paso h , definimos $x_i = x_0 + ih$ con $i = 0, 1, 2, \dots$. Si aproximamos $y(x_i) \approx y_i$, el

método de Heun obtiene dichos valores mediante el siguiente método predictor–corrector:

$$\begin{aligned} y_{i+1}^* &= y_i + hf(x_i, y_i) , \\ y_{i+1} &= y_i + \frac{h}{2} \left[f(x_i, y_i) + f(x_{i+1}, y_{i+1}^*) \right] , \quad \text{para } i = 0, 1, \dots \end{aligned}$$

En nuestro caso se tiene $h = 1/2$ de modo que

$$\begin{aligned} y_{i+1}^* &= y_i + \frac{1}{2} \frac{x_i y_i}{3x_i^2 - y_i^4} , \\ y_{i+1} &= y_i + \frac{1}{4} \left[\frac{x_i y_i}{3x_i^2 - y_i^4} + \frac{x_{i+1} y_{i+1}^*}{3x_{i+1}^2 - y_{i+1}^{*4}} \right] , \quad \text{para } i = 0, 1, \dots \end{aligned}$$

Teniendo en cuenta que $(x_0, y_0) = (1, 1)$, realizamos la anterior iteración y mostramos los resultados en la tabla siguiente.

i	x_i	y_i^*	y_i
0	1	1	1
1	1.5	1.25	1.23379
2	2	1.44255	1.43221

Se tiene pues la aproximación $y(2) \approx y_2 = 1,43221$.

Puesto que sabemos del apartado (ii) que la solución exacta es $y(x) = \sqrt{x}$, el error absoluto Δ cometido en la anterior aproximación es

$$\Delta = |y(2) - y_2| = |\sqrt{2} - 1,43221| = 0,0179964 .$$

Problema 7.18 Consideremos un péndulo simple de longitud ℓ sin rozamiento. Se puede demostrar que, si el punto de suspensión del péndulo sigue la trayectoria plana $(x(t), y(t))$ parametrizada con el tiempo t en un sistema de referencia cartesiano, entonces el ángulo φ que forma el péndulo con la vertical (eje y) verifica la ecuación diferencial no lineal

$$\ell \ddot{\varphi} + (\ddot{y} + g) \sin \varphi + \ddot{x} \cos \varphi = 0 ,$$

siendo g la aceleración de la gravedad. Si $\ell = 1$, $g = 9,8$ y la trayectoria del punto de suspensión es $(x(t), y(t)) = (t^2 + t, 7t)$, aproximar el valor de la velocidad angular $\dot{\varphi}$ cuando el tiempo es $t = 2$ sabiendo que inicialmente el péndulo se encuentra en el estado $(\varphi(0), \dot{\varphi}(0)) = (3, 1)$. Usar el método de Euler con una longitud de paso $h = 1$.

Solución. Teniendo en cuenta que la posición en función del tiempo del punto de suspensión es $(x(t), y(t)) = (t^2 + t, 7t)$, su aceleración es $(\ddot{x}(t), \ddot{y}(t)) = (2, 0)$. Entonces, introduciendo el dato $\ell = 1$, la función $\varphi(t)$ satisface el problema de Cauchy de segundo orden

$$\ddot{\varphi} = -g \sin \varphi - 2 \cos \varphi , \quad (\varphi(0), \dot{\varphi}(0)) = (3, 1) .$$

Definimos la velocidad angular w como $\dot{\varphi} = w$ y transformamos el problema de Cauchy de segundo orden anterior en un problema de Cauchy de primer orden. En concreto,

$$\dot{z} = f(t, z), \quad z(0) = (3, 1),$$

siendo $z = (\varphi, w) \in \mathbb{R}^2$ y $f(t, z) = (w, -g \sin \varphi - 2 \cos \varphi) \in \mathbb{R}^2$. Tomando los tiempos equiespaciados $t_i = t_0 + ih = i$ con $i \in \mathbb{N}$ y aplicando el método de Euler $z_{i+1} = z_i + hf(t_i, z_i)$, se obtiene la siguiente iteración:

$$\begin{pmatrix} \varphi_{i+1} \\ w_{i+1} \end{pmatrix} = \begin{pmatrix} \varphi_i \\ w_i \end{pmatrix} + \begin{pmatrix} w_i \\ -g \sin \varphi_i - 2 \cos \varphi_i \end{pmatrix}$$

con la condición inicial $(\varphi_i, w_i) = (3, 1)$. Realizando dicha iteración, se obtiene la tabla siguiente:

i	t_i	(φ_i, w_i)
0	0	(3,1)
1	1	(4, 1.59701)
2	2	(5.59701, 10.321)

Se concluye que la velocidad angular $w(2)$ se aproxima por $w(2) \approx w_2 = 10,321$.

Problema 7.19 Sea $z(t)$ la solución del problema de valor inicial $\dot{z} = \ln t$, $z(1) = 3$. Obtener el error absoluto que se comete al aproximar el valor de $z(2)$ mediante el método de Taylor de segundo orden con una longitud de paso $h = 1/2$.

Solución. La solución $z(t)$ del problema de valor inicial $\dot{z} = f(t, z) = \ln t$, $z(1) = 3$, se puede aproximar mediante el método de Taylor de segundo orden con una longitud de paso $h = 1/2$ de la forma

$$z_{i+1} = z_i + hf(t_i, z_i) + \frac{h^2}{2} \left[\frac{\partial f(t_i, z_i)}{\partial t} + \frac{\partial f(t_i, z_i)}{\partial z} f(t_i, z_i) \right].$$

En nuestro problema, las derivadas parciales de primer orden de la función f son

$$\frac{\partial f(t, z)}{\partial t} = \frac{1}{t}, \quad \frac{\partial f(t, z)}{\partial z} = 0,$$

de manera que el método de Taylor adopta la forma

$$z_{i+1} = z_i + \frac{1}{2} \ln t_i + \frac{1}{8t_i}.$$

Realizando esta iteración con el valor inicial $(t_0, z_0) = (1, 3)$ obtenemos la tabla siguiente

i	t_i	z_i
0	1	3
1	3/2	25/8
2	2	3.41107

Se finaliza con la aproximación $z(2) \approx z_2 = 3,41107$.

El problema de valor inicial $\dot{z} = \ln t$, $z(1) = 3$ se puede resolver de manera exacta puesto que la ecuación diferencial es de variables separables. Se tiene entonces que $z(t) - z(1) = \int_1^t \ln s \, ds$, es decir, $z(t) = 3 + \int_1^t \ln s \, ds$. La integral que aparece se puede calcular mediante integración por partes obteniéndose

$$\int_1^t \ln s \, ds = \left[s \ln s \right]_1^t - \int_1^t ds = t \ln t - t + 1 .$$

En definitiva $z(t) = t \ln t - t + 4$, de modo que $z(2) = 2(1 + \ln 2) = 3,38629$. Entonces, se concluye que el error absoluto Δ que se comete al aproximar el valor de $z(2)$ mediante el método de Taylor de segundo orden es

$$\Delta = |z(2) - z_2| = |3,38629 - 3,41107| = 0,02478 .$$

Capítulo 8

Solución de Ecuaciones no Lineales

8.1. Resumen de teoría

Son muchos los problemas en ciencia e ingeniería que se pueden modelar matemáticamente como una ecuación $f(x) = 0$, siendo f una función de la variable x . Los valores de x soluciones de dicha ecuación son llamados *ceros* de la función f o *raíces* de la ecuación. Existen muchas de estas ecuaciones que no admiten que su solución pueda ser expresada a través de funciones elementales.

El objetivo del presente capítulo es el estudio de algunas de las técnicas numéricas que nos permitan abordar este tipo de problemas. Es importante destacar el hecho de que las técnicas que estudiaremos en este capítulo son siempre *iterativas*, es decir, partiremos de una aproximación inicial x_0 del cero exacto x^* de f y posteriormente construiremos una sucesión de números reales $\{x_n\}_{n=0}^{\infty}$ que converja hacia x^* cuando $n \rightarrow \infty$.

Teorema 8.1 (Bolzano) Sea f una función continua en $[a, b]$ tal que cambia de signo en los extremos, es decir, $f(a)f(b) < 0$. Entonces, existe al menos un $\xi \in (a, b)$ tal que $f(\xi) = 0$.

El *algoritmo de bisección* se basa en el teorema de Bolzano y consiste en ir estrechando de manera sistemática el intervalo en el cual una función continua cambia de signo. De esta forma conseguiremos obtener un intervalo arbitrariamente pequeño que contenga el cero de la función. El *método de bisección* para una función $f \in C[a_0, b_0]$ procede de la forma siguiente: $(a_0, b_0) \supset (a_1, b_1) \supset \cdots \supset (a_n, b_n) \supset \cdots$. La sucesión de puntos medios $\{c_n\}_{n=0}^{\infty}$ siendo $c_n = (a_n + b_n)/2$, verifica $\lim_{n \rightarrow \infty} c_n = x^*$ con $|x^* - c_n| \leq (b - a)/2^{n+1}$.

- *Método de Newton-Raphson:*

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

Teorema 8.2 (Convergencia de la sucesión de Newton-Raphson)

Supongamos que $f \in C^2[a, b]$ con $f(a)f(b) < 0$ y $f'(x)$ y $f''(x)$ son de signo constante $\forall x \in [a, b]$. Entonces, tomando $x_0 \in \{a, b\}$ que verifique $\text{sign}f(x_0) = \text{sign}f''(x_0)$, la sucesión de Newton-Raphson $\{x_n\}_{n \in \mathbb{N}}$ converge hacia la única raíz de f en $[a, b]$.

■ *Método de la Secante:*

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

Consideremos un sistema de n ecuaciones con n incógnitas x^1, x^2, \dots, x^n . Dicho sistema puede ser escrito de la forma $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, siendo

$$\mathbf{x} = (x^1, x^2, \dots, x^n)^T.$$

La aplicación $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tiene las funciones componentes $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ para $i = 1, 2, \dots, n$, es decir

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x}))^T.$$

Sea \mathbf{x}^* un cero de la función \mathbf{f} , es decir $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$. El *método de Newton-Raphson* para aproximar las soluciones del sistema $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ es

$$\mathbf{x}_{i+1} = \mathbf{x}_i - (D\mathbf{f}(\mathbf{x}_i))^{-1} \mathbf{f}(\mathbf{x}_i) \quad \text{con } i = 0, 1, \dots, \quad (8.1)$$

siendo $D\mathbf{f}(\mathbf{x}_i)$ la matriz Jacobiana de la función \mathbf{f} en el punto \mathbf{x}_i , es decir

$$D\mathbf{f}(\mathbf{x}_i) = \begin{pmatrix} \frac{\partial f_1(\mathbf{x}_i)}{\partial x^1} & \dots & \frac{\partial f_1(\mathbf{x}_i)}{\partial x^n} \\ \vdots & & \vdots \\ \frac{\partial f_n(\mathbf{x}_i)}{\partial x^1} & \dots & \frac{\partial f_n(\mathbf{x}_i)}{\partial x^n} \end{pmatrix}.$$

Si la matriz Jacobiana $D\mathbf{f}(\mathbf{x}_i)$ es no singular, es decir si $\det D\mathbf{f}(\mathbf{x}_i) \neq 0$, entonces se puede realizar la iteración.

8.2. Problemas resueltos

Problema 8.1 Consideremos el intervalo $I = [0, 1/2]$ y sea $g : I \rightarrow I$ una función derivable en I cuya derivada está acotada de la forma $|g'(x)| \leq 1/2$ para todo $x \in I$. Partiendo de un punto arbitrario $x_0 \in I$, se realiza la iteración $x_{n+1} = g(x_n)$.

- (i) Explicar si dicha iteración será convergente o no. Si lo fuera, es decir si $\lim_{n \rightarrow \infty} x_n = x^*$, averiguar el valor $g(x^*)$.
- (ii) Calcular una cota superior de $|x_8 - x^*|$ en función de la distancia entre los dos primeros términos de la sucesión $|x_1 - x_0|$.

Solución. La solución del problema viene dada por el siguiente teorema: Si $g \in C[a, b]$ y $g([a, b]) \subseteq [a, b]$, entonces existe al menos un punto fijo $x^* \in [a, b]$ de la función g . Si además $g \in D(a, b)$ y es contractiva, es decir $|g'(x)| \leq L < 1$ para todo $x \in (a, b)$ entonces g tiene un único punto fijo $x^* \in [a, b]$. Además la sucesión generada mediante la iteración $x_{n+1} = g(x_n)$ converge hacia x^* para cualquier valor inicial $x_0 \in [a, b]$. Además $|x^* - x_n| \leq \frac{L^n}{1-L} |x_1 - x_0|$.

(i) Como la función del enunciado es $g : I \rightarrow I$, entonces si se toma $I = [a, b]$ es evidente que $g([a, b]) \subseteq [a, b]$. Además, como g es una función derivable en I entonces se tiene que también es continua en I , es decir $g \in C[a, b]$. Finalmente, del enunciado se desprende que la constante $L = 1/2 < 1$. Se concluye que la sucesión generada con la recurrencia $x_{n+1} = g(x_n)$ converge hacia el único punto fijo x^* de g y por lo tanto $g(x^*) = x^*$.

(ii) La cota superior es

$$|x^* - x_8| \leq \frac{(1/2)^8}{1 - 1/2} |x_1 - x_0| = \frac{1}{128} |x_1 - x_0| = 0,0078125 |x_1 - x_0|.$$

Problema 8.2 Considerar el polinomio $P(x) = x^5 + 5x + 8$.

- (i) Encontrar dos números enteros negativos y consecutivos a y $a+1$ de manera que $P(x)$ tenga una única raíz $x^* \in I = [a, a+1]$.
- (ii) Averiguar cual de los dos extremos del intervalo I es el adecuado para iniciar la iteración de Newton-Raphson si se pretende asegurar la convergencia hacia x^* de la sucesión obtenida. Calcular la aproximación $x^* \approx x_2$.

Solución. (i) Como $P(x)$ es un polinomio entonces P es continuo en todo \mathbb{R} . Entonces, si encontramos dos números enteros consecutivos en los cuales cambie el signo de P se podrá aplicar el Teorema de Bolzano y concluir que entre esos dos puntos existe al menos una raíz de P . En concreto se tiene que $P(-1) = 2 > 0$ y $P(-2) = -34 < 0$ de manera que existe al menos un $x^* \in [-2, -1]$.

Para demostrar que x^* es único vamos a demostrar que $P(x)$ tiene derivada de signo constante en $[-2, -1]$ de manera que sólo puede cortar una vez al eje de abscisas. En concreto se tiene que $P'(x) = 5x^4 + 5 = 5(x^4 + 1) > 0$ para todo $x \in \mathbb{R}$.

(ii) Para resolver la segunda cuestión utilizaremos el siguiente teorema: Sea $f \in C^2[a, b]$ con $f(a)f(b) < 0$ tal que $f'(x)$ y $f''(x)$ tienen signo constante para todo $x \in [a, b]$. Sea $c \in \{a, b\}$ tal que $\text{sign}f(c) = \text{sign}f''(c)$. Entonces el método de Newton-Raphson para f con valor inicial $x_0 = c$ converge hacia la única raíz de f en $[a, b]$.

Ya se ha visto que $P'(x)$ tiene signo constante. Como $P''(x) = 20x^3 < 0$ para todo $x < 0$ se tiene que P'' es negativa en $[-2, -1]$. Finalmente, como $\text{sign}P(-2) = \text{sign}P''(-2)$, el anterior teorema asegura la convergencia de la sucesión generada con el método de Newton-Raphson para P tomando como

valor inicial $x_0 = -2$.

Realizando la iteración de Newton-Raphson

$$x_{n+1} = x_n - \frac{P(x_n)}{P'(x_n)} = x_n - \frac{x_n^5 + 5x_n + 8}{5(x_n^4 + 1)}$$

tomando $x_0 = -2$ se obtiene

n	x_n
0	-2
1	-1.58025
2	-1.27709

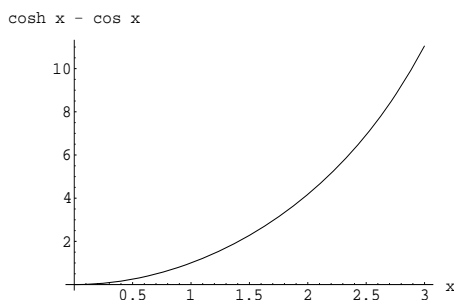
Se concluye que $x^* \approx x_2 = -1,27709$.

Problema 8.3 Considerar la función $g(x) = \sinh x + \sin x - 3x$.

(i) Demostrar que $g(x)$ tiene un extremo relativo en el intervalo $[1, 2]$.

INDICACIÓN: Las curvas $y = \sinh x$, $y = \sin x$ no tienen ningún punto de corte con abscisa en el intervalo $[1, 2]$.

(ii) Aproximar la abscisa x^* de dicho extremo mediante dos iteraciones del método de Newton-Raphson. Comenzar la recurrencia en el valor más adecuado $x_0 = 1$ o bien $x_0 = 2$. INDICACIÓN: Ver la figura adjunta.



Solución. (i) Una condición necesaria para que la función $g(x)$ tenga un extremo relativo en x^* es que la función $f(x) := g'(x)$ se anule para $x = x^*$. Notemos que $f(x) = \cosh x + \cos x - 3$ es una función continua en todo \mathbb{R} y por lo tanto en particular $f \in \mathcal{C}[1, 2]$. Como además $f(1)f(2) = (-0,916617) \times 0,346049 < 0$ se puede aplicar el Teorema de Bolzano de manera que se concluye con la existencia de al menos un valor $x^* \in (1, 2)$ tal que $f(x^*) = 0$. Una condición suficiente para que x^* sea un extremo relativo de g es que $g''(x^*) \neq 0$. Pero esta condición está garantizada puesto que si las curvas $y = \sinh x$, $y = \sin x$ no tiene ningún punto de corte en el intervalo $[1, 2]$ entonces la función $g''(x) = \sinh x - \sin x$ no

tiene ningún zero en dicho intervalo.

(ii) El extremo relativo x^* de $g(x)$ viene dado por la raíz de la función $f(x)$ en el intervalo $[1, 2]$. Recordemos que si $f \in \mathcal{C}^2[1, 2]$ y además $f''(x)$ tiene signo constante para toda $x \in [1, 2]$ entonces la recta tangente trazada desde el extremo del intervalo $[1, 2]$ adecuada para hallar la raíz de f por el método de Newton-Raphson es el extremo p (con $p = 1$ o $p = 2$) en el cual $\text{sig} f(p) = \text{sig} f''(p)$. A partir de la gráfica adjunta en el enunciado se comprueba que $f''(x)$ tiene signo constante en $[1, 2]$. Además, como $f(2) = 0,3460 > 0$ y $f''(2) = 4,1783 > 0$ resulta $p = 2$.

La iteración de Newton-Raphson viene dada por

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{\cosh x_n + \cos x_n - 3}{\sinh x_n - \sin x_n}, \quad x_0 = 2$$

de donde se obtiene la siguiente tabla

n	x_n	$f(x_n)$
0	2	0.346049
1	1.87266	0.0323507
2	1.8581	0.000382472

La aproximación buscada es $x^* \approx x_2 = 1,8581$.

Problema 8.4 La ecuación de ondas bidimensional $c^2(u_{xx} + u_{yy}) = u_{tt}$, donde c es la velocidad constante de propagación de la onda, se utiliza en ingeniería mecánica para modelar las vibraciones $u(x, y, t)$ de una placa elástica rectangular en las coordenadas espaciales (x, y) para tiempo t . Tomando el sistema de referencia cartesiano $x-y$ con origen en el vértice inferior izquierdo de la placa y sabiendo que, para un cierto instante de tiempo, el desplazamiento de la placa respecto de su posición de equilibrio viene dado por la función

$$u(x, y) = \sin x + \sin y + \cos(x + y),$$

se pretende hallar en qué punto (x^*, y^*) de la placa cuadrada de lado 2π se tiene un mínimo relativo de $u(x, y)$.

- (i) Aproximar dicho punto mediante dos iteraciones del método de Newton, sabiendo que $(x^*, y^*) \in [4\pi/3, 7\pi/4] \times [4\pi/3, 7\pi/4]$. Comenzar la iteración en el extremo inferior izquierdo del anterior cuadrado.
- (ii) Hallar los errores absolutos cometido en el anterior apartado sabiendo que la solución exacta es $(x^*, y^*) = (3\pi/2, 3\pi/2)$.
- (iii) Demostrar que la solución hallada es la única.

Solución. (i) Puesto que se ha de hallar un mínimo relativo de la función $u(x, y)$, en primer lugar obtendremos el sistema de ecuaciones que debe verificar

dicho punto, es decir, $\nabla u(x, y) := (u_x(x, y), u_y(x, y)) = (0, 0)$. Calculando las derivadas parciales de u se tiene

$$u_x(x, y) = \cos x - \sin(x + y) = 0, \quad u_y(x, y) = \cos y - \sin(x + y) = 0. \quad (8.2)$$

La matriz Jacobiana de la aplicación ∇u calculada en el punto (x_n, y_n) es

$$\begin{aligned} D_{\nabla u}(x_n, y_n) &= \begin{pmatrix} u_{xx}(x_n, y_n) & u_{xy}(x_n, y_n) \\ u_{yx}(x_n, y_n) & u_{yy}(x_n, y_n) \end{pmatrix} \\ &= \begin{pmatrix} -\sin x_n - \cos(x_n + y_n) & -\cos(x_n + y_n) \\ -\cos(x_n + y_n) & -\sin y_n - \cos(x_n + y_n) \end{pmatrix}, \end{aligned}$$

siendo su matriz inversa

$$D_{\nabla u}^{-1}(x_n, y_n) = \frac{1}{\delta} \begin{pmatrix} -\sin y_n - \cos(x_n + y_n) & \cos(x_n + y_n) \\ \cos(x_n + y_n) & -\sin x_n - \cos(x_n + y_n) \end{pmatrix},$$

donde $\delta := \det[D_{\nabla u}(x_n, y_n)] = \sin x_n \cos(x_n + y_n) + \sin y_n \cos(x_n + y_n) + \sin x_n \sin y_n$.

La iteración de Newton para aproximar una solución del sistema $\nabla u(x, y) = (0, 0)$ viene dada por

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - D_{\nabla u}^{-1}(x_n, y_n) \begin{pmatrix} u_x(x_n, y_n) \\ u_y(x_n, y_n) \end{pmatrix}. \quad (8.3)$$

En el primer paso, tomando el valor inicial $(x_0, y_0) = (4\pi/3, 4\pi/3)$ se tiene

$$\begin{aligned} D_{\nabla u}^{-1}(x_0, y_0) &= \frac{2}{3 + 2\sqrt{3}} \begin{pmatrix} 1 + \sqrt{3} & -1 \\ -1 & 1 + \sqrt{3} \end{pmatrix}, \\ (u_x(x_0, y_0), u_y(x_0, y_0)) &= (-[1 + \sqrt{3}]/2, -[1 + \sqrt{3}]/2), \end{aligned}$$

de manera que el primer iterado vale

$$\begin{aligned} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} &= \begin{pmatrix} \frac{4\pi}{3} \\ \frac{4\pi}{3} \end{pmatrix} - \frac{2}{3 + 2\sqrt{3}} \begin{pmatrix} 1 + \sqrt{3} & -1 \\ -1 & 1 + \sqrt{3} \end{pmatrix} \begin{pmatrix} -\frac{1+\sqrt{3}}{2} \\ -\frac{1+\sqrt{3}}{2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{9+12\pi+(3+8\pi)\sqrt{3}}{9+6\sqrt{3}} \\ \frac{9+12\pi+(3+8\pi)\sqrt{3}}{9+6\sqrt{3}} \end{pmatrix} \approx \begin{pmatrix} 4,92084 \\ 4,92084 \end{pmatrix}. \end{aligned}$$

Los cálculos en el segundo paso son los siguientes.

$$\begin{aligned} D_{\nabla u}^{-1}(x_1, y_1) &= D_{\nabla u}^{-1}(4,92084, 4,92084) \\ &= \begin{pmatrix} 0,689186 & -0,33294 \\ -0,33294 & 0,689186 \end{pmatrix}, \\ (u_x(x_1, y_1), u_y(x_1, y_1)) &= (u_x(4,92084, 4,92084), u_y(4,92084, 4,92084)) \\ &= (0,611877, 0,611877). \end{aligned}$$

Así, el segundo iterado del método de Newton es

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4,92084 \\ 4,92084 \end{pmatrix} - \begin{pmatrix} 0,689186 & -0,33294 \\ -0,33294 & 0,689186 \end{pmatrix} \begin{pmatrix} 0,611877 \\ 0,611877 \end{pmatrix} \\ = \begin{pmatrix} 4,70286 \\ 4,70286 \end{pmatrix}.$$

Se tiene pues la aproximación $(x^*, y^*) \approx (x_2, y_2) = (4,70286, 4,70286)$.

(ii) Los errores absolutos cometidos son

$$\Delta_{x^*} = |x^* - x_2| = \left| \frac{3\pi}{2} - 4,70286 \right| = 9,529 \times 10^{-3}, \\ \Delta_{y^*} = |y^* - y_2| = \left| \frac{3\pi}{2} - 4,70286 \right| = 9,529 \times 10^{-3}.$$

(iii) Resolvamos de manera exacta el sistema (8.2). En primer lugar, despejando $\cos x$ y $\cos y$ de la primera y segunda ecuación respectivamente del sistema (8.2) se obtiene que $\cos x = \cos y$, de manera que, recordando que $0 \leq x \leq 2\pi$ y $0 \leq y \leq 2\pi$, se tiene $x = y$ o bien $x + y = 2\pi$. Tenemos pues dos casos por analizar.

- Si $x + y = 2\pi$, de la primera ecuación del sistema (8.2) se tiene $\cos x = 0$, es decir $x = \pi/2$ o $x = 3\pi/2$. En definitiva se obtiene la pareja de soluciones

$$(x_1^*, y_1^*) = \left(\frac{\pi}{2}, \frac{3\pi}{2} \right), \quad (x_2^*, y_2^*) = \left(\frac{3\pi}{2}, \frac{\pi}{2} \right).$$

- Sea $x = y$. En este caso, la primera ecuación del sistema (8.2) adopta la forma $\cos x = \sin(2x)$, es decir $\cos x = 2 \sin x \cos x$. Entonces, o bien $\cos x = 0$ o $\sin x = 1/2$. De la primera opción se obtienen las soluciones

$$(x_3^*, y_3^*) = \left(\frac{\pi}{2}, \frac{\pi}{2} \right), \quad (x_4^*, y_4^*) = \left(\frac{3\pi}{2}, \frac{3\pi}{2} \right),$$

mientras que de la segunda opción se tiene

$$(x_5^*, y_5^*) = \left(\frac{\pi}{6}, \frac{\pi}{6} \right), \quad (x_6^*, y_6^*) = \left(\frac{5\pi}{6}, \frac{5\pi}{6} \right).$$

A continuación clasificaremos los puntos (x_i^*, y_i^*) con $i = 1, 2, 3, 4, 5, 6$ estudiando la matriz Jacobiana de ∇u en dichos puntos $D_{\nabla u}(x_i^*, y_i^*)$.

- La matriz Jacobiana de ∇u en (x_1^*, y_1^*) es

$$D_{\nabla u}(x_1^*, y_1^*) = \begin{pmatrix} -2 & -1 \\ -1 & 0 \end{pmatrix},$$

y sus determinantes menores principales son $\Delta_1 = -2 < 0$ y $\Delta_2 = -1 < 0$. Por lo tanto $D_{\nabla u}(x_1^*, y_1^*)$ no es definida y (x_1^*, y_1^*) es un punto de silla de la función $u(x, y)$.

- La matriz Jacobiana de ∇u en el punto (x_2^*, y_2^*) es

$$D_{\nabla u}(x_2^*, y_2^*) = \begin{pmatrix} 0 & -1 \\ -1 & 2 \end{pmatrix},$$

y sus determinantes menores principales son $\Delta_1 = 0$ y $\Delta_2 = -1 \neq 0$. Entonces $D_{\nabla u}(x_2^*, y_2^*)$ es no definida y el punto (x_2^*, y_2^*) es un punto de silla de la función $u(x, y)$.

- La matriz Jacobiana de ∇u en (x_3^*, y_3^*) es

$$D_{\nabla u}(x_3^*, y_3^*) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

siendo $\Delta_1 = 0$ y $\Delta_2 = -1 \neq 0$ sus determinantes menores principales. Concluimos que $D_{\nabla u}(x_3^*, y_3^*)$ no es definida y por lo tanto (x_3^*, y_3^*) es un punto de silla de la función $u(x, y)$.

- La matriz Jacobiana de ∇u en (x_4^*, y_4^*) es

$$D_{\nabla u}(x_4^*, y_4^*) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Como sus determinantes menores principales son $\Delta_1 = 2 > 0$ y $\Delta_2 = 3 > 0$ se tiene que $D_{\nabla u}(x_4^*, y_4^*)$ es definida positiva y por lo tanto (x_4^*, y_4^*) es un mínimo de la función $u(x, y)$.

- La matriz Jacobiana de ∇u en los puntos (x_5^*, y_5^*) y (x_6^*, y_6^*) coincide y vale

$$D_{\nabla u}(x_5^*, y_5^*) = D_{\nabla u}(x_6^*, y_6^*) = \begin{pmatrix} -1 & -1/2 \\ -1/2 & -1 \end{pmatrix}.$$

Sus determinantes menores principales son $\Delta_1 = -1 < 0$ y $\Delta_2 = 3/4 > 0$, de manera que la anterior matriz es definida negativa y por lo tanto los puntos (x_5^*, y_5^*) y (x_6^*, y_6^*) son máximos de la función $u(x, y)$.

En definitiva, la función $u(x, y)$ tiene un único mínimo relativo en la placa que es justo el punto (x_4^*, y_4^*) .

Problema 8.5 La deformación u de una esfera de radio R calentada no uniformemente con una distribución de temperaturas esféricamente simétrica $T(r)$ viene dada por

$$u(r) = \alpha \frac{1 + \sigma}{3(1 - \sigma)} \left[\frac{1}{r^2} \int_0^r r^2 T(r) dr + \frac{2(1 - 2\sigma)}{1 + \sigma} \frac{r}{R^3} \int_0^R r^2 T(r) dr \right],$$

donde σ y α son el coeficiente de Poisson y el coeficiente de dilatación térmica característicos del material que forma la esfera. Con esta formulación la temperatura $T(r)$ se mide tomando como origen la temperatura en la superficie de la esfera, es decir, $T(R) = 0$. DATOS: $R = 4$, $\alpha = 1$, $\sigma = 80$.

- (i) Sea la distribución de temperaturas $T(r) = (r - R) \exp(\sin r)$. Aplicando la regla de los Trapecios con una longitud de paso $h = 1$, aproximar el valor de la deformación sufrida por los puntos de la esfera que se encuentran a una distancia radial $r = R/2$. Obtener además una cota del error cometido sabiendo que la gráfica de la segunda derivada $[r^2 T(r)]''$ viene dada en la Figura 8.1.
- (ii) Supongamos que la distribución de temperaturas viene dada por $T(r) = (r - R)^3$.
- (ii.1) Demostrar que existe una única capa esférica donde la deformación es nula.
- (ii.2) Aproximar el radio de la capa esférica de deformación nula utilizando 3 iteraciones del método de Newton-Raphson con un valor inicial adecuado que nos asegure a priori la convergencia del método.

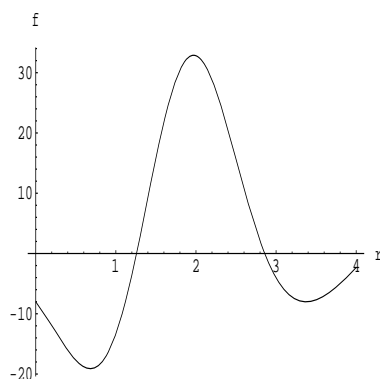


Figura 8.1: Gráfica de $f(r) = [r^2 T(r)]''$ para los valores $0 \leq r \leq 4$ cuando $T(r) = (r - R) \exp(\sin r)$.

Solución. Con los datos del enunciado $R = 4$, $\alpha = 1$, $\sigma = 80$, la expresión

para la deformación de la esfera es

$$u(r) = -\frac{27}{79r^2} \int_0^r r^2 T(r) \, dr + \frac{53}{2528} r \int_0^4 r^2 T(r) \, dr . \quad (8.4)$$

(i) Tomando la distribución de temperaturas $T(r) = (r - R) \exp(\sin r)$, la deformación de los puntos de la esfera situados en $r = R/2 = 2$ viene dada por

$$u(2) = -\frac{27}{316} \int_0^2 f(r) \, dr + \frac{53}{1264} \int_0^4 f(r) \, dr ,$$

siendo $f(r) = r^2 T(r) = r^2(r - 4) \exp(\sin r)$. Aproximaremos el valor de las dos integrales anteriores utilizando la regla de los Trapecios con una longitud de paso $h = 1$. Recordando que dicha regla es $\int_a^b f(r) \, dr \approx T_h(f) = h(f_0/2 + f_1 + \dots + f_{n-1} + f_n/2)$ se tiene que

$$\begin{aligned} \int_0^2 f(r) \, dr &\approx T_1(f) = \frac{f(0)}{2} + f(1) + \frac{f(2)}{2} \\ &= \frac{0}{2} - 3 \exp(\sin 1) - \frac{8 \exp(\sin 2)}{2} \\ &= -6,95933 - 9,93031 = -16,8896 , \\ \int_0^4 f(r) \, dr &\approx T_1(f) = \frac{f(0)}{2} + f(1) + f(2) + f(3) + \frac{f(4)}{2} \\ &= \frac{0}{2} - 3 \exp(\sin 1) - 8 \exp(\sin 2) - 9 \exp(\sin 3) + \frac{0}{2} \\ &= -6,95933 - 19,8606 - 10,3641 = -37,184 . \end{aligned}$$

En definitiva, se tiene

$$u(2) \approx -\frac{27}{316}(-16,8896) + \frac{53}{1264}(-37,184) = -0,116041 .$$

Por otra parte, se sabe que el error absoluto Δ cometido al utilizar la regla de los Trapecios viene dado por $\Delta := \int_a^b f(r) \, dr - T_h(f) = -\frac{h^2}{12}(b - a)f''(\xi)$ con $\xi \in (a, b)$. En nuestro problema hemos realizado dos aproximaciones y por lo tanto se tienen dos errores

$$\begin{aligned} \Delta_1 &:= \left| \int_0^2 f(r) \, dr - T_1(f) \right| = \frac{1}{6} |f''(\xi_1)| , \\ \Delta_2 &:= \left| \int_0^4 f(r) \, dr - T_1(f) \right| = \frac{1}{3} |f''(\xi_2)| , \end{aligned}$$

donde $\xi_1 \in (0, 2)$ y $\xi_2 \in (0, 4)$. A partir de la gráfica de la Figura 8.1 se obtienen, por ejemplo, las cotas siguientes: $|f''(\xi_1)| < 40$ y $|f''(\xi_2)| < 40$. Concluimos pues que $\Delta_1 < 20/3$ y $\Delta_2 < 40/3$.

Finalmente, a partir de la fórmula de propagación de errores absolutos se tiene que el error absoluto $\Delta_{u(2)}$ cometido en la aproximación efectuada de $u(2)$

viene dado por

$$\Delta_{u(2)} \approx \left| -\frac{27}{316} \right| \Delta_1 + \left| \frac{53}{1264} \right| \Delta_2 < \frac{27}{316} \frac{20}{3} + \frac{53}{1264} \frac{40}{3} = \frac{535}{474} = 1,12869 .$$

(ii.1) Tomando la distribución de temperaturas $T(r) = (r - R)^3 = (r - 4)^3$ en la ecuación (8.4) se tiene

$$u(r) = -\frac{27}{79r^2} \int_0^r r^2 (r - 4)^3 dr + \frac{53}{2528} r \int_0^4 r^2 (r - 4)^3 dr .$$

Integrando esta expresión es fácil ver que

$$u(r) = \frac{r}{2370} P(r) ,$$

donde $P(r) = 13888 - 9720r + 1944r^2 - 135r^3$. Los radios r^* de las capas esféricas que no sufren deformación vendrán dados por las raíces reales y positivas del polinomio P que pertenezcan al intervalo $[0, R] = [0, 4]$. Notemos que, por ser P un polinomio cúbico (grado impar) con coeficientes reales, se sabe que tendrá al menos una raíz real. Sin embargo hemos de demostrar que P tiene una única raíz real positiva en el intervalo $[0, 4]$. Una forma de verlo es la siguiente: en primer lugar demostraremos que P tiene al menos una raíz real r^* en el intervalo $[0, 4]$ mediante el Teorema de Bolzano y posteriormente mostraremos que es única estudiando la monotonía del polinomio P en $[0, 4]$.

Como $P(r)$ es un polinomio, entonces es continuo en todo \mathbb{R} . Además se tiene que $P(0) = 1388 > 0$ y $P(4) = -2528 < 0$ de manera que, por el Teorema de Bolzano, existe al menos un $r^* \in [0, 4]$ tal que $P(r^*) = 0$. Finalmente, como $P'(r) = -9720 + 3888r - 405r^2$ es un polinomio de segundo grado con discriminante negativo, es fácil comprobar que $P'(r) < 0$ para todo $r \in \mathbb{R}$. Esto implica que el polinomio P es monótono decreciente y por lo tanto hemos demostrado que r^* es única.

(ii.2) Iniciaremos la búsqueda iterativa de r^* utilizando el valor del extremo adecuado del intervalo $[0, 4]$ que nos asegure la convergencia del método de Newton-Raphson hacia la raíz r^* del polinomio P . Puesto que $P \in C^2[0, 4]$ con $P(0)P(4) < 0$ y además $P'(r)$ y $P''(r)$ tienen signo constante para todo $r \in [0, 4]$, se elige el valor $r_0 \in \{0, 4\}$ que verifique $\text{sign}P(r_0) = \text{sign}P''(r_0)$ y este r_0 será el valor inicial para comenzar la iteración de Newton-Raphson.

Ya se ha visto en el apartado anterior que $P'(r) < 0$ para todo r . Como $P''(r) = 3888 - 810r > 0$ para todo $r \in [0, 4]$ se tiene $\text{sign}P(0) = \text{sign}P''(0)$, de manera que se tomará valor inicial $r_0 = 0$.

Realizando la iteración de Newton-Raphson

$$r_{n+1} = r_n - \frac{P(r_n)}{P'(r_n)} = r_n - \frac{13888 - 9720r_n + 1944r_n^2 - 135r_n^3}{-9720 + 3888r_n - 405r_n^2}$$

con el valor inicial $r_0 = 0$ que asegura la convergencia hacia la única raíz de P en el intervalo $[0, 4]$ se obtiene la siguiente sucesión.

n	r_n
0	0
1	1.42881
2	2.14499
3	2.34559

Se tiene pues la aproximación $r^* \approx r_3 = 2,34559$ del radio de la capa con deformación nula.

Problema 8.6 Una partícula de masa m puede deslizarse sin rozamiento a lo largo de un alambre. El alambre gira entorno al origen, en un plano vertical y a una velocidad angular constante w . Tomando en el alambre la coordenada ξ para el movimiento de la partícula se puede demostrar que, si las condiciones iniciales vienen dadas por el alambre en posición horizontal y la partícula en reposo respecto del alambre (es decir $\xi(0) = 0$ y $\dot{\xi}(0) = 0$), entonces la posición de la partícula sobre el alambre en función del tiempo t es

$$\xi(t) = \frac{g}{2w^2} [\cosh(wt) - \cos(wt)] ,$$

siendo $g = 9.8$ la aceleración de la gravedad. Suponed que $w = 3$.

- (i) Demostrar que antes de $t = 1$ la partícula alcanza una única vez la posición $\xi = 4$.
- (ii) Sea t^* el tiempo para el cual la partícula alcanza la posición $\xi = 4$. Aproximar t^* mediante dos iteraciones con el método de la secante.

Solución. (i) Para demostrar que antes de $t = 1$ la partícula alcanza al menos una vez la posición $\xi = 4$ hemos de ver que existe al menos un tiempo $t^* \in (0, 1)$ tal que $\xi(t^*) = 4$. Para demostrarlo, utilizaremos el Teorema de Bolzano aplicado a la función $f(t) := \xi(t) - 4$. En concreto, $f \in C[0, 1]$ y además $f(0) = \xi(0) - 4 = -4 < 0$, $f(1) = \xi(1) - 4 = 6.02028 - 4 > 0$. Por lo tanto existe al menos un tiempo $t^* \in (0, 1)$ tal que $f(t^*) = 0$ o equivalentemente $\xi(t^*) = 4$.

Finalmente, para ver que t^* es único estudiaremos el crecimiento o decrecimiento de la función $f(t)$ en el intervalo $(0, 1)$. En concreto

$$f'(t) = \frac{g}{2w} [\sinh(wt) + \sin(wt)] = \frac{9.8}{6} [\sinh(3t) + \sin(3t)] .$$

Puesto que $\sinh(3t) > 0$ para todo $t > 0$ y $\sin(3t) > 0$ para $t \in (0, 1)$ se concluye que $f'(t) > 0$ para $t \in (0, 1)$ y por lo tanto $f(t)$ es una función creciente en el intervalo $(0, 1)$. En definitiva esto muestra que t^* es único.

(ii) Aproximaremos el cero t^* de la función f mediante el método de la secante, es decir, a través de la iteración

$$t_{i+1} = t_i - f(t_i) \frac{t_i - t_{i-1}}{f(t_i) - f(t_{i-1})} ,$$

donde, para comenzar la iteración, tomaremos los valores $t_0 = 0$ y $t_1 = 1$ debido a que sabemos que del apartado anterior que $t^* \in (0, 1)$. La tabla generada es la siguiente

i	t_i	$f(t_i)$
0	0	-4
1	1	2.02028
2	0.664421	-1.74172
3	0.819787	-0.369878

Por lo tanto, $t^* \approx t_3 = 0.819787$.

Problema 8.7 Una cadena uniforme, de longitud ℓ , está colocada sobre una tabla horizontal, libre de fricción, de tal manera que una longitud b de la cadena cae por el borde. Es fácil demostrar que el tiempo T que tardará la cadena en deslizarse completamente hacia abajo viene dado por

$$T = \sqrt{\frac{\ell}{g}} \ln \left(\frac{\ell + \sqrt{\ell^2 - b^2}}{b} \right), \quad (8.5)$$

siendo $g = 9,8$ la intensidad del campo gravitatorio.

- (i) Se han realizado las mediciones $\ell = 5 \pm 0,1$ y $b = 2 \pm 0,5$. Aproximar la precisión con la que se puede calcular el tiempo T .
- (ii) ¿Es la ecuación (8.5) trascendente si su incógnita es la longitud ℓ ? ¿Y si la incógnita es la longitud b ?
- (iii) Si $T = 15$ y $b = 5$, demostrar que existe al menos una raíz ℓ^* positiva de la ecuación (8.5). Calcular el mínimo número n de iteraciones necesarias en el algoritmo de bisección para aproximar ℓ^* con un error absoluto menor que 10^{-5} , partiendo del intervalo $[130, 140]$.

Solución. (i) Utilizando la fórmula de propagación de errores absolutos se tiene que

$$\Delta_T \approx \left| \frac{\partial T}{\partial \ell} \right| \Delta_\ell + \left| \frac{\partial T}{\partial b} \right| \Delta_b.$$

Calculamos las derivadas parciales

$$\frac{\partial T}{\partial \ell} = \frac{1}{2} \sqrt{\frac{\ell}{g}} \left(\frac{2}{\sqrt{\ell^2 - b^2}} + \frac{\ln \left(\frac{\ell + \sqrt{\ell^2 - b^2}}{b} \right)}{\ell} \right), \quad \frac{\partial T}{\partial b} = -\frac{\ell \sqrt{\ell/g}}{b \sqrt{\ell^2 - b^2}},$$

de manera que, sustituyendo los datos del enunciado se obtiene

$$\begin{aligned} \Delta_T &\approx \left| \frac{1}{2} \sqrt{\frac{5}{9,8}} \left(\frac{2}{\sqrt{5^2 - 2^2}} + \frac{\ln \left(\frac{5 + \sqrt{5^2 - 2^2}}{2} \right)}{5} \right) \right| \Delta_\ell + \left| -\frac{5 \sqrt{5/9,8}}{2 \sqrt{5^2 - 2^2}} \right| \Delta_b \\ &= |0,267784|0,1 + |-0,389675|0,5 = 0,221616. \end{aligned}$$

(ii) La ecuación (8.5) es trascendente respecto de la variable ℓ puesto que no es posible despejarla de la ecuación (8.5) mediante funciones elementales.

Sin embargo, la ecuación (8.5) no es trascendente respecto de la variable b . De hecho es fácil despejar de tal ecuación para obtener

$$b = \frac{2\ell \exp\left(\frac{T}{\sqrt{\ell/g}}\right)}{1 + \exp\left(\frac{2T}{\sqrt{\ell/g}}\right)} = \ell \operatorname{sech}\left(\frac{\sqrt{g}T}{\sqrt{\ell}}\right).$$

(iii) Tomando $T = 15$, $b = 5$ y pasando todos los términos de la ecuación (8.5) a la izquierda, dicha ecuación adopta la forma $f(\ell) = 0$, siendo

$$f(\ell) = 15 - \sqrt{\frac{\ell}{9,8}} \ln\left(\frac{\ell + \sqrt{\ell^2 - 25}}{5}\right).$$

Demstrar que existe al menos una raíz ℓ^* positiva de la ecuación (8.5) es equivalente a demostrar que existe $\ell^* > 0$ tal que $f(\ell^*) = 0$. Para ello nos basaremos en el Teorema de Bolzano. En primer lugar, observemos que, la expresión del tiempo T dada por (8.5) sólo tiene sentido físico si $\ell \geq b > 0$. Matemáticamente observamos que si no se verifica la desigualdad $\ell \geq b$, debido a la raíz cuadrada que aparece en (8.5), la función T no es real.

De esta forma, sólo tiene sentido que estudiemos la función $f(\ell)$ para $\ell > 5$. Además, bajo esta restricción, los únicos problemas de continuidad de la función f , es decir, la raíz cuadrada y el logaritmo, desaparecen ya que sus argumentos son mayor o igual que cero y positivo respectivamente. En definitiva la función $f \in C[5, \infty)$. En particular $f \in C[5, 150]$ y verifica $f(5)f(150) = 15 \times (-1,01723) < 0$, de manera que, aplicando el Teorema de Bolzano concluimos que existe $\ell^* \in (5, 150)$ tal que $f(\ell^*) = 0$.

Se sabe que el error absoluto Δ_n cometido en la aproximación $\ell^* \approx c_n$ siendo n el número de iteraciones realizadas con el algoritmo de bisección está acotado por

$$\Delta_n := |\ell^* - c_n| \leq \frac{|b_0 - a_0|}{2^{n+1}},$$

siendo $[a_0, b_0]$ el intervalo inicial de búsqueda. Así, se tiene que

$$\Delta_n \leq \frac{10}{2^{n+1}} < 10^{-5},$$

de donde obtenemos que

$$n > \log_2 10^{-6} - 1 = \frac{\log 10^6}{\log 2} - 1 = \frac{6}{\log 2} - 1 = 18,93.$$

Se puede asegurar entonces que, con 19 iteraciones del método de bisección obtenemos la longitud ℓ^* con un error menor que 10^{-5} .

Problema 8.8 Consideremos una partícula de masa m con un grado de libertad sometida a un campo de fuerzas conservativo con energía potencial asociada $U(x)$. A partir del teorema de la conservación de la energía mecánica E se tiene

$$\frac{1}{2}m\dot{x}^2 + U(x) = E. \quad (8.6)$$

Puesto que la energía cinética es una magnitud semidefinida positiva, el movimiento de la partícula tiene lugar únicamente en las regiones del espacio para las cuales se verifique $U(x) \leq E$. Los puntos \tilde{x}_i en los que se verifique $U(\tilde{x}_i) = E$ determinan los límites del movimiento y son llamados puntos de retorno puesto que allí la velocidad es nula. Sean \tilde{x}_1 y \tilde{x}_2 dos puntos de retorno tales que $U(x) < E$ para todo $x \in (\tilde{x}_1, \tilde{x}_2)$. Entonces, la partícula realiza un movimiento oscilatorio cuyo periodo T es el doble del tiempo necesario para ir de \tilde{x}_1 a \tilde{x}_2 , es decir, separando las variables en (8.6) e integrando se obtiene

$$T = \sqrt{2m} \int_{\tilde{x}_1}^{\tilde{x}_2} \frac{dx}{\sqrt{E - U(x)}}. \quad (8.7)$$

- (i) Si la partícula está sometida a un campo cuya energía potencial es $U(x) = Ax^n$, siendo A una constante real no nula y $n \in \mathbb{N}$ un número par, es fácil demostrar que el periodo T de oscilación de su movimiento dado a partir de (8.7) no se puede expresar, en general, mediante funciones elementales.
- (i.1) Suponiendo que $A = 1$, $n = 4$, $m = 2\text{Kg}$ y $E = 16\text{J}$, aproximar el valor del periodo T mediante la fórmula de integración Gauss-Legendre con 2 puntos. INDICACIÓN: Los polinomios de Legendre $P_k(x)$ de grado k son $P_k(x) = \frac{1}{k!2^k} \frac{d^k}{dx^k} (x^2 - 1)^k$ y forman una sucesión ortogonal en el intervalo $[-1, 1]$ respecto de la función peso $w(x) = 1$.
- (i.2) Suponiendo que $A = 1$, $n = 2$, $m = 2\text{Kg}$ y $E = 16\text{J}$, calcular el mínimo número de puntos N necesarios de manera que se pueda calcular de forma exacta el periodo T utilizando la fórmula de integración Gauss-Chebyshev con N puntos. Calcular con dicha fórmula el periodo T .
- (ii) Supongamos que la partícula tiene una energía potencial $U(x) = 5x - \ln x^3$ con $x > 0$ y una energía mecánica $E = 5\text{J}$.
- (ii.1) Es claro que $\tilde{x}_2 = 1$ es un punto de retorno. Demostrar que existe otro único punto de retorno \tilde{x}_1 con $0 < \tilde{x}_1 < 1$.
- (ii.2) Utilizando el método de Newton-Raphson, aproximar el anterior punto de retorno \tilde{x}_1 mediante tres iteraciones y partiendo de un punto inicial adecuado que asegure la convergencia hacia \tilde{x}_1 .
- (ii.3) Explicar detalladamente qué problema se encuentra si se desea aproximar el periodo T dado por (8.7) mediante una fórmula de Newton-Cotes.
- (ii.4) A partir de la ecuación diferencial (8.6), sabiendo que la partícula tiene una masa $m = 2\text{Kg}$ y que inicialmente se encuentra en el punto $x(0) = 0,9$ con una velocidad negativa, aproximar el valor $x(0,1)$ de la posición para tiempo $t = 0,1$ mediante un paso con el algoritmo de Runge-Kutta de orden 4.

Solución. (i.1) Si la partícula de masa $m = 2$ tiene una energía mecánica $E = 16$ y está sometida a una energía potencial $U(x) = x^4$, sus puntos de retorno \tilde{x}_i serán soluciones de la ecuación $U(x) = E$, es decir, $\tilde{x}_i^4 = 16$. Se tienen pues dos puntos de retorno $\tilde{x}_1 = -2$ y $\tilde{x}_2 = 2$. De esta forma, el periodo T de la partícula confinada a moverse en la región del espacio $-2 \leq x \leq 2$ viene dado por

$$T = 2 \int_{-2}^2 f(x) dx, \quad f(x) = \frac{1}{\sqrt{16 - x^4}}. \quad (8.8)$$

Un estudio más detallado sobre la integral que aparece en (8.8) muestra que $T = 2\sqrt{\pi}\Gamma(5/4)/\Gamma(3/4)$ siendo Γ la *función gamma de Euler* definida como $\Gamma(z) = \int_0^\infty t^{z-1} \exp(-t) dt$. Sin embargo en este problema se realizará un estudio numérico del valor del periodo T a partir de la fórmula de integración de Gauss-Legendre con 2 puntos, es decir,

$$\int_{-1}^1 g(x) dx \approx \sum_{k=0}^n \alpha_k g(x_k), \quad (8.9)$$

con $n = 1$ y siendo x_0 y x_1 las dos raíces del polinomio de Legendre de segundo grado $P_2(x)$. A partir de la fórmula generadora de los polinomios de Legendre, se tiene que

$$P_2(x) = \frac{1}{2!2^2} \frac{d^2}{dx^2} (x^2 - 1)^2 = \frac{1}{2} (3x^2 - 1),$$

de modo que $x_0 = -1/\sqrt{3}$, $x_1 = 1/\sqrt{3}$. Los coeficientes de la fórmula de Gauss-Legendre (8.9) con $n = 1$ se obtendrán de imponer que (8.9) tenga grado de precisión n , es decir que sea exacta para cualquier polinomio g de grado menor o igual que $n = 1$. Así, según el método de los coeficientes indeterminados y tomando la base canónica $\{1, x\}$ del espacio vectorial $\mathbb{R}_1[x]$ de los polinomios de grado menor o igual que 1, se tiene que α_0 y α_1 verifican el sistema lineal de ecuaciones

$$\int_{-1}^1 dx = \alpha_0 + \alpha_1, \quad \int_{-1}^1 x dx = \alpha_0 x_0 + \alpha_1 x_1,$$

es decir

$$2 = \alpha_0 + \alpha_1, \quad 0 = \frac{1}{\sqrt{3}}(-\alpha_0 + \alpha_1),$$

cuya solución es $\alpha_0 = \alpha_1 = 1$. En definitiva, la fórmula de Gauss-Legendre con 2 puntos es

$$\int_{-1}^1 g(x) dx \approx g\left(\frac{-1}{\sqrt{3}}\right) + g\left(\frac{1}{\sqrt{3}}\right), \quad (8.10)$$

y tiene un grado de precisión $2n + 1 = 3$. Obviamente, para poder aproximar mediante la fórmula (8.10) el periodo T dado por (8.8), previamente se ha de realizar un cambio de variable que lleve el intervalo $x \in [-2, 2]$ al intervalo $y \in [-1, 1]$. El cambio lineal que verifica lo anterior es $y = x/2$. En la nueva variable, (8.8) se escribe como

$$T = 4 \int_{-1}^1 f(2y) dy = \int_{-1}^1 g(y) dy, \quad g(y) = \frac{1}{\sqrt{1 - y^4}}.$$

Aproximando el anterior valor de T mediante (8.10) se tiene

$$T \approx g\left(\frac{-1}{\sqrt{3}}\right) + g\left(\frac{1}{\sqrt{3}}\right) = \frac{3}{2\sqrt{2}} + \frac{3}{2\sqrt{2}} = \frac{3}{\sqrt{2}} \approx 2,12132.$$

(i.2) Ahora la partícula de masa $m = 2$ tiene una energía mecánica $E = 16$ y una energía potencial $U(x) = x^2$. Nótese que la partícula realiza entonces un movimiento armónico simple. Los puntos de retorno \tilde{x}_i son las soluciones de la ecuación $\tilde{x}_i^2 = 16$ y por lo tanto $\tilde{x}_1 = -4$ y $\tilde{x}_2 = 4$. En definitiva, el periodo T del oscilador armónico es

$$T = 2 \int_{-4}^4 \frac{dx}{\sqrt{16 - x^2}}. \quad (8.11)$$

Realizamos el cambio de variable $x = 4y$, de manera que (8.11) se reescribe como una integral definida en el intervalo $[-1, 1]$ de la forma

$$T = 2 \int_{-1}^1 \frac{dy}{\sqrt{1 - y^2}}, \quad (8.12)$$

pudiéndose calcular de forma exacta, es decir, $T = 2[\arcsin y]_{-1}^1 = 2\pi$. Obsérvese que el valor obtenido concuerda con el periodo $T = 2\pi/w$ de un oscilador armónico simple siendo su energía potencial $U(x) = \frac{1}{2}mw^2x$. Sin embargo, se pretende calcular de forma exacta el valor de T dado por (8.12) a través de la fórmula de Gauss-Chebyshev con N puntos. Recordemos que dicha fórmula es

$$\int_{-1}^1 \frac{g(y)}{\sqrt{1 - y^2}} dy \approx \frac{\pi}{n+1} \sum_{k=0}^n g(y_k),$$

siendo $n = N - 1$ y y_k las abscisas de Chebyshev. Además tiene un grado de precisión $2n + 1 = 2N - 1$, es decir, la anterior fórmula de Gauss-Chebyshev es exacta para cualquier polinomio g de grado menor o igual que $2N - 1$. Como la integral que aparece en (8.12) tiene $g(y) = 1$ y este es un polinomio de grado cero, se debe imponer $2N - 1 = 0$, es decir utilizar la fórmula de Gauss-Chebyshev con $N = 1$ punto.

(ii.1) Los puntos de retorno \tilde{x}_i son las soluciones de la ecuación $U(x) = E$. Definiendo la función $f(x) = U(x) - E$, los puntos de retorno \tilde{x}_i son los ceros de la función $f(x)$ con $x > 0$. Puesto que $E = 5$ y $U(x) = 5x - \ln x^3$, se tiene $f(x) = 5(x - 1) - \ln x^3$. Es obvio que $\tilde{x}_2 = 1$ es un cero de f . Se ha de demostrar que f tiene un único cero \tilde{x}_1 verificando $0 < \tilde{x}_1 < 1$.

Calculando la primera y segunda derivada de f se obtiene $f'(x) = 5 - 3/x$ y $f''(x) = 3/x^2$. Ya que $f'(3/5) = 0$ y además $f''(3/5) > 0$ se comprueba que f tiene un único mínimo relativo en $x = 3/5 = 0,6$ de manera que $f'(x) < 0$ para todo $x \in (0, 0,6)$ y $f'(x) > 0$ para todo $x \in (0,6, \infty)$. Además $f \in C(0, \infty)$, $f(0,6) < 0$, $\lim_{x \rightarrow 0^+} f(x) = \infty$ y $f(1) = 0$, de manera que, ver Figura 8.2, se ha demostrado que f tiene un único cero $\tilde{x}_1 \in (0, 1)$.

Otra forma de proceder es la siguiente. Utilizando el Teorema de Bolzano se puede ver que f tiene al menos un cero $\tilde{x}_1 \in (0, 1)$. Efectivamente, puesto que

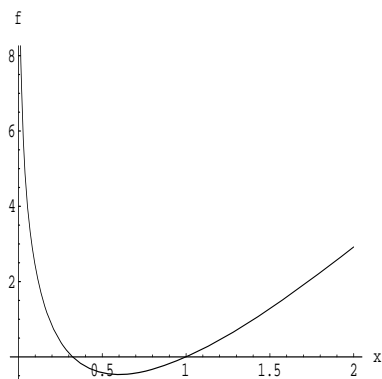


Figura 8.2: Gráfica de $f(x) = 5(x-1) - \ln x^3$.

f es una función continua para todo $x > 0$, en particular $f \in C[0,1,0,5]$ por ejemplo. Además $f(0,1)f(0,5) = f(0,1)f(0,5) < 0$. Entonces, se concluye que f tiene al menos un cero $\tilde{x}_1 \in [0,1,0,5]$. Para ver que dicho cero es único en este intervalo la función f es decreciente y por lo tanto f tiene un único cero en $\tilde{x}_1 \in [0,1,0,5]$.

(ii.2) Una vez se sabe que f tiene un único cero $\tilde{x}_1 \in [0,1,0,5]$, se utilizará el método de Newton-Raphson con el objetivo de aproximarlos, es decir

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{5(x_n - 1) - \ln x_n^3}{5 - 3/x_n}. \quad (8.13)$$

Como $f \in C^2[0,1,0,5]$, $f(0,1)f(0,5) < 0$ y tanto f' como f'' tienen signo constante en el intervalo $[0,1,0,5]$, se tomará, con el objetivo de asegurar la convergencia de la sucesión $\{x_n\}_{n \in \mathbb{N}}$ hacia \tilde{x}_1 , el punto inicial x_0 de la iteración de Newton-Raphson verificando $x_0 \in \{0,1,0,5\}$ tal que $\text{sign}(f(x_0)) = \text{sign}(f''(x_0))$. Es evidente que el valor inicial es $x_0 = 0,1$. De esta forma, la sucesión $\{x_n\}_{n \in \mathbb{N}}$ generada mediante la iteración de Newton-Raphson (8.13) con el valor inicial $x_0 = 0,1$ tiene los siguientes primeros términos de la sucesión.

n	x_n
0	0.1
1	0.19631
2	0.280509
3	0.318437

Se concluye que el punto de retorno \tilde{x}_1 es aproximado por $\tilde{x}_1 \approx x_3 = 0,318437$.

(ii.3) El periodo T dado por la fórmula integral (8.7) es una integral impropia de segunda especie ya que la función integrando

$$f(x) = \frac{1}{\sqrt{E - U(x)}} ,$$

no está acotada en el intervalo de integración $[\tilde{x}_1, \tilde{x}_2]$. Esto es debido a que $U(\tilde{x}_1) = U(\tilde{x}_2) = E$ y, como $E > U(x)$ para todo $x \in (\tilde{x}_1, \tilde{x}_2)$, se tiene

$$\lim_{x \rightarrow \tilde{x}_1^+} \frac{1}{\sqrt{E - U(x)}} = \lim_{x \rightarrow \tilde{x}_2^-} \frac{1}{\sqrt{E - U(x)}} = \infty .$$

Puesto que cualquier fórmula de Newton-Cotes evaluará la función integrando f en los dos extremos del intervalo de integración $[\tilde{x}_1, \tilde{x}_2]$, el resultado que nos proporcionaría es siempre $T = \infty$.

(ii.4) La ecuación diferencial (8.6) se puede escribir en la forma habitual

$$\dot{x} = f(t, x) , \quad \text{siendo } f(t, x) = \pm \sqrt{\frac{2(E - U(x))}{m}} .$$

Con los datos del problema se tiene el problema de Cauchy

$$\dot{x} = f(t, x) , \quad x(0) = 0,9 ,$$

siendo $f(t, x) = -\sqrt{5(1-x) + \ln x^3}$. Obsérvese que se ha utilizado la rama negativa de la raíz cuadrada debido a que el enunciado asegura que la velocidad \dot{x} es negativa. Utilizando el algoritmo RK4 para aproximar la solución del anterior problema de Cauchy se genera la sucesión $\{x_n\}_{n \in \mathbb{N}}$, siendo $x(t_i) \approx x_i$ para tiempos equiespaciados $t_i = t_0 + ih$, donde t_0 es el tiempo inicial y h una constante. En el caso que nos ocupa $t_0 = 0$ y $x_0 = 0,9$. Además, la ecuación diferencial es autónoma, es decir, la función f no depende del tiempo. Por lo tanto, el algoritmo RK4 aplicado a este problema consiste en calcular la sucesión definida de forma recursiva siguiente

$$\begin{aligned} x_{n+1} &= x_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4) , \\ k_1 &= f(x_n) , \\ k_2 &= f\left(x_n + \frac{h}{2}k_1\right) , \\ k_3 &= f\left(x_n + \frac{h}{2}k_2\right) , \\ k_4 &= f(x_n + hk_3) . \end{aligned}$$

Teniendo en cuenta que $f(x) = -\sqrt{5(1-x) + \ln x^3}$ y que se quiere aproximar $x(0,1)$, se tomará una longitud de paso es $h = 0,1$, de manera que con una iteración del algoritmo RK4 se obtiene

n	t_n	x_n
0	0	0.9
1	0.1	0.853091

En definitiva $x(0,1) \approx x_1 = 0,853091$.

Problema 8.9 Para la radiación de un cuerpo negro se sabe que, la energía radiada por unidad de volumen en el rango de longitudes de onda entre λ y $\lambda + d\lambda$ viene dada por $u(\lambda) d\lambda$, donde

$$u(\lambda) = \frac{8\pi}{\lambda^5} \frac{hc}{\exp\left(\frac{hc}{\lambda KT}\right) - 1},$$

siendo T la temperatura del cuerpo, c la velocidad de la luz, h la constante de Planck y K la constante de Boltzmann.

- (i) Demostrar que la longitud de onda λ_{\max} para la cual $u(\lambda)$ es máxima se puede escribir como $\lambda_{\max} = \alpha hc/(KT)$, siendo $\alpha \in \mathbb{R}$ una raíz de la ecuación trascendente

$$e^{1/\alpha}(5\alpha - 1) - 5\alpha = 0. \quad (8.14)$$

- (ii) Demostrar que la ecuación (8.14) tiene una única raíz en el intervalo $[0,15,0,25]$.

- (iii) Aproximar el valor de α mediante dos iteraciones del método de Newton-Raphson, partiendo de un punto inicial tal que nos asegure la convergencia hacia la raíz buscada.

- (iv) Supongamos que la temperatura del cuerpo negro es $T = 350$ grados Kelvin. Aproximar con qué exactitud se puede calcular la energía radiada por unidad de volumen u sabiendo que la longitud de onda vale $\lambda = 5 \times 10^{-6} \pm 10^{-7}$ metros. DATOS: En el sistema internacional $c = 3 \times 10^8$, $K = 1,381 \times 10^{-23}$ y $h = 6,626 \times 10^{-34}$.

Solución. (i) Puesto que $u(\lambda)$ es máxima para $\lambda = \lambda_{\max}$ se debe verificar que la derivada $du/d\lambda$ se anule para $\lambda = \lambda_{\max}$. La primera derivada de la función u respecto de λ es

$$\frac{du}{d\lambda} = \frac{8ch\pi \left[\exp\left(\frac{hc}{\lambda KT}\right) (hc - 5KT\lambda) + 5KT\lambda \right]}{KT\lambda^7 \left[\exp\left(\frac{hc}{\lambda KT}\right) - 1 \right]^2},$$

de modo que, sustituyendo en el numerador de esta expresión el valor de λ por el de $\lambda_{\max} = \alpha hc/(KT)$ e igualando a cero se obtiene $-8c^2h^2\pi f(\alpha) = 0$, siendo

$$f(\alpha) := e^{1/\alpha}(5\alpha - 1) - 5\alpha,$$

con lo cual queda demostrado.

(ii) La ecuación (8.14) se escribe de la forma $f(\alpha) = 0$. Puesto que $f \in C[0,15,0,25]$ y además $f(0,15)f(0,25) < 0$, por el Teorema de Bolzano se sabe que la ecuación (8.14) tiene al menos una raíz en el intervalo $[0,15,0,25]$.

Para ver que dicha raíz es única demostraremos que la función f es monótona en el intervalo $[0,15,0,25]$.

La primera y segunda derivada de f valen

$$\begin{aligned} f'(\alpha) &= \frac{1}{\alpha^2} \left[e^{1/\alpha} (1 - 5\alpha + 5\alpha^2) - 5\alpha^2 \right] , \\ f''(\alpha) &= \frac{1}{\alpha^4} e^{1/\alpha} (3\alpha - 1) . \end{aligned}$$

Puesto que $f''(\alpha) < 0$ en el intervalo $[0,15,0,25]$, esto significa que $f'(\alpha)$ es decreciente en el intervalo $[0,15,0,25]$. Como además $f' \in C[0,15,0,25]$, $f'(0,15) > 0$ y $f'(0,25) > 0$ se concluye que $f'(\alpha) > 0$ en el intervalo $[0,15,0,25]$ y, en consecuencia, $f(\alpha)$ es monótona creciente en el intervalo $[0,15,0,25]$.

(iii) La iteración de Newton-Raphson para aproximar la solución de la ecuación $f(\alpha) = 0$ viene dada por

$$\alpha_{n+1} = \alpha_n - \frac{f(\alpha_n)}{f'(\alpha_n)} , \quad n = 0, 1, 2, \dots$$

En nuestro caso se tiene

$$\alpha_{n+1} = \alpha_n - \frac{e^{1/\alpha_n} (5\alpha_n - 1) - 5\alpha_n}{\frac{1}{\alpha_n^2} [e^{1/\alpha_n} (1 - 5\alpha_n + 5\alpha_n^2) - 5\alpha_n^2]} .$$

Puesto que $f \in C^2[0,15,0,25]$ y se ha demostrado en el apartado anterior que $f(0,15) = -197,193 < 0$, $f(0,25) = 12,3995 > 0$, $f''(\alpha) < 0$ y $f'(\alpha) > 0$ para todo $\alpha \in [0,15,0,25]$, tomando el valor inicial de la iteración de Newton-Raphson $\alpha_0 = 0,15$ se asegura la convergencia de la sucesión $\{\alpha_n\}_{n \in \mathbb{N}}$ generada hacia el cero de f en el intervalo $[0,15,0,25]$. Esto es debido a que el signo de f y de f'' coinciden en α_0 .

Realizando la iteración con este valor inicial se tiene

n	α_n
0	0.15
1	0.165583
2	0.181035

(iv) Puesto que el único error se comete en la medida de la longitud de onda λ , mediante la fórmula de propagación de errores se tiene que el error absoluto Δ_u en el cálculo de u es aproximado por

$$\Delta_u \approx \left| \frac{du}{d\lambda} \right| \Delta_\lambda ,$$

siendo $\Delta_\lambda = 10^{-7}$. Sabemos del apartado (i) que

$$\frac{du}{d\lambda} = \frac{8ch\pi \left[\exp\left(\frac{hc}{\lambda KT}\right) (hc - 5KT\lambda) + 5KT\lambda \right]}{KT\lambda^7 \left[\exp\left(\frac{hc}{\lambda KT}\right) - 1 \right]^2},$$

de modo que, sustituyendo los datos del problema, se tiene

$$\Delta_u \approx |276461| \times 10^{-7} = 0,027646.$$

Problema 8.10 Consideremos la ecuación $4 - x - \tan x = 0$.

- (i) ¿Cuántas soluciones tiene dicha ecuación? ¿Es una ecuación trascendente?
- (ii) Demostrar que existe una única solución x^* de dicha ecuación perteneciente al intervalo $[1, 1.5]$.
- (iii) Averiguar el mínimo número de pasos a realizar con el algoritmo de bisección si se pretende aproximar x^* con un error menor que 10^{-1} . Dar dicha aproximación.

Solución. (i) Es fácil ver que la ecuación $f(x) = 0$, con $f(x) = 4 - x - \tan x$, tiene infinitas soluciones reales. Para verlo basta con notar que dicha ecuación se puede reescribir de la forma $g_1(x) = g_2(x)$, siendo $g_1(x) = 4 - x$ y $g_2(x) = \tan x$. Finalmente, dibujando las gráficas de las funciones g_1 y g_2 se comprueba que existen infinitos puntos de corte (siendo las abscisas de dichos puntos soluciones de la ecuación $f(x) = 0$). Ver la Figura 8.3.

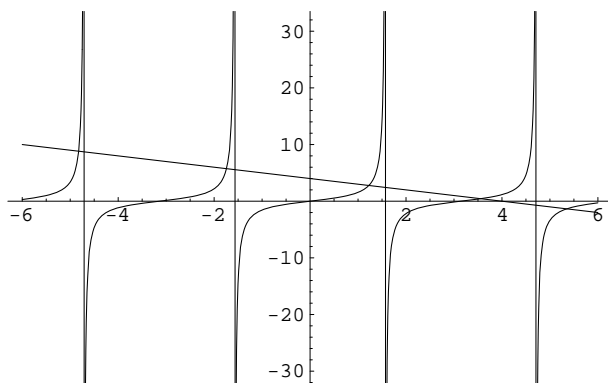


Figura 8.3: Gráfica superpuesta de las funciones $g_1(x) = 4 - x$ y $g_2(x) = \tan x$.

A pesar de tener infinitas soluciones, la ecuación $f(x) = 0$ es trascendente puesto que es imposible despejar la incógnita x , es decir, expresar x mediante funciones elementales a partir de la ecuación.

(ii) Notar que la función $\tan x$ (y por lo tanto la función f) es continua en todo \mathbb{R} excepto en los puntos $x_k = \pm(2k + 1)\pi/2$ con $k \in \mathbb{N} \cup \{0\}$. Puesto

que $x_k \notin [1, 1.5]$ para ningún k se concluye que la función $f(x)$ es continua en el intervalo $[1, 1.5]$. Como además $f(1)f(1.5) = 1,44259 \times (-11,6014) < 0$, aplicando el Teorema de Bolzano podemos asegurar que existe al menos una solución x^* de la ecuación $f(x) = 0$ perteneciente al intervalo $[1, 1.5]$.

Finalmente, mostraremos que x^* es única viendo la monotonicidad de la función f . Puesto que la función derivada de f es $f'(x) = -1 - 1/\cos^2 x < 0$ para todo $x \in [1, 1.5]$, se tiene que f es monótona decreciente en el intervalo $[1, 1.5]$ y, en consecuencia, existe una única solución x^* perteneciente al intervalo $[1, 1.5]$.

(iii) El método de bisección consiste en hallar una sucesión de intervalos anidados

$$(1, 1.5) = (a_0, b_0) \supset (a_1, b_1) \supset \cdots \supset (a_n, b_n) \supset \cdots$$

de manera que la anchura del siguiente intervalo es la mitad que la anchura del intervalo anterior y además x^* pertenezca a todos los intervalos. Entonces, la sucesión de puntos medios $\{c_n\}_{n=0}^{\infty}$ verifica $\lim_{n \rightarrow \infty} c_n = x^*$ con $|x^* - c_n| \leq (b_0 - a_0)/2^{n+1}$.

Queremos hallar n de modo que

$$|x^* - c_n| \leq \frac{b_0 - a_0}{2^{n+1}} = \frac{1/2}{2^{n+1}} < 10^{-1}.$$

De la anterior desigualdad se tiene $2^{n+2} > 10$, de donde, tomando logaritmos neperianos y aislando n obtenemos

$$n > \frac{\ln 10}{\ln 2} - 2 = 1.322.$$

Se tiene, en definitiva, que el mínimo número de iteraciones en el algoritmo de bisección aplicado a la ecuación $f(x) = 0$ y partiendo del intervalo inicial $[1, 1.5]$ es 2 si se pretende aproximar la solución x^* con un error menor que 10^{-1} .

Aplicamos el método de bisección 2 veces de modo que aproximaremos $x^* \approx c_2$, siendo la sucesión de puntos medios $\{c_n\}_{n=0}^{\infty}$ definida como $c_n = (a_n + b_n)/2$. En primer lugar calculamos $f(a_0) = f(1) = 1.442 > 0$ y $f(b_0) = f(1.5) = -11.601 < 0$.

- Ya que $f(c_0) = f(1.25) = -0.259 < 0$, el siguiente intervalo es $(a_1, b_1) = (a_0, c_0) = (1, 1.25)$.
- Como $f(c_1) = f(1.125) = 0.782 > 0$, el siguiente intervalo es $(a_2, b_2) = (c_1, b_1) = (1.125, 1.25)$, siendo su punto medio $c_2 = 1.187$.

Tenemos en resumen la siguiente tabulación

n	c_n	$f(c_n)$
0	1.25	-0.259
1	1.125	0.782
2	1.187	0.332

con la cual aproximamos $x^* \approx c_2 = 1,187$.

Problema 8.11 *Demostrar que la ecuación trascendente $e^x + \sin x = 2$ tiene una única solución x^* en el intervalo $[0, 1]$. Aproximar dicha solución utilizando 2 iteraciones del método de Newton-Raphson con una condición inicial adecuada que garantice la convergencia de la sucesión hacia x^* .*

Solución. Reescribimos la ecuación trascendente como $f(x) = e^x + \sin x - 2 = 0$. Observemos que $f \in C^\infty(\mathbb{R})$, de modo que, en particular $f \in C[0, 1]$. Puesto que además verifica $f(0)f(1) = (-1)(1,55975) < 0$, aplicando el Teorema de Bolzano se demuestra que la ecuación $f(x) = 0$ tiene al menos una solución $x^* \in (0, 1)$. Para ver que x^* es único basta con probar que f es una función monótona en el intervalo $[0, 1]$. Como $f'(x) = e^x + \cos x$, se tiene que $f'(x) > 0$ para todo $x > 0$ ya que $e^x > 1$ para todo $x > 0$ y $|\cos x| \leq 1$ para todo $x \in \mathbb{R}$. En particular $f'(x) > 0$ para todo $x \in [0, 1]$ con lo cual existe una única solución $x^* \in [0, 1]$ de la ecuación $f(x) = 0$.

Observemos que $f''(x) = e^x - \sin x$. Con un argumento similar al dado anteriormente se tiene $f''(x) > 0$ para todo $x \in [0, 1]$, de modo que se tomará el punto inicial x_0 de la iteración de Newton-Raphson como $x_0 \in \{0, 1\}$ tal que $f(x_0)f''(x_0) > 0$. En nuestro caso se tiene $x_0 = 1$.

Aproximaremos a continuación x^* utilizando 2 iteraciones del método de Newton-Raphson, es decir,

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} = x_i - \frac{e^{x_i} + \sin x_i - 2}{e^{x_i} + \cos x_i}, \quad x_0 = 1.$$

Se obtiene de esta forma la siguiente tabla

n	x_n
0	1
1	0.52134
2	0.44988

de modo que la aproximación deseada es $x^* \approx x_2 = 0,44988$.

Problema 8.12 *Consideremos un oscilador lineal con rozamiento forzado. Supongamos que su posición $x(t)$ en función del tiempo verifica la ecuación diferencial $\ddot{x} + \dot{x} + 2x = F(t)$, siendo la fuerza externa aplicada $F(t) = t \sin t$.*

- (i) *Demostrar que existe un tiempo t^* tal que $1 \leq t^* \leq 2$ para el cual la fuerza externa ejercida sobre el oscilador vale 1. Aproximar t^* mediante dos iteraciones del método de bisección.*
- (ii) *Sabiendo que inicialmente el oscilador se encuentra en reposo en la posición $x(0) = 5$, hallar una aproximación de la velocidad del oscilador para tiempo $t = 2$ mediante el método de Euler con una longitud de paso $h = 1$.*

Solución. (i) El tiempo $t^* \in [1, 2]$, si existe, verifica $F(t^*) = 1$. Definimos la función $f(t) = F(t) - 1$. Es claro por construcción que t^* , si existe, es un cero de la función $f(t)$. Veamos que efectivamente t^* existe. Para ello observamos que $f(t) = F(t) - 1 = t \sin t - 1$ es una función de clase $C^\infty(\mathbb{R})$, de modo que, en particular, $f \in C[1, 2]$. Además, $f(1)f(2) = -0,158 \times 0,818 < 0$ por lo tanto la función f cambia de signo en los extremos del intervalo $[1, 2]$. Se concluye que, aplicando el Teorema de Bolzano, existe al menos un tiempo $t^* \in [1, 2]$ tal que $f(t^*) = 0$, o equivalentemente, $F(t^*) = 1$.

Aplicando el método de bisección, generaremos una sucesión de intervalos encajados $[a_0, b_0] \supset [a_1, b_1] \supset [a_2, b_2]$ de modo que $t^* \in [a_i, b_i]$, siendo la anchura del siguiente intervalo la mitad que el del anterior. Finalmente aproximaremos $t^* \approx c_2$, siendo $c_i = (a_i + b_i)/2$ el punto medio del intervalo $[a_i, b_i]$.

Partimos del intervalo $[a_0, b_0] = [1, 2]$ con punto medio $c_0 = 1,5$. Puesto que $f(c_0) = f(1,5) = 0,496242 > 0$, esto significa que $t^* \in [a_0, c_0] := [a_1, b_1] = [1, 1,5]$. El punto medio $c_1 = (1 + 1,5)/2 = 1,25$, de modo que $f(c_1) = f(1,25) = 0,186231 > 0$. Entonces $t^* \in [a_1, c_1] := [a_2, b_2] = [1, 1,25]$. Finalizamos con la aproximación

$$t^* \approx c_2 = \frac{1 + 1,25}{2} = 1,125 .$$

(ii) Antes de aplicar el método de Euler, convertimos mediante el cambio de variable habitual $\dot{x} = y$ la ecuación diferencial de segundo orden del oscilador en un sistema de ecuaciones diferenciales de primer orden, es decir,

$$\begin{aligned} \dot{x} &= f(t, x, y) = y , \\ \dot{y} &= g(t, x, y) = -y - 2x + F(t) = -y - 2x + t \sin t . \end{aligned}$$

Según los datos del enunciado, inicialmente el oscilador se encuentra en reposo en la posición 5. Esto significa que el anterior sistema de ecuaciones diferenciales está acompañado de la condición inicial $(x(0), y(0)) = (5, 0)$ convirtiéndose de esta forma en un problema de Cauchy. Definiendo las aproximaciones $x(t_i) \approx x_i$, $y(t_i) \approx y_i$, con tiempos equiespaciados $t_i = ih = i$, el método de Euler

$$\begin{aligned} x_{i+1} &= x_i + hf(t_i, x_i, y_i), \\ y_{i+1} &= y_i + hg(t_i, x_i, y_i), \end{aligned}$$

genera la sucesión de puntos (x_i, y_i) con $i = 1, 2, \dots$ partiendo de la condición inicial $(x_0, y_0) = (5, 0)$. Aplicando esta iteración al oscilador del enunciado se tiene

$$\begin{aligned} x_{i+1} &= x_i + y_i, \\ y_{i+1} &= -2x_i + i \sin i, \end{aligned}$$

dando lugar a la siguiente tabla

i	t_i	x_i	y_i
0	0	5	0
1	1	5	-10
2	2	-5	-9.15853

Se concluye que la velocidad del oscilador para tiempo $t = 2$ es aproximadamente

$$\dot{x}(2) = y(2) \approx y_2 = -9,15853 .$$

Problema 8.13 *Los laureados con el premio Nobel de física E. Fermi, J. Pasta y S. Ulam descubrieron mediante experimentos numéricos con un ordenador que un sistema de osciladores no lineales acoplados no verificaban una propiedad de equilibrio termodinámico llamada equipartición de la energía. Posteriormente, M. Toda fue capaz de estudiar el problema de forma analítica mediante la introducción de un potencial de interacción del tipo*

$$V(r) = \frac{\alpha}{\beta} \exp(-\beta r) + \alpha r .$$

Tomando $\alpha = 2$ y $\beta = 1$:

- (i) *Demostrar que existe un único $r^* > 0$ verificando $V(r^*) = 4$. ¿Es posible hallar r^* de manera exacta?*
- (ii) *Aplicar 2 iteraciones del algoritmo de bisección con el objetivo de aproximar r^* . Iniciar el algoritmo en un intervalo del tipo $[m, m+1] \subset \mathbb{R}$ con $m \in \mathbb{N}$.*
- (iii) *Partiendo del intervalo de menor anchura obtenido en el apartado anterior tras 2 iteraciones del algoritmo de bisección, refinar la aproximación de r^* mediante 2 iteraciones del método de Newton-Raphson. Iniciar la iteración en un punto adecuado que asegure la convergencia hacia r^* de la sucesión de Newton-Raphson.*
- (iv) *Hallar una aproximación del error relativo que se comete al evaluar la función $V(r)$ cuando $r = 10 \pm 1$.*

Solución. (i) Puesto que $V(r^*) = 4$, es claro que r^* es un cero de la función

$$f(r) = V(r) - 4 = 2 \exp(-r) + 2r - 4 .$$

Calculando la primera derivada

$$f'(r) = 2(1 - \exp(-r)) ,$$

es inmediato comprobar que $f'(r) > 0$ para todo $r > 0$. Por lo tanto la función $f(r)$ es creciente en el intervalo $(0, \infty)$. Como además verifica que

$$f(0) = -2 < 0 , \quad \lim_{r \rightarrow \infty} f(r) = \infty > 0 ,$$

queda demostrado que exista un único $r^* > 0$ verificando $f(r^*) = 0$.

Por supuesto, la ecuación $f(r) = 0$ es trascendente y no se puede resolver de manera exacta.

(ii) Observemos en primer lugar que la función $f(r) \in C(\mathbb{R})$. Entonces, el intervalo $[m, m+1] \subset \mathbb{R}$ con $m \in \mathbb{N}$ para iniciar el algoritmo de bisección sólo debe de verificar la condición $f(m)f(m+1) < 0$. Se halla calculando el signo de $f(m)$ con $m = 0, 1, 2, \dots$ y parando cuando se observe un cambio de signo. En concreto se tiene $f(0) = -2 < 0$, $f(1) = -1,26424 < 0$ y $f(2) = 0,270671 > 0$ por lo tanto el intervalo inicial es $[a_0, b_0] = [m, m+1] = [1, 2]$. Iniciamos el algoritmo calculando el punto medio $c_0 = (1 + 2)/2 = 1,5$. Como $f(c_0) = f(1,5) = -0,55374 < 0$, entonces el nuevo intervalo es $[a_1, b_1] = [c_0, b_0] = [1,5, 2]$ y su punto medio $c_1 = (1,5 + 2)/2 = 1,75$. Como $f(c_1) = f(1,75) = -0,152452 < 0$, entonces el siguiente intervalo es $[a_2, b_2] = [c_1, b_1] = [1,75, 2]$ con punto medio $c_2 = 1,875$. En definitiva, $r^* \approx c_2 = 1,875$.

(iii) Según los resultados anteriores, $f \in C^2[1,75, 2]$, $f(1,75)f(2) < 0$ y $f'(r) > 0$ para todo $r > 0$. Además, como

$$f''(r) = 2 \exp(-r) ,$$

la función $f(r)$ tiene curvatura positiva en todo \mathbb{R} puesto que $f''(r) > 0$ para todo $r \in \mathbb{R}$. En definitiva, como el signo de $f(2)$ coincide con el signo de $f''(2)$, tomamos el punto inicial $r_0 = 2$ en la iteración de Newton-Raphson

$$r_{n+1} = r_n - \frac{f(r_n)}{f'(r_n)} = r_n - \frac{\exp(-r_n) + r_n - 2}{1 - \exp(-r_n)} .$$

De esta forma se obtiene la siguiente tabla

n	r_n
0	2
1	1.8434
2	1.8414

de donde se concluye que $r^* \approx r_2 = 1,8414$.

(iv) Como $r = 10 \pm 1$, se conoce r con un error relativo

$$\epsilon_r = \left| \frac{\Delta_r}{r} \right| = \left| \frac{1}{10} \right| = 0,1 .$$

Entonces, al evaluar $V(r) = 2(\exp(-r) + r)$ en $r = 10 \pm 1$, el error relativo cometido es

$$\begin{aligned} \epsilon_r &\approx \left| \frac{r}{V(r)} \frac{dV(r)}{dr} \right| \epsilon_r = \left| -\frac{r}{\exp(-r) + r} (-\exp(-r) + 1) \right| 0,1 \\ &= \left| -\frac{10}{\exp(-10) + 10} (-\exp(-10) + 1) \right| 0,1 = 0,099995 . \end{aligned}$$

Problema 8.14 Consideremos una partícula de masa m con un grado de libertad sometida a un campo de fuerzas conservativo con energía potencial asociada $U(x)$, siendo x la posición de la partícula. A partir del teorema de la conservación de la energía mecánica E y suponiendo que la partícula verifica la condición inicial $x(t_0) = x_0$, es fácil ver que

$$t - t_0 = \sqrt{\frac{m}{2}} \int_{x_0}^x \frac{ds}{\sqrt{E - U(s)}} , \quad (8.15)$$

siendo t el tiempo para el cual la posición es x .

- (i) Si $m = 2$, $E = 10$, $U(x) = x^4$ y la partícula se encuentra inicialmente en $x(0) = 0$, calcular la longitud de paso en la regla de los Trapecios compuesta si se pretende aproximar con ella el tiempo t que tarda la partícula en llegar por primera vez a la posición $x = 1$ con un error absoluto menor que 10^{-4} .
- (ii) Puesto que la energía cinética es una magnitud semidefinida positiva, el movimiento de la partícula tiene lugar únicamente en las regiones del espacio para las cuales se verifique $U(x) \leq E$. Los puntos en los que se verifique $U(x) = E$ determinan los límites del movimiento y son llamados puntos de retorno puesto que allí la velocidad es nula. Suponed ahora que $E = 0,4$ y $U(x) = x^2 e^x$.
- (ii.1) Demostrar que existe un único punto de retorno positivo.
- (ii.2) Aproximar el anterior punto de retorno mediante 3 iteraciones del método de bisección partiendo de un intervalo inicial de extremos formados por dos enteros consecutivos. Hallar además una cota del error absoluto cometido.

Solución. (i) Puesto que $m = 2$, $E = 10$, $U(x) = x^4$ y la partícula se encuentra inicialmente en $x(0) = 0$, la ecuación (8.15) adopta la forma

$$t = \int_0^x \frac{ds}{\sqrt{10 - s^4}} .$$

Entonces, el tiempo t que tarda la partícula en llegar por primera vez a la posición $x = 1$ es

$$t = \int_0^1 f(s) ds , \quad f(s) = \frac{1}{\sqrt{10 - s^4}} .$$

Recordemos, por otra parte, que la regla de los Trapecios compuesta con longitud de paso h es $\int_a^b f(s) ds \approx T_h(f)$ con un error absoluto $\Delta = \frac{h^2}{12}(b-a)|f''(\xi)|$, $\xi \in (a, b)$. En nuestro caso se tiene

$$\Delta = \frac{h^2}{12}|f''(\xi)| , \quad \xi \in (0, 1) .$$

Puesto que la función derivada segunda de f es

$$f''(\xi) = \frac{6\xi^2(10 + \xi^4)}{(10 - \xi^4)^{5/2}}$$

y además $\xi \in (0, 1)$, es obvio que $f''(\xi) > 0$. Además, al aumentar ξ aumenta el numerador y disminuye el denominador de $f''(\xi)$ con lo que vemos que la función $f''(\xi)$ es creciente¹ en el intervalo $\xi \in (0, 1)$. En definitiva, se puede acotar el error absoluto de la forma

$$\Delta < \frac{h^2}{12} f''(1) = \frac{h^2}{12} \frac{22}{81} = \frac{11}{486} h^2.$$

Si imponemos que la longitud de paso h verifique la desigualdad

$$\frac{11}{486} h^2 < 10^{-4}$$

nos aseguramos de que $\Delta < 10^{-4}$ como se desea. De esta forma se tiene que

$$h < \sqrt{10^{-4} \frac{486}{11}} = 0,066.$$

(ii) En primer lugar observamos que los puntos de retorno x^* son los ceros de la función $f(x) = U(x) - E$. En nuestro caso se tiene $f(x) = x^2 e^x - 0,4$.

(ii.1) Hemos de demostrar que la función $f(x)$ tiene un único cero positivo x^* . Para verlo, notemos que $f(0) < 0$ y $\lim_{x \rightarrow \infty} f(x) = \infty$ de modo que, como la función $f(x)$ es continua, esta tiene al menos un cero $x^* > 0$. Para ver que x^* es único, basta con demostrar que $f(x)$ es una función creciente en el intervalo $(0, \infty)$. Esto es cierto puesto que $f'(x) = x(2 + x)e^x > 0$ para todo $x > 0$.

(ii.2) Como $f \in C[0, 1]$, $f(0) = -0,4 < 0$ y $f(1) = 2,31828 > 0$, aplicando el Teorema de Bolzano sabemos que existe $x^* \in [0, 1]$ tal que $f(x^*) = 0$. De hecho, por el apartado anterior, sabemos que x^* es único. Aplicando 3 iteraciones en el método de bisección, obtendremos una colección de intervalos anidados $(a_0, b_0) = (0, 1) \supset (a_2, b_2) \supset (a_3, b_3)$ de anchura del siguiente la mitad que la del anterior tal que $[a_i, b_i] \ni x^* \approx c_3$, siendo c_i el punto medio del intervalo i -ésimo. Un simple cálculo nos muestra la siguiente tabla

i	c_i	$f(c_i)$
0	0.5	0.0121803
1	0.25	-0.319748
2	0.375	-0.195392
3	0.4375	-0.103544

¹Otra forma de ver que $f''(\xi)$ es creciente en el intervalo $\xi \in (0, 1)$ es calculando $f'''(\xi) = 24x(50 + 35x^4 + x^8)/(10 - x^4)^{7/2}$ y viendo que, efectivamente, $f'''(\xi) > 0$ para todo $\xi \in (0, 1)$.

Se concluye que $x^* \approx c_3 = 0,4375$.

Para hallar una cota del error absoluto Δ_3 cometido en la aproximación $x^* \approx c_3$ recordamos que, en el método de bisección se verifica que $\Delta_n := |x^* - c_n| \leq (b_0 - a_0)/2^{n+1}$. En nuestro caso se tiene

$$\Delta_3 = |x^* - c_3| \leq \frac{1}{2^4}.$$

Problema 8.15 Consideremos el siguiente sistema no lineal de ecuaciones dependiendo del parámetro a .

$$x^2 - y = a, \quad x^2 + 4y^2 = 4.$$

- (i) Hallar, mediante un método gráfico, el número de soluciones $(x^*, y^*) \in \mathbb{R}^2$ que tiene el sistema para los casos $a = 0$ y $a = -1$.
- (ii) Si $a = 0$, realizar dos iteraciones del método de Newton-Raphson partiendo del punto inicial $(x_0, y_0) = (1, 1)$.
- (iii) Si $a = -1$, explicar el problema que se tendrá al aplicar el método de Newton-Raphson para aproximar la solución del sistema.
- (iv) Si $a = 0$, transforma el sistema de ecuaciones en una única ecuación para la incógnita x . Demostrar, para esta ecuación, que existe una única solución $x^* \in (1/2, 1)$. Utilizar dos iteraciones del método de Newton-Raphson para aproximar x^* , partiendo de un punto inicial x_0 que asegure la convergencia del método.

Solución. (i) Notemos que, para cualquier valor del parámetro $a \in \mathbb{R}$, las soluciones del sistema

$$x^2 - y = a, \quad x^2 + 4y^2 = 4,$$

son representadas geoméricamente por los puntos de corte entre la parábola $y = x^2 - a$ y la cónica $x^2 + 4y^2 = 4$. Además, es fácil ver que la cónica es en realidad la elipse

$$\frac{x^2}{2^2} + \frac{y^2}{1^2} = 1,$$

centrada en el origen de coordenadas y de semiejes 2 y 1. Se tiene pues que las gráficas superpuestas de la parábola y la elipse viene dadas en las Figuras 8.4 y 8.5.

Concluimos que si $a = 0$ el sistema tiene dos soluciones reales y si $a = -1$ sólo tiene una solución que es justo $(x^*, y^*) = (0, 1)$.

(ii) Tomando $a = 0$, el sistema se puede escribir de la forma vectorial siguiente $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, siendo $\mathbf{x} = (x, y)^T$ y la aplicación $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ de finida de la forma

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} x^2 - y \\ x^2 + 4y^2 - 4 \end{pmatrix}.$$

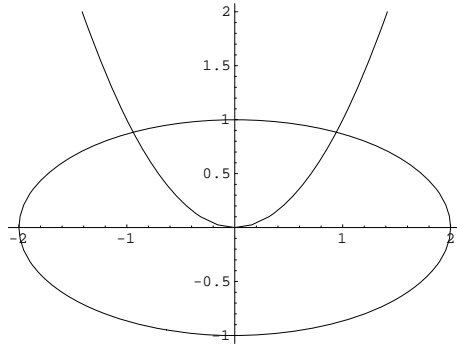


Figura 8.4: Gráfica de las cónicas en el plano $x - y$ con $a = 0$.

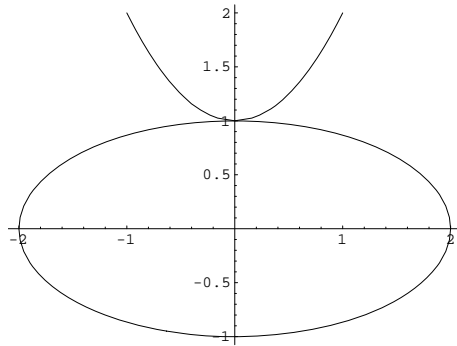


Figura 8.5: Gráfica de las cónicas en el plano $x - y$ con $a = -1$.

Sea \mathbf{x}^* un cero de la función \mathbf{f} , es decir $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$. El método de Newton-Raphson para aproximar las soluciones del sistema $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ consiste en realizar la iteración vectorial

$$\mathbf{x}_{i+1} = \mathbf{x}_i - (D\mathbf{f}(\mathbf{x}_i))^{-1}\mathbf{f}(\mathbf{x}_i) \quad \text{con } i = 0, 1, \dots, \quad (8.16)$$

siendo $D\mathbf{f}(\mathbf{x}_i)$ la matriz Jacobiana de la función \mathbf{f} en el punto $\mathbf{x}_i = (x_i, y_i) \in \mathbb{R}^2$, es decir

$$D\mathbf{f}(\mathbf{x}_i) = \begin{pmatrix} \frac{\partial f_1(\mathbf{x}_i)}{\partial x} & \frac{\partial f_1(\mathbf{x}_i)}{\partial y} \\ \frac{\partial f_2(\mathbf{x}_i)}{\partial x} & \frac{\partial f_2(\mathbf{x}_i)}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x_i & -1 \\ 2x_i & 8y_i \end{pmatrix}.$$

Si la matriz Jacobiana $D\mathbf{f}(\mathbf{x}_i)$ es no singular, es decir si $\det D\mathbf{f}(\mathbf{x}_i) = 2x_i(1 + 8y_i) \neq 0$, entonces se puede realizar la iteración (8.16). En concreto, se tiene

$$\begin{pmatrix} x_{i+1} \\ y_{i+1} \end{pmatrix} = \begin{pmatrix} x_i \\ y_i \end{pmatrix} - \frac{1}{2x_i(1 + 8y_i)} \begin{pmatrix} 8y_i & 1 \\ -2x_i & 2x_i \end{pmatrix} \begin{pmatrix} x_i^2 - y_i \\ x_i^2 + 4y_i^2 - 4 \end{pmatrix},$$

partiendo del punto inicial $(x_0, y_0) = (1, 1)$. Obtenemos la siguiente tabla a partir de la anterior iteración

i	x_i	y_i
0	1	1
1	0.9444	0.8888
2	0.9395	0.8828

La sucesión de puntos (x_i, y_i) parece converger hacia la solución (x^*, y^*) del primer cuadrante.

(iii) Si $a = -1$ y procedemos de igual forma que en el apartado anterior, se tiene que el sistema se puede escribir de la forma vectorial siguiente $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, siendo

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} x^2 - y + 1 \\ x^2 + 4y^2 - 4 \end{pmatrix}.$$

Para aplicar el método de Newton-Raphson, la matriz Jacobiana es la misma que en el anterior apartado

$$D\mathbf{f}(\mathbf{x}_i) = \begin{pmatrix} 2x_i & -1 \\ 2x_i & 8y_i \end{pmatrix},$$

y sabemos que no es invertible si $\det D\mathbf{f}(\mathbf{x}_i) = 2x_i(1 + 8y_i) = 0$. Puesto que, con $a = -1$, el sistema sólo tiene una solución que es justo $(x^*, y^*) = (0, 1)$, es claro que $\det D\mathbf{f}(0, 1) = 0$, de modo que, la sucesión de puntos que se genera con el método de Newton-Raphson tendería, en caso de converger, hacia un punto donde la matriz $D\mathbf{f}$ no es invertible. Esto originará problemas numéricos a la hora de calcular la matriz inversa de $D\mathbf{f}$ en puntos suficientemente cercanos a la solución $(0, 1)$.

NOTA: Geométricamente ya se ve que la solución $(0, 1)$ del sistema con $a = -1$ presentará problemas con el método de Newton-Raphson puesto que en ese punto la intersección de las dos cónicas (parábola $f_1(x, y) = 0$ y elipse $f_2(x, y) = 0$) no es transversal si no que es tangencial. Entonces, los vectores gradiente $\nabla f_1(0, 1)$ y $\nabla f_2(0, 1)$ son proporcionales y en consecuencia, el determinante de la matriz Jacobiana $\det D\mathbf{f}(0, 1) = 0$.

(iv) Volvemos al caso $a = 0$. Entonces, despejando la incógnita y de la primera ecuación y sustituyéndola en la segunda ecuación se obtiene una única ecuación para la incógnita x . En concreto se llega a que x verifica la ecuación

$$0 = f(x) = -4 + x^2 + 4x^4.$$

Puesto que $f(x)$ es un polinomio, es claro que $f \in C^\infty(\mathbb{R})$. En particular, $f \in C([1/2, 1])$ y $f(1/2)f(1) = (-7/2) \times 1 < 0$ de modo que, aplicando el Teorema de Bolzano, se tiene que existe al menos un cero x^* de f con $x^* \in (1/2, 1)$. Además, como $f'(x) = 2x(1 + 8x^2)$ se tiene $f'(x) > 0$ para todo $x > 0$. Entonces, $f(x)$ es una función creciente en el intervalo $x \in (1/2, 1)$ de modo que el cero $x^* \in (1/2, 1)$ de f es único.

Utilizaremos a continuación dos iteraciones del método de Newton-Raphson para aproximar x^* , partiendo de un punto inicial x_0 que asegure la convergencia

del método. Para elegir x_0 , observemos que $f''(x) = 2(1 + 24x^2) > 0$ para todo $x \in \mathbb{R}$. En definitiva, aplicando el Teorema de convergencia de la sucesión de Newton–Raphson, se tiene que, puesto que i) $f \in C^2[1/2, 1]$; ii) $f(1/2)f(1) < 0$; iii) $f'(x) > 0$ y $f''(x) > 0$ para todo $x \in [1/2, 1]$, concluimos que una elección de x_0 que asegura la convergencia de la sucesión de Newton–Raphson hacia x^* es $x_0 = 1$ ya que el signo de $f(1)$ y de $f''(1)$ coinciden.

Realizaremos dos iteraciones del método de Newton–Raphson para aproximar x^* . La iteración es

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} = x_i - \frac{-4 + x_i^2 + 4x_i^4}{2x_i(1 + 8x_i^2)},$$

con condición inicial $x_0 = 1$. Obtenemos la siguiente tabla.

i	x_i	$f(x_i)$
0	1	1
1	0.9444	0.0737
2	0.9396	0.0005

Se tiene pues que $x^* \approx x_2 = 0,9396$.

Problema 8.16 Considerar la ecuación $e^x = x + 2$.

- (i) Demostrar gráficamente que esta ecuación tiene sólo dos soluciones $x_1^* > 0$, $x_2^* < 0$. Obtener dos intervalos entre dos enteros consecutivos $I_i = [m_i, m_i + 1] \subset \mathbb{R}$ con $m_i \in \mathbb{Z}$ para $i = 1, 2$, tal que $x_i^* \in I_i$.
- (ii) Calcular el número de pasos que se deben ejecutar en el algoritmo de bisección partiendo del intervalo I_1 si se desea aproximar x_1^* con un error absoluto menor que 10^{-2} .
- (iii) Utilizar 2 iteraciones del método de Newton–Raphson para aproximar x_2^* . Usar un punto inicial en la iteración de modo que quede garantizada la convergencia de la sucesión hacia x_2^* .

Solución. (i) Las soluciones de la ecuación $e^x = x + 2$ son las abscisas de los puntos de corte entre las gráficas de la función exponencial $y = e^x$ y la recta $y = x + 2$, ver la gráfica de la Figura 8.6.

Es claro que dichas gráficas sólo tienen 2 puntos de corte (x_1^*, y_1^*) y (x_2^*, y_2^*) con $x_1^* > 0$ y $x_2^* < 0$. Definiendo la función $f(x) = e^x - (x + 2)$ se verificará que $f(x_i^*) = 0$ para $i = 1, 2$. Además, si tomamos los intervalos $I_1 = [1, 2]$ y $I_2 = [-2, -1]$ se tiene que $f(1)f(2) = -0,281718 \times 3,38906 < 0$ y $f(-2)f(-1) = 0,135335 \times (-0,632121) < 0$. De este modo, teniendo en cuenta que $f \in C(\mathbb{R})$, aplicando el teorema de Bolzano a la función f en los intervalos I_1 e I_2 concluimos que $x_i^* \in I_i$ para $i = 1, 2$.

(ii) Sea n el número de pasos que se deben ejecutar en el algoritmo de bisección partiendo del intervalo $I_1 = [a_0, b_0] = [1, 2]$ para aproximar x_1^* con un

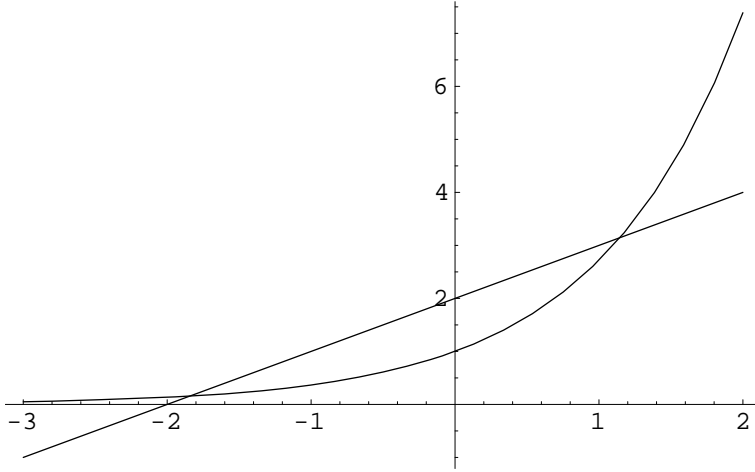


Figura 8.6: Superposición de las gráficas de las funciones $y = e^x$, $y = x + 2$.

error absoluto $\Delta_n < 10^{-2}$. Recordemos que el algoritmo de Bisección genera una sucesión de intervalos encajados de anchura del siguiente la mitad que la del anterior $(a_0, b_0) \supset (a_1, b_1) \supset \dots \supset (a_n, b_n) \supset \dots$ de modo que, la sucesión de puntos medios $\{c_n = (a_n + b_n)/2\}_{n=0}^{\infty}$ verifica $\lim_{n \rightarrow \infty} c_n = x_1^*$. De este modo, aproximando $x_1^* \approx c_n$, el error absoluto cometido es $\Delta_n = |x_1^* - c_n| \leq (b_0 - a_0)/2^{n+1}$. En nuestro caso se tiene $\Delta_n \leq (2 - 1)/2^{n+1} = 1/2^{n+1}$. Imponiendo pues que $\Delta_n < 10^{-2}$ se llega a una inequación para n de la forma

$$\frac{1}{2^{n+1}} < 10^{-2}.$$

Invirtiendo y tomando logaritmos neperianos se obtiene $(n + 1) \ln 2 > 2 \ln 10$, de donde

$$n > 2 \frac{\ln 10}{\ln 2} - 1 = 5,64386.$$

Puesto que $n \in \mathbb{N}$, realizaremos $n = 6$ iteraciones.

(iii) Sabemos que $x_2^* \in I_2 = [-2, -1]$ es solución de la ecuación $f(x) = 0$ siendo $f(x) = e^x - (x + 2)$. Queremos aproximar x_2^* mediante 2 iteraciones del algoritmo de Newton-Raphson

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

partiendo de un punto inicial x_0 adecuado que asegure la convergencia de la sucesión. Sabemos que $f \in C^2[-2, -1]$ y que $f(-2)f(-1) < 0$. Además, $f'(x) = e^x - 1$ y $f''(x) = e^x$, de modo que $f'(x)$ sólo se anula en $x = 0$ y $f''(x) > 0$ para todo $x \in \mathbb{R}$. En definitiva, para todo $x \in [-2, -1]$ se tiene que $f'(x) < 0$

y $f''(x) > 0$. Entonces, tomando $x_0 \in \{-2, -1\}$ que verifique $\text{sign}f(x_0) = \text{sign}f''(x_0)$ se asegura la convergencia de la sucesión de Newton–Raphson hacia la única raíz de f en $[-2, -1]$. En nuestro caso es claro que $x_0 = -2$. Realizando 2 iteraciones de Newton–Raphson se obtiene la siguiente tabla

n	x_n	$f(x_n)$
0	-2	0.135335
1	-1.84348	0.00174769
2	-1.84141	3.40602×10^{-7}

Obtenemos pues $x_2^* \approx x_2 = -1,84141$.

Problema 8.17 *Se pretende construir un depósito en forma de paralelepípedo sin tapa superior a partir de una plancha rectangular de dimensiones L y H . La construcción se basa en plegar la plancha tras extraer sus 4 vértices mediante el recortado de 4 cuadrados de lado ℓ . Es obvio que el volumen V de la caja construida verifica la ecuación $V = \ell(L - 2\ell)(H - 2\ell)$. Sabiendo que $L = 12$ y $H = 20$, queremos aproximar el valor de ℓ más pequeño posible de modo que $V = 49,5$. Localizar dicho valor de ℓ en un intervalo entre dos enteros consecutivos y utilizar dos iteraciones del método de Newton–Raphson asegurando la convergencia de la sucesión.*

Solución. La longitud ℓ satisface la ecuación polinomial cúbica $f(\ell) = \ell(L - 2\ell)(H - 2\ell) - V = 0$. Con los datos del enunciado se obtiene

$$f(\ell) = \ell(12 - 2\ell)(20 - 2\ell) - 49,5 .$$

Puesto que $f(\ell)$ es un polinomio cúbico, es fácil esbozar de manera cualitativa su gráfica. Para ello, primero calculamos las derivadas $f'(\ell) = 4(3\ell^2 - 32\ell + 60)$ y $f''(\ell) = 8(3\ell - 16)$. En concreto, $f'(\ell) = 0$ si y sólo si $\ell = 2/3(8 \pm \sqrt{19})$, es decir, $\ell = 2,4274$ y $\ell = 8,23927$. Además, $f''(\ell) = 0$ sólo para $\ell = 16/3 = 5,3333$. En definitiva, la gráfica de la función $f(\ell)$ tiene un máximo relativo en $\ell = 2,4274$, un mínimo relativo en $\ell = 8,23927$ y un punto de inflexión en $\ell = 5,3333$. Al ser un polinomio, dicha gráfica no tiene asíntotas horizontales ni verticales. De este modo, y teniendo en cuenta que $f(0) < 0$, $f(2,4274) > 0$ y $f(8,23927) < 0$, se obtiene la gráfica cualitativa de la Figura 8.2.

En particular, el cero ℓ^* más pequeño de la función f se encuentra en el intervalo $(0, 2,4274)$ que es un intervalo de decrecimiento y de curvatura positiva de la función f .

En resumen, como $f(0) = -49,5 < 0$, $f(1) = 130,5 > 0$ y $f \in C[0, 1]$, por el Teorema de Bolzano se tiene que existe al menos una raíz $\ell^* \in (0, 1)$ de la ecuación $f(\ell) = 0$. Además, esta raíz es única ya que $f'(\ell) > 0$ para todo $\ell \in [0, 1]$, de modo que f es creciente en el intervalo $[0, 1]$. Como además $f''(\ell) < 0$ para todo $\ell \in [0, 1]$, se concluye que podemos tomar el punto inicial $\ell_0 = 0$ en la iteración de Newton–Raphson asegurando de antemano su convergencia hacia ℓ^* . En definitiva, realizando la iteración

$$\ell_{n+1} = \ell_n - \frac{f(\ell_n)}{f'(\ell_n)} = \ell_n - \frac{\ell_n(12 - 2\ell_n)(20 - 2\ell_n) - 49,5}{4(3\ell_n^2 - 32\ell_n + 60)} , \quad \ell_0 = 0 ,$$

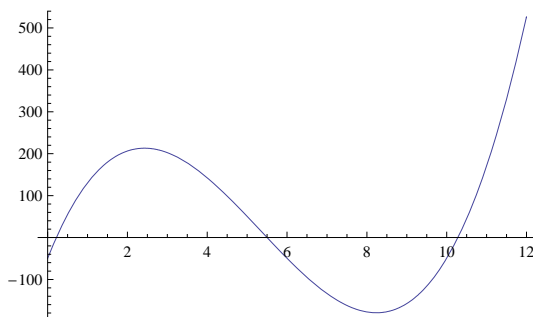


Figura 8.7: Gráfica del polinomio cúbico $f(\ell) = \ell(12 - 2\ell)(20 - 2\ell) - 49,5$.

se obtiene la siguiente tabla.

n	ℓ_n
0	0
1	0.20625
2	0.218801

La aproximación buscada de la longitud del cuadrado es $\ell^* \approx 0,218801$.

Problema 8.18 Si C es la cantidad total en euros de un préstamo, n es el número de pagos (mensualidades) a efectuar, a el importe de cada pago y r el interés mensual aplicado, se satisface la siguiente relación

$$Cr^n = a \frac{r^n - 1}{r - 1}.$$

Supongamos que nos ofrecen un crédito de 10000 euros a devolver en 100 mensualidades de 150 euros. Realizar dos iteraciones del método de Newton-Raphson para aproximar el tipo de interés mensual r que se ha aplicado sabiendo que $r > 1$. Tomar como valor inicial de la anterior sucesión un número natural adecuado y explicar dicha elección de modo que se garantice la convergencia de la sucesión de Newton-Raphson hacia r .

AYUDA. Escribir la ecuación que verifica r de la forma $f(r) = 0$ con f un polinomio.

Solución. A partir de la relación

$$Cr^n = a \frac{r^n - 1}{r - 1},$$

es claro que r satisface la ecuación polinomial $f(r) = 0$ siendo $f(r)$ el polinomio de grado $n + 1$ dado por

$$f(r) = Cr^n(r - 1) - a(r^n - 1) = 10000r^{100}(r - 1) - 150(r^{100} - 1).$$

La derivada de f es

$$\begin{aligned} f'(r) &= C(n+1)r^n - Cnr^{n-1} - anr^{n-1} = r^{n-1}[C(n+1)r - n(C+a)] \\ &= 5000r^{99}(202r - 203) . \end{aligned}$$

El método de Newton–Raphson para aproximar las soluciones de $f(r) = 0$ consiste en generar la sucesión $\{r_n\}_{n \in \mathbb{N}}$ definida de modo recurrente según

$$r_{i+1} = r_i - \frac{f(r_i)}{f'(r_i)} = r_i - \frac{10000r_i^{100}(r_i - 1) - 150(r_i^{100} - 1)}{5000r_i^{99}(202r_i - 203)}$$

Para elegir de modo adecuado el valor inicial r_0 , haremos un análisis cualitativo de la gráfica de $f(r)$. En primer lugar, los puntos críticos de $f(r)$, es decir los ceros de $f'(r)$, son $r = 0$ y $r = 203/202 \approx 1,00495$. En particular, vemos que si $r \in (0, 203/202)$ entonces $f'(r) < 0$ (y por lo tanto $f(r)$ decrece) mientras que si $r > 203/202$ entonces $f'(r) > 0$ y en consecuencia $f(r)$ crece. Finalmente, como $f(1) = 0$ por simple inspección y sabiendo que los polinomios no pueden tener asíntotas horizontales, se deduce que existe un único $r^* > 203/202$ tal que $f(r^*) = 0$. Este valor r^* es el interés mensual que se ha aplicado.

La segunda derivada de f es

$$f''(r) = 5000r^{98}(20200r - 20097) ,$$

de modo que los puntos de inflexión de $f(r)$ son los ceros de $f''(r)$, es decir, $r = 0$ y $r = 20097/20200 \approx 0,994901$. Tras el análisis efectuado, es fácil ver que un intervalo $[\alpha, \beta] \subset \mathbb{R}$ tal que $r^* \in [\alpha, \beta]$ y además el signo de $f'(r)$ y de $f''(r)$ sea constante en todo $[\alpha, \beta]$ debe ser de la forma $\alpha > 203/202$ con $f(\alpha) < 0$ y $f(\beta) > 0$. Además, para asegurar la convergencia de la sucesión $\{r_n\}_{n \in \mathbb{N}}$ generada por el método de Newton–Raphson hacia r^* , se debe tomar el valor inicial $r_0 = \beta$ puesto que el signo de $f(\beta)$ coincide con el signo de $f''(\beta)$. Calculando vemos que $f(2) > 0$ de manera que tomaremos $r_0 = 2$. Entonces, se obtiene la siguiente tabla.

i	r_i
0	2
1	1.9804
2	1.96099

La aproximación buscada es $r^* \approx r_2 = 1,96099$.

Problema 8.19 Considerar la ecuación trascendente $x^2 + \cos x - 2 = 0$.

- (i) Demostrar gráficamente que dicha ecuación tiene exactamente dos soluciones reales.
- (ii) Acotando la función $\cos x$, demostrar que si x^* es una solución de la ecuación, entonces $x^* \in [-\sqrt{3}, -1] \cup [1, \sqrt{3}]$.
- (iii) Demostrar que existe un único $x^* > 0$ solución de la ecuación.
- (iv) Aproximar el valor de la solución $x^* > 0$ mediante dos iteraciones del método de Newton-Raphson asegurando la convergencia de la sucesión.

Solución. (i) Las soluciones reales de la ecuación $x^2 + \cos x - 2 = 0$ corresponden con las abscisas de los puntos de corte de las gráficas de las funciones $y = \cos x$ e $y = 2 - x^2$. Dichas gráficas están dibujadas a continuación.

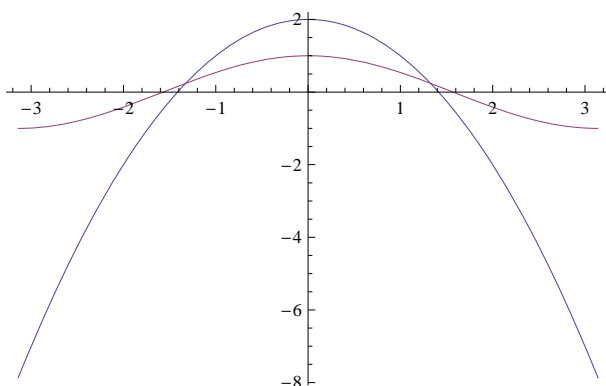


Figura 8.8: Gráficas superpuestas de las funciones $y = \cos x$ e $y = 2 - x^2$.

A la vista de la gráfica, es evidente que existen exactamente dos soluciones reales de dicha ecuación, siendo una positiva y la otra negativa. De hecho, puesto que la función $f(x) = x^2 + \cos x - 2$ es una función par (es decir $f(-x) = f(x)$ para todo $x \in \mathbb{R}$), es claro que si x^* es un cero de f , entonces $-x^*$ también lo es.

(ii) Reordenando la ecuación de la forma $x^2 = 2 - \cos x$ y teniendo en cuenta que $-1 \leq \cos x \leq 1$, es evidente que $(x^*)^2 \in [1, 3]$, de modo que, $x^* \in [-\sqrt{3}, -1] \cup [1, \sqrt{3}]$.

(iii) Definamos la función $f(x) = x^2 + \cos x - 2$. Teniendo en cuenta el resultado del apartado anterior, para demostrar que existe un único cero $x^* > 0$ de la función f basta con demostrar que f tiene un único cero en el intervalo $[1, \sqrt{3}]$. En primer lugar utilizaremos el Teorema de Bolzano. Puesto que $f \in C([1, \sqrt{3}])$ y además $f(1) > 0$ y $f(\sqrt{3}) < 0$, por el Teorema de Bolzano se sabe que f tiene

al menos un cero en el intervalo $[1, \sqrt{3}]$. Finalmente mostraremos la unicidad de dicha raíz puesto que $f'(x) = 2x - \sin x \neq 0$ para todo $x \in [1, \sqrt{3}]$. De hecho, se tiene que $f'(x) > 1$ para todo $x \in [1, \sqrt{3}]$ debido a que $-1 \leq \sin x \leq 1$.

(iv) Una vez sabemos que existe un único $x^* \in [1, \sqrt{3}]$ tal que $f(x^*) = 0$, vamos a aproximarlos mediante dos iteraciones del método de Newton–Raphson $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$. Para asegurar la convergencia de la sucesión $\{x_n\}_{n \in \mathbb{N}}$, hemos de elegir adecuadamente el punto inicial x_0 . Tomamos el intervalo $[a, b] = [1, \sqrt{3}]$. Por los cálculos realizados en el apartado (iii) sabemos que $f \in C^2[a, b]$ con $f(a)f(b) < 0$. Además, $f'(x) > 0$ para todo $x \in [a, b]$. Sólo falta estudiar el signo de $f''(x)$ en el intervalo $[a, b]$. Puesto que $f''(x) = 2 - \cos x$, es claro que $f''(x) > 0$ para todo $x \in [a, b]$. De este modo, se tomará el valor inicial $x_0 \in \{a, b\}$ que verifique $\text{sign} f(x_0) = \text{sign} f''(x_0)$. En definitiva, $x_0 = a = 1$. Realizando la iteración de Newton–Raphson

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 + \cos x_n - 2}{2x_n - \sin x_n}, \quad x_0 = 1,$$

se obtienen los primeros términos

n	x_n	$f(x_n)$
0	1	-0.459698
1	1.39679	0.12416
2	1.32815	0.00425146

Se concluye con la aproximación $x^* \approx x_2 = 1,32815$.

Problema 8.20 Consideremos la ecuación $x^3 + 10x = -\cos x$.

- (i) Demostrar gráfica y analíticamente que la ecuación tiene una única solución $x^* \in \mathbb{R}$. Además, acotarla entre dos números enteros consecutivos, es decir, hallar $m \in \mathbb{Z}$ tal que $x^* \in [m, m+1]$.
- (ii) Hallar el mínimo número de iteraciones que se deben realizar con el algoritmo de bisección partiendo del intervalo $[m, m+1]$ dado en el apartado (i) para obtener una aproximación de x^* con un error menor que 10^{-1} . Obtener mediante ese número de iteraciones el intervalo $I \subset \mathbb{R}$ que contiene a x^* .
- (iii) Mejorar la aproximación de x^* dada en el apartado anterior aplicando 2 iteraciones del método de Newton–Raphson partiendo del intervalo I hallado en el apartado (ii). Usar una condición inicial adecuada de modo que se garantice la convergencia de la sucesión Newton–Raphson hacia x^* .

Solución. (i) Escribimos la ecuación de la forma $g(x) = h(x)$ siendo $g(x) = x^3 + 10x = x(x^2 + 10)$ y $h(x) = -\cos x$. Es fácil hallar cualitativamente la gráfica de las dos funciones $g(x)$ y $h(x)$. En particular, $|h(x)| \leq 1$, la gráfica de $g(x)$ sólo corta al eje de abscisas en $x = 0$ y, puesto que $g'(x) = 3x^2 + 10 > 0$

para todo $x \in \mathbb{R}$, se concluye que el polinomio $g(x)$ es una función creciente en todo \mathbb{R} . Teniendo en cuenta además que $h(x)$ oscila tomando valores entre -1 y 1 , que $h(\pm\pi/2) = 0$ y que $g(\pm\pi/2) = \pm 19,58$ se concluye que las gráficas de las funciones $g(x)$ y $h(x)$ sólo se cortan en un punto de abscisa $x^* \in [-\pi/2, 0]$. Por supuesto, el valor x^* es pues la única solución de la ecuación $g(x) = h(x)$.

Para hallar $m \in \mathbb{Z}$ tal que $x^* \in [m, m+1]$, primero reescribimos la ecuación de la forma $f(x) = 0$, siendo $f(x) = x^3 + 10x + \cos x$. Como $f \in C(\mathbb{R})$ y además $f(0) = 1 > 0$ y $f(-1) = -11 + \cos(1) = -10,4597 < 0$, aplicando el Teorema de Bolzano se tiene que existe $x^* \in [-1, 0]$ tal que $f(x^*) = 0$.

(ii) Utilizando el algoritmo de bisección, se obtiene una sucesión de intervalos encajados $(a_0, b_0) = (-1, 0) \supset (a_1, b_1) \supset \cdots \supset (a_n, b_n) \supset \cdots$, de modo que la sucesión de puntos medios $\{c_n\}_{n=0}^\infty$ con $c_n = (a_n + b_n)/2$ satisface $\lim_{n \rightarrow \infty} c_n = x^*$. Además, el error absoluto $\Delta = |x^* - c_n|$ cometido en la aproximación $x^* \approx c_n$ tras realizar n iteraciones con el algoritmo de bisección está acotado por $\Delta \leq (b_0 - a_0)/2^{n+1}$. Imponiendo que $\Delta \leq 1/2^{n+1} < 10^{-1}$ se tiene que

$$n > -1 + \frac{\ln 10}{\ln 2} = 2,32,$$

de modo que $n = 3$.

Procedemos a calcular, según el método de bisección, los intervalos $(a_0, b_0) = (-1, 0) \supset (a_1, b_1) \supset (a_2, b_2) \supset I = (a_3, b_3)$. El punto medio $c_0 = (a_0 + b_0)/2 = -1/2$ verifica $f(-1/2) = -4,24742 < 0$ de modo que $(a_1, b_1) = (-1/2, 0)$. Luego, el punto medio $c_1 = (a_1 + b_1)/2 = -1/4$ satisface $f(-1/4) = -1,54671 < 0$ con lo que $(a_2, b_2) = (-1/4, 0)$. Finalmente, el punto medio $c_2 = (a_2 + b_2)/2 = -1/8$, con lo que $f(-1/8) = -0,259755 < 0$ y se tiene $I = (a_3, b_3) = (-1/8, 0)$. Concluimos que $x^* \in I = (-1/8, 0)$.

(iii) Sabemos de los apartados anteriores que existe un único $x^* \in I = [-1/8, 0]$ tal que $f(x^*) = 0$. Además, $f \in C^2(I)$ con $f(-1/8)f(0) < 0$ ya que $f(-1/8) < 0$ y $f(0) > 0$. Por otra parte, $f'(x) = 3x^2 + 10 - \sin x \geq 9 > 0$ para todo $x \in \mathbb{R}$ puesto que $3x^2 + 10 \geq 10$ y $-1 \leq \sin x \leq 1$. Entonces, en particular $f'(x) > 0$ para todo $x \in I$. Finalmente, es evidente que $f''(x) = 6x - \cos x < 0$ para todo $x \in I$. Entonces, tomando el punto inicial x_0 de la sucesión Newton-Raphson como $x_0 \in \{-1/8, 0\}$ que verifique $\text{sign} f(x_0) = \text{sign} f''(x_0)$, es decir, tomando $x_0 = -1/8$ la sucesión de Newton-Raphson $\{x_n\}_{n \in \mathbb{N}}$ converge hacia la única raíz x^* de f en I .

Acabamos el problema realizando 2 iteraciones de Newton-Raphson

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} = x_i - \frac{x_i^3 + 10x_i + \cos x_i}{3x_i^2 + 10 - \sin x_i}$$

y obteniendo la tabla siguiente

i	x_i
0	-1/8
1	-0.0994625
2	-0.0994081

Se concluye que $x^* \approx x_2 = -0,0994081$.

Capítulo 9

Aproximación de Funciones

9.1. Resumen de teoría

El problema básico en este capítulo consiste en reconstruir la función $f(x)$ teniendo en cuenta la clase funcional a la cual pertenece. La idea principal consiste en representar la función f mediante una combinación lineal de funciones linealmente independientes $\varphi_1(x), \dots, \varphi_n(x)$ pertenecientes a dicha clase funcional, es decir

$$f(x) \approx f^*(x) := c_1 \varphi_1(x) + \dots + c_n \varphi_n(x) .$$

Se tienen que calcular los coeficientes de la combinación lineal c_i para $i = 1, \dots, n$, de manera que f^* aproxime a f de la “mejor manera posible”.

Un planteamiento general del problema de la aproximación de funciones sería el siguiente: dada una función f perteneciente a un cierto espacio de funciones \mathcal{F} dotado de una norma $\|\cdot\|$ y un subespacio $\mathcal{G} \subset \mathcal{F}$ de dimensión finita, encontrar la función $f^* \in \mathcal{G}$ que mejor aproxima a la función f con la norma dada, es decir, tal que

$$\|f - f^*\| \leq \|f - g\| \quad \text{para toda función } g \in \mathcal{G} .$$

En otras palabras, la función f^* es la *mejor aproximación* de la función f respecto al subespacio \mathcal{G} con la norma $\|\cdot\|$ si

$$\|f - f^*\| = \inf_{g \in \mathcal{G}} \|f - g\| .$$

La aproximación efectuada será *discreta* si sólo se necesitan un número finito de puntos para calcular la norma. En caso contrario, es decir, si es necesario evaluar la función en todos los puntos de un intervalo $[a, b]$ para calcular la norma se dice que la aproximación es *continua*. Existe otra clasificación de las aproximaciones atendiendo a la norma utilizada. De esta manera, si se utiliza la norma del infinito o del máximo $\|\cdot\|_\infty$ se habla de aproximación *min-max* y si se

utiliza la norma euclidiana $\|\cdot\|_2$ se llama aproximación por *mínimos cuadrados*. En concreto, se tiene:

- La aproximación *min-max discreta* utiliza la seminorma

$$\|f\|_\infty = \max_{1 \leq i \leq n} |f(x_i)| .$$

- La aproximación *min-max continua* utiliza la norma

$$\|f\|_\infty = \max_{x \in [a, b]} |f(x)| .$$

- La aproximación por *mínimos cuadrados discreta* utiliza la seminorma

$$\|f\|_2 = \sqrt{\sum_{i=1}^n [f(x_i)]^2} .$$

- La aproximación por *mínimos cuadrados continua* utiliza la norma

$$\|f\|_2 = \sqrt{\int_a^b [f(x)]^2 dx} .$$

Teorema 9.1 Sea \mathcal{F} un espacio vectorial de funciones normado. Sean $\varphi_1, \dots, \varphi_n \in \mathcal{F}$ un conjunto de funciones linealmente independientes. Entonces, para cada función $f \in \mathcal{F}$ existe una única combinación lineal $f^* = \sum_{i=1}^n c_i \varphi_i$ tal que $\|f - f^*\| \leq \|f - g\|$ para cualquier otra combinación lineal $g = \sum_{i=1}^n \tilde{c}_i \varphi_i$.

Nuestro objetivo consiste en calcular el valor de los coeficientes c_i de manera que $\|f - f^*\|$ o equivalentemente $\|f - f^*\|^2$ tenga un valor mínimo. Para conseguirlo se ha de imponer la anulación de las derivadas parciales

$$\frac{\partial}{\partial c_j} \|f - f^*\|^2 = 0 \quad j = 1, \dots, n .$$

Desarrollando esta expresión obtenemos las *ecuaciones normales*

$$\begin{pmatrix} \langle \varphi_1, \varphi_1 \rangle & \cdots & \langle \varphi_1, \varphi_n \rangle \\ \vdots & \ddots & \vdots \\ \langle \varphi_n, \varphi_1 \rangle & \cdots & \langle \varphi_n, \varphi_n \rangle \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \langle f, \varphi_1 \rangle \\ \vdots \\ \langle f, \varphi_n \rangle \end{pmatrix} , \quad (9.1)$$

donde $\langle \cdot, \cdot \rangle$ denota el producto escalar asociado a la norma $\|\cdot\|$, es decir, $\|f\|^2 = \langle f, f \rangle$.

Otra forma alternativa de llegar a las ecuaciones normales es mediante los sistemas lineales sobredeterminados.

Teorema 9.2 Sea A una matriz $m \times n$ con $m > n$ y $b \in \mathbb{R}^m$. Si $x \in \mathbb{R}^n$ verifica las ecuaciones normales $A^T(b - Ax) = 0$, entonces $\|b - Ax\|_2 \leq \|b - Ay\|_2$ para todo $y \in \mathbb{R}^n$.

9.2. Problemas resueltos

Problema 9.1 *Un ingeniero en una central nuclear mide las siguientes oscilaciones de la temperatura T en función del tiempo t en el interior de un dispositivo termonuclear*

t_i (s)	0	1	2	3
$T(t_i)$ (K)	27	34	25	38

(i) *Si el modelo físico al que deben responder las medidas anteriores es*

$$T(t) = \alpha + \beta \sin t + \gamma \cos^2 t ,$$

calcular las constantes α , β y γ de manera que la función $T(t)$ sea la mejor aproximación, según la teoría de mínimos cuadrados, a los datos experimentales.

(ii) *Se pretende ahora aproximar la función $T(t)$ por un polinomio $P(t)$ de grado menor o igual que 2 que pase por los tres primeros puntos $(t_i, T(t_i))$ con $i = 0, 1, 2$.*

(a) *Utilizando el método de las diferencias divididas de Newton, obtener el polinomio $P(t)$.*

(b) *Supongamos que ciertas consideraciones físicas indican que la función tercera derivada de la temperatura respecto del tiempo $T^{(3)}(t)$ debe estar acotada de la forma $|T^{(3)}(t)| \leq 0,1$ para todo tiempo $t \in [0, 2]$. Obtener una cota (independiente del tiempo) del error absoluto cometido al realizar la aproximación $T(t) \approx P(t)$ para cualquier tiempo $t \in [0, 2]$.*

(c) *Utilizar el polinomio del apartado (a) para realizar la aproximación $\int_0^2 T(t) dt \approx \int_0^2 P(t) dt$. Explicar cual es la fórmula de Newton-Cotes con la que se obtiene la misma aproximación.*

Solución. (i) La temperatura en función del tiempo se quiere aproximar por una función del tipo

$$T(t) = \alpha \varphi_1(t) + \beta \varphi_2(t) + \gamma \varphi_3(t) ,$$

siendo $\varphi_1(t) = 1$, $\varphi_2(t) = \sin t$ y $\varphi_3(t) = \cos^2 t$. Las ecuaciones normales para obtener los parámetros α , β y γ según la teoría de mínimos cuadrados son

$$\begin{pmatrix} \langle \varphi_1, \varphi_1 \rangle & \langle \varphi_1, \varphi_2 \rangle & \langle \varphi_1, \varphi_3 \rangle \\ \langle \varphi_2, \varphi_1 \rangle & \langle \varphi_2, \varphi_2 \rangle & \langle \varphi_2, \varphi_3 \rangle \\ \langle \varphi_3, \varphi_1 \rangle & \langle \varphi_3, \varphi_2 \rangle & \langle \varphi_3, \varphi_3 \rangle \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \langle T, \varphi_1 \rangle \\ \langle T, \varphi_2 \rangle \\ \langle T, \varphi_3 \rangle \end{pmatrix} ,$$

donde \langle, \rangle denota el producto escalar discreto euclidean. Se tiene pues que

$$\langle \varphi_1, \varphi_1 \rangle = \sum_{i=1}^4 \varphi_1^2(t_i) = \sum_{i=1}^4 1 = 4 ,$$

$$\begin{aligned}
\langle \varphi_2, \varphi_2 \rangle &= \sum_{i=1}^4 \varphi_2^2(t_i) = \sum_{i=1}^4 \sin^2(t_i) = 0 + 0,708073 + 0,826822 \\
&\quad + 0,0199149 = 1,55481, \\
\langle \varphi_3, \varphi_3 \rangle &= \sum_{i=1}^4 \varphi_3^2(t_i) = \sum_{i=1}^4 \cos^4(t_i) = 1 + 0,0852211 + 0,0299907 \\
&\quad + 0,960567 = 2,07578, \\
\langle \varphi_1, \varphi_2 \rangle &= \sum_{i=1}^4 \varphi_1(t_i) \varphi_2(t_i) = \sum_{i=1}^4 \sin(t_i) = 0 + 0,841471 + 0,909297 \\
&\quad + 0,14112 = 1,89189, \\
\langle \varphi_1, \varphi_3 \rangle &= \sum_{i=1}^4 \varphi_1(t_i) \varphi_3(t_i) = \sum_{i=1}^4 \cos^2(t_i) = 1 + 0,291927 + 0,173178 \\
&\quad + 0,980085 = 2,44519, \\
\langle \varphi_2, \varphi_3 \rangle &= \sum_{i=1}^4 \varphi_2(t_i) \varphi_3(t_i) = \sum_{i=1}^4 \sin(t_i) \cos^2(t_i) = 0,541428, \\
\langle T, \varphi_1 \rangle &= \sum_{i=1}^4 T(t_i) \varphi_1(t_i) = 27 + 34 + 25 + 38 = 124, \\
\langle T, \varphi_2 \rangle &= \sum_{i=1}^4 T(t_i) \varphi_2(t_i) = 56,705, \\
\langle T, \varphi_3 \rangle &= \sum_{i=1}^4 T(t_i) \varphi_3(t_i) = 78,4982.
\end{aligned}$$

Finalmente, como el producto escalar verifica la propiedad conmutativa, las ecuaciones normales son

$$\begin{pmatrix} 4 & 1,89189 & 2,44519 \\ 1,89189 & 1,55481 & 0,541428 \\ 2,44519 & 0,541428 & 2,07578 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 124 \\ 56,705 \\ 78,4982 \end{pmatrix},$$

y su solución es $(\alpha, \beta, \gamma) = (-87,0139, 102,828, 113,495)$.

Otra forma de resolver el problema consiste en imponer que la función $T(t)$ pase por todos los puntos experimentales $(t_i, T(t_i))$ con $i = 1, 2, 3, 4$. De esta forma se obtiene un sistema lineal sobredeterminado $Ax = b$ para $x = (\alpha, \beta, \gamma)^T$. En concreto se tiene

$$\begin{pmatrix} 1 & \sin 0 & \cos^2 0 \\ 1 & \sin 1 & \cos^2 1 \\ 1 & \sin 2 & \cos^2 2 \\ 1 & \sin 3 & \cos^2 3 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 27 \\ 34 \\ 25 \\ 38 \end{pmatrix}.$$

Según la teoría de mínimos cuadrados, la mejor aproximación vendrá dada por la solución de las ecuaciones normales $A^t Ax = A^t b$. En este problema se

tiene

$$A = \begin{pmatrix} 1 & \sin 0 & \cos^2 0 \\ 1 & \sin 1 & \cos^2 1 \\ 1 & \sin 2 & \cos^2 2 \\ 1 & \sin 3 & \cos^2 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0,841471 & 0,291927 \\ 1 & 0,909297 & 0,173178 \\ 1 & 0,14112 & 0,980085 \end{pmatrix},$$

de manera que las ecuaciones normales adoptan la forma

$$\begin{pmatrix} 4 & 1,89189 & 2,44519 \\ 1,89189 & 1,55481 & 0,541428 \\ 2,44519 & 0,541428 & 2,07578 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 124 \\ 56,705 \\ 78,4982 \end{pmatrix},$$

siendo su solución es $(\alpha, \beta, \gamma) = (-86,7386, 102,583, 113,234)$.

Observar que los dos métodos utilizados llevan a las mismas ecuaciones normales, aunque debido a los errores de redondeo en los cálculos efectuados exista algún dígito diferente en ellas.

(iia) Realizamos el siguiente esquema de diferencias divididas

0	27		
		7	
1	34		-8
		-9	
2	25		

El polinomio interpolador es

$$P(t) = T[t_0] + T[t_0, t_1](t - t_0) + T[t_0, t_1, t_2](t - t_0)(t - t_1) = 27 + 7t - 8t(t - 1),$$

que desarrollado resulta $P(t) = 27 + 15t - 8t^2$.

(iib) Se sabe que el error absoluto Δ cometido al realizar la aproximación $T(t) \approx P(t)$ para cualquier tiempo $t \in [0, 2]$ viene dado por

$$\Delta = |T(t) - P(t)| = \left| \frac{T^{(3)}(\xi)}{3!} \prod_{i=0}^2 (t - t_i) \right|,$$

siendo $\xi \in (0, 2)$. Como tanto t como t_i pertenecen al intervalo $[0, 2]$ se tiene que $|t - t_i| \leq 2$ y por lo tanto $\prod_{i=0}^2 |t - t_i| \leq 2^3 = 8$. Además el enunciado asegura que $|T^{(3)}(t)| \leq 0,1$ para todo tiempo $t \in [0, 2]$ de manera que $|T^{(3)}(\xi)| \leq 0,1$. Con estas acotaciones se tiene que

$$\Delta = |T(t) - P(t)| \leq \frac{0,1}{3!} 8 = 0,13333.$$

(iic) La aproximación requerida es

$$\int_0^2 T(t) dt \approx \int_0^2 P(t) dt = \int_0^2 [27 + 7t - 8t(t - 1)] dt = \frac{188}{3} \approx 62,66.$$

Como $P(t)$ es el polinomio que interpola a la función $T(t)$ en las abscisas equiespaciadas $a = t_0 < t_1 < \cdots < t_n = b$ y además la aproximación efectuada es $\int_a^b T(t) dt \approx \int_a^b P(t) dt$ es evidente que se ha utilizado una fórmula de Newton-Cotes. Además como $n = 2$ la fórmula es la de Simpson.

Problema 9.2 La viscosidad μ de un cierto fluido varía con la temperatura T a la que está sometido el fluido. Un ingeniero realiza las siguientes mediciones

T_k ($^{\circ}C$)	1	10	38
μ_k (centipoises)	50	10	5

- (i) Suponiendo que se tiene el siguiente modelo físico $\mu(T) = \mu_0 T^k$, calcular las constantes μ_0 y k de manera que la función $\mu(T)$ sea la mejor aproximación por mínimos cuadrados a los datos experimentales de la tabla.
- (ii) Aproximar el valor $\mu(20)$ mediante interpolación polinomial de Lagrange con todos los datos de la tabla de valores experimentales. En vista del resultado obtenido, ¿se puede afirmar que es una buena aproximación?
- (iii) Supongamos que podemos realizar las tres mediciones experimentales (T_k, μ_k) con $k = 0, 1, 2$, de la tabla para cualquier valor $T_k \in [1, 38]$. Sabiendo que posteriormente se pretende realizar una aproximación por interpolación polinomial como en el anterior apartado, dar de manera razonada las tres temperaturas T_k donde se efectuarían las mediciones con el objetivo de disminuir el error en la aproximación interpolatoria.

Solución. (i) Puesto que el parámetro k no interviene de forma lineal en la expresión de la función $\mu(T) = \mu_0 T^k$, realizamos una estrategia de linearización tomando logaritmos en ambos miembros de la anterior igualdad. De esta forma, utilizando las propiedades de la función logaritmo se tiene que

$$\ln \mu(T) = \ln \mu_0 + k \ln T. \quad (9.2)$$

En segundo lugar, realizamos una nueva tabulación de la nueva variable $\ln T$ y la nueva función $\ln \mu$ a partir de la tabla de valores del enunciado.

$\ln T_k$	0	2.30259	3.63759
$\ln \mu_k$	3.91202	2.30259	1.60944

A continuación imponemos que la función (9.2) verifique todos los puntos de la nueva tabulación. Así, se obtiene el siguiente sistema lineal sobredeterminado

$$\begin{pmatrix} 1 & 0 \\ 1 & 2.30259 \\ 1 & 3.63759 \end{pmatrix} \begin{pmatrix} \log \mu_0 \\ k \end{pmatrix} = \begin{pmatrix} 3.91202 \\ 2.30259 \\ 1.60944 \end{pmatrix}.$$

Buscaremos la “mejor solución posible” del anterior sistema lineal, dentro de la teoría de mínimos cuadrados, multiplicando por la izquierda el anterior

sistema lineal por la matriz traspuesta de los coeficientes del sistema. De esta forma se obtiene el sistema lineal de ecuaciones normales

$$\begin{pmatrix} 3 & 5,94018 \\ 5,94018 & 18,534 \end{pmatrix} \begin{pmatrix} \log \mu_0 \\ k \end{pmatrix} = \begin{pmatrix} 7,82405 \\ 11,1564 \end{pmatrix}, \quad (9.3)$$

cuya solución es $(\log \mu_0, k) = (3,87571, -0,64023)$. Por lo tanto

$$\mu_0 = \exp(3,87571) = 48,217,$$

y, en definitiva, la mejor aproximación por mínimos cuadrados viene dada por

$$\mu(T) = 48,217 T^{-0,64023}.$$

Existe otra forma de resolver el problema, basada en la obtención de las ecuaciones normales (9.3) a partir del producto escalar euclideo discreto que denotaremos por $\langle \cdot, \cdot \rangle$. En concreto, tomando la función $f(T) = \ln \mu(T)$ y las funciones $\varphi_1(T) = 1$, $\varphi_2(T) = \ln T$, se tiene que $f(T)$ es una combinación lineal de $\varphi_1(T)$, $\varphi_2(T)$, es decir, $f(T) = \ln \mu_0 \varphi_1(T) + k \varphi_2(T)$. De esta forma, se sabe que las ecuaciones normales adoptan la forma

$$\begin{pmatrix} \langle \varphi_1, \varphi_1 \rangle & \langle \varphi_1, \varphi_2 \rangle \\ \langle \varphi_2, \varphi_1 \rangle & \langle \varphi_2, \varphi_2 \rangle \end{pmatrix} \begin{pmatrix} \log \mu_0 \\ k \end{pmatrix} = \begin{pmatrix} \langle f, \varphi_1 \rangle \\ \langle f, \varphi_2 \rangle \end{pmatrix}. \quad (9.4)$$

Calculando los productos escalares euclideos discretos asociados a la tabla del enunciado e involucrados en este sistema se tiene

$$\begin{aligned} \langle \varphi_1, \varphi_1 \rangle &= \sum_{k=0}^2 \varphi_1^2(T_k) = 1^2 + 1^2 + 1^2 = 3, \\ \langle \varphi_2, \varphi_2 \rangle &= \sum_{k=0}^2 \varphi_2^2(T_k) \\ &= \ln^2 T_0 + \ln^2 T_1 + \ln^2 T_2 = \ln^2 1 + \ln^2 10 + \ln^2 38 = 18,534, \\ \langle \varphi_1, \varphi_2 \rangle &= \sum_{k=0}^2 \varphi_1(T_k) \varphi_2(T_k) = \ln 1 + \ln 10 + \ln 38 = 5,94018, \\ \langle f, \varphi_1 \rangle &= \sum_{k=0}^2 f(T_k) \varphi_1(T_k) \\ &= \ln \mu(1) + \ln \mu(10) + \ln \mu(38) = \ln 50 + \ln 10 + \ln 5 \\ &= 7,82405, \\ \langle f, \varphi_2 \rangle &= \sum_{k=0}^2 f(T_k) \varphi_2(T_k) = \ln 50 \ln 1 + \ln 10 \ln 10 + \ln 5 \ln 38 \\ &= 11,1564. \end{aligned}$$

De estos cálculos se desprende que las ecuaciones normales (9.4) coinciden con las ecuaciones (9.3) y por lo tanto el problema sigue de igual manera.

(ii) En este segundo apartado del problema, realizaremos una aproximación polinomial $\mu(T) \approx P_2(T)$, donde $P_2(T)$ es el polinomio de grado menor o igual que 2 que interpola a todos los datos (T_k, μ_k) de la tabla del enunciado.

Utilizando el formalismo de Lagrange se tiene

$$P_2(T) = \sum_{k=0}^2 \mu_k \ell_k(T) = 50\ell_0(T) + 10\ell_1(T) + 5\ell_2(T) ,$$

donde los polinomios básicos de Lagrange $\ell_k(T)$ son

$$\begin{aligned}\ell_0(T) &= \frac{T - T_1}{T_0 - T_1} \frac{T - T_2}{T_0 - T_2} = \frac{T - 10}{1 - 10} \frac{T - 38}{1 - 38} , \\ \ell_1(T) &= \frac{T - T_0}{T_1 - T_0} \frac{T - T_2}{T_1 - T_2} = \frac{T - 1}{10 - 1} \frac{T - 38}{10 - 38} , \\ \ell_2(T) &= \frac{T - T_0}{T_2 - T_0} \frac{T - T_1}{T_2 - T_1} = \frac{T - 1}{38 - 1} \frac{T - 10}{38 - 10} .\end{aligned}$$

Finalmente, la aproximación que se desea efectuar es

$$\begin{aligned}\mu(20) &\approx P_2(20) = 50\ell_0(20) + 10\ell_1(20) + 5\ell_2(20) \\ &= 50 \times \left(\frac{-20}{37} \right) + 10 \times \frac{19}{14} + 5 \times \frac{95}{518} = -12,5386 .\end{aligned}$$

Notamos que la aproximación que acabamos de realizar no es buena puesto que hemos obtenido que cuando la temperatura es $20^\circ C$, la viscosidad del fluido es $\mu(20) < 0$, lo cual no concuerda con la realidad física, ya que $\mu(T) > 0$ para $-273,16 < T < \infty$. El motivo ha sido la oscilación no deseada del polinomio interpolador $P_2(T)$.

(iii) Recordemos que, si se quiere realizar una interpolación polinomial con $n + 1$ nodos en el intervalo $[-1, 1]$ para posteriormente aproximar una función suficientemente derivable en dicho intervalo, la mejor elección de los nodos con el objetivo de minimizar el error en dicha aproximación con la norma del máximo o del infinito vienen dados por las abscisas de Chebychev $x_k \in (-1, 1)$ definidas como

$$x_k = \cos \left(\frac{2k + 1}{2n + 2} \pi \right) , \quad k = 0, 1, \dots, n .$$

Si se pretende realizar una interpolación polinomial en el intervalo $[a, b]$, lo único que se debe realizar es un cambio lineal de variables que lleve el intervalo $[-1, 1]$ al intervalo $[a, b]$.

En concreto, en el problema planteado, los nodos $T_k \in [a, b] = [1, 38]$, de manera que el cambio de variable

$$T = \frac{39}{2} + \frac{37}{2}x ,$$

lleva los valores $x \in [-1, 1]$ a valores $T \in [1, 38]$. Por lo tanto, los nodos que tomaremos son

$$T_k = \frac{39}{2} + \frac{37}{2}x_k = \frac{39}{2} + \frac{37}{2} \cos\left(\frac{2k+1}{2n+2}\pi\right), \quad k = 0, 1, \dots, n.$$

Finalmente, como $n = 2$, las temperaturas T_k donde se realizarán las mediciones de la viscosidad serán

$$\begin{aligned} T_0 &= \frac{39}{2} + \frac{37}{2} \cos\left(\frac{\pi}{6}\right) = \frac{39}{2} + \frac{37}{2} \frac{\sqrt{3}}{2} = \frac{78 + 37\sqrt{3}}{4} = 35,5215^\circ C, \\ T_1 &= \frac{39}{2} + \frac{37}{2} \cos\left(\frac{\pi}{2}\right) = \frac{39}{2} = 19,5^\circ C, \\ T_2 &= \frac{39}{2} + \frac{37}{2} \cos\left(\frac{5\pi}{6}\right) = \frac{39}{2} - \frac{37}{2} \frac{\sqrt{3}}{2} = \frac{78 - 37\sqrt{3}}{4} = 3,47853^\circ C. \end{aligned}$$

Problema 9.3 La intensidad I de la radiación de una fuente radioactiva depende del tiempo de la forma $I(t) = I_0 \exp(-\alpha t)$. Determinar, mediante el método de mínimos cuadrados, la intensidad inicial de la radiación I_0 y la constante de desintegración α sabiendo que se han realizado las mediciones siguientes

t	0.2	0.3	0.4	0.5	0.6	0.7	0.8
I	3.16	2.38	1.75	1.34	1.00	0.74	0.56

Solución. Observemos que el parámetro α no interviene de forma lineal en la expresión de la función a aproximar $I(t)$. Por lo tanto se realizará una linearización de la forma siguiente. Tomando logaritmos neperianos en ambos miembros de la ecuación $I(t) = I_0 \exp(-\alpha t)$ se obtiene que $\ln I(t) = \ln I_0 - \alpha t$. De este modo, los nuevos parámetros $\ln I_0$ y α intervienen de forma lineal. Calculamos una nueva tabulación a partir de la dada en el enunciado

t	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$\ln I$	1.15057	0.8671	0.559616	0.29267	0	-0.301105	-0.579818

Ahora imponemos que la ecuación $\ln I(t) = \ln I_0 - \alpha t$ sea verificada para todos los puntos de esta tabulación. Obtenemos, en definitiva, el siguiente sistema lineal sobredeterminado $Ax = b$ para $x = (\ln I_0, \alpha)^T$. En concreto se tiene

$$\begin{pmatrix} 1 & -0,2 \\ 1 & -0,3 \\ 1 & -0,4 \\ 1 & -0,5 \\ 1 & -0,6 \\ 1 & -0,7 \\ 1 & -0,8 \end{pmatrix} \begin{pmatrix} \ln I_0 \\ \alpha \end{pmatrix} = \begin{pmatrix} 1,15057 \\ 0,8671 \\ 0,559616 \\ 0,29267 \\ 0 \\ -0,301105 \\ -0,579818 \end{pmatrix}.$$

Por supuesto, el anterior sistema lineal sobredeterminado es incompatible, sin embargo, en el contexto de la teoría de mínimos cuadrados, la “mejor aproximación” vendrá dada por la solución de las ecuaciones normales $A^t Ax = A^t b$,

siendo A^t la matriz traspuesta de A . Aplicado al anterior sistema sobredeterminado, las ecuaciones normales resultan

$$\begin{pmatrix} 7 & -3,5 \\ -3,5 & 2,03 \end{pmatrix} \begin{pmatrix} \ln I_0 \\ \alpha \end{pmatrix} = \begin{pmatrix} 1,98903 \\ -0,185798 \end{pmatrix}.$$

La solución de estas ecuaciones normales es $(\ln I_0, \alpha) = (1,72829, 2,88828)$. Por lo tanto, la solución es

$$(I_0, \alpha) = (5,63102, 2,88828).$$

Se puede comprobar el ajuste obtenido mediante la gráfica de la Figura 9.1.

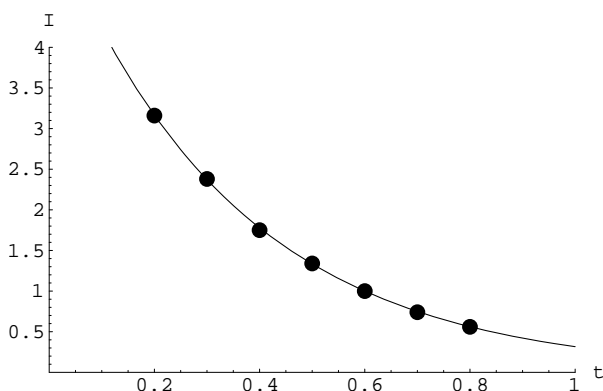


Figura 9.1: Gráfica de la intensidad radioactiva $I(t)$ en función del tiempo superpuesta a la nube de puntos experimentales.

Problema 9.4 La posición x en función del tiempo t para cualquier movimiento rectilíneo uniforme es $x(t) = x_0 + v_0 t + \frac{1}{2}at^2$, siendo x_0 , v_0 y a la posición inicial, velocidad inicial y aceleración de móvil. Se han realizado las siguientes mediciones experimentales

t_i	0	1	2	3	4
$x(t_i)$	1.1	0.3	-6.5	-21	-39

Hallar x_0 , v_0 y a por el método de mínimos cuadrados.

Solución. Imponiendo que la función $x(t) = x_0 + v_0 t + \frac{1}{2}at^2$ pase por todos los puntos de la tabulación del enunciado obtenemos el sistema lineal sobredeterminado

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1/2 \\ 1 & 2 & 2 \\ 1 & 3 & 9/2 \\ 1 & 4 & 8 \end{pmatrix} \begin{pmatrix} x_0 \\ v_0 \\ a \end{pmatrix} = \begin{pmatrix} 1,1 \\ 0,3 \\ -6,5 \\ -21 \\ -39 \end{pmatrix}$$

Multiplicamos por la izquierda ambos miembros de esta ecuación por la matriz traspuesta de la matriz de coeficientes del sistema obteniendo las ecuaciones normales

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1/2 \\ 1 & 2 & 2 \\ 1 & 3 & 9/2 \\ 1 & 4 & 8 \end{pmatrix}^T \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1/2 \\ 1 & 2 & 2 \\ 1 & 3 & 9/2 \\ 1 & 4 & 8 \end{pmatrix} \begin{pmatrix} x_0 \\ v_0 \\ a \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1/2 \\ 1 & 2 & 2 \\ 1 & 3 & 9/2 \\ 1 & 4 & 8 \end{pmatrix}^T \begin{pmatrix} 1,1 \\ 0,3 \\ -6,5 \\ -21 \\ -39 \end{pmatrix}.$$

Operando, se llega a que

$$\begin{pmatrix} 5 & 10 & 15 \\ 10 & 30 & 50 \\ 15 & 50 & 177/2 \end{pmatrix} \begin{pmatrix} x_0 \\ v_0 \\ a \end{pmatrix} = \begin{pmatrix} -65,1 \\ -231,7 \\ -419,35 \end{pmatrix},$$

cuya solución es

$$(x_0, v_0, a) = (1,26571, 1,87857, -6,01429).$$

En la Figura 9.2 se muestra una superposición de la parábola $x(t) = 1,26571 + 1,87857t - 6,01429/2t^2$ y de la nube de puntos (t_i, x_i) del enunciado.

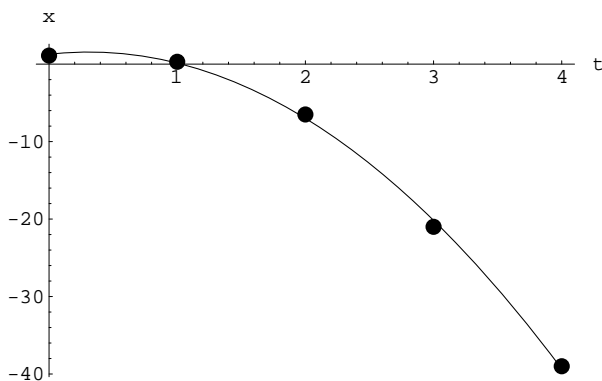


Figura 9.2: Parábola $x(t) = 1,26571 + 1,87857t - 6,01429/2t^2$ y nube de puntos experimentales.

Problema 9.5

La siguiente tabla muestra el volumen v de un gas en un cilindro cuando la presión p en un pistón es aumentada

p_k (Pa)	5	6	7
v_k (m^3)	2.5	1	0.5

- (i) Aproximar, con todos los datos de la tabla adjunta, la función $v(p)$ mediante una interpolación polinómica de Lagrange.
- (ii) Posteriormente se averigua que la evolución termodinámica del gas es adiabática y por lo tanto verifica la ecuación $pv^\gamma = c$ con c y γ constantes. Calcular dichas constantes por el método de mínimos cuadrados a partir de las mediciones experimentales dadas.
- (iii) Supongamos que se realiza una tabulación del volumen v del gas en los $n + 1$ nudos dados por las presiones $3 = p_0 < p_1 < \dots < p_{n-1} < p_n = 5$ y posteriormente se calcula el polinomio de interpolación de Lagrange $P_n(p)$. Se sabe que se tienen $N = 15$ moles de gas a temperatura constante $T = 300$ K y que además suponemos válida la aproximación de gas ideal, es decir, se verifica la ecuación de estado $pv = NRT$ con $R = 8,314$ J/K mol. Acotar el error absoluto Δ_n que se comete cuando se realiza la aproximación $v(p) \approx P_n(p)$ para cualquier valor $p \in [3, 5]$ en el caso de que $n = 3$. Calcular (únicamente en función de n) la anterior acotación de Δ_n para un valor arbitrario de n y obtener $\lim_{n \rightarrow \infty} \Delta_n$.
- (iv) Un modelo de gas más realista donde se tienen en cuenta las interacciones entre sus moléculas así como su tamaño mediante las constantes a y b características del gas viene dado por la ecuación de estado de van der Waals

$$\left(p + \frac{N^2 a}{v^2}\right)(v - Nb) = NRT.$$

Si el gas que se estudia es el Oxígeno, se tiene $a = 137,4$, $b = 0,031$. Suponiendo que $T = 300$, $N = 15$ y $p = 2000$

- (iv.a) Reescribir la ecuación de estado de van der Waals en la forma $f(v) = 0$, siendo f un polinomio cúbico.
- (iv.b) Demostrar que el volumen v^* ocupado por el Oxígeno está en el intervalo $[17, 20]$. Partiendo de dicho intervalo, calcular el número de iteraciones necesarias en el algoritmo de bisección para aproximar v^* con un error absoluto menor que 10^{-5} .
- (iv.c) Aproximar el valor de v^* mediante 2 iteraciones con el método de la secante partiendo de los valores iniciales $v_0 = 17$ y $v_1 = 20$.

Solución. (i) Sea (p_k, v_k) con $k = 0, 1, 2$, las parejas presión-volumen de la

tabulación dada en el enunciado. Según el método de interpolación polinómica de Lagrange, la función $v(p) \approx P_2(p)$, siendo $P_2(p)$ un polinomio de grado menor o igual que 2. La forma explícita de $P_2(p)$ es

$$P_2(p) = \sum_{k=0}^2 v_k \ell_k(p) = v_0 \ell_0(p) + v_1 \ell_1(p) + v_2 \ell_2(p) ,$$

donde $\ell_k(p)$ con $k = 0, 1, 2$, son los llamados polinomios básicos de Lagrange y vienen dados por

$$\ell_k(p) = \prod_{\substack{j=0 \\ j \neq k}}^2 \frac{p - p_j}{p_k - p_j} .$$

En definitiva se tiene

$$\begin{aligned} \ell_0(p) &= \frac{p - p_1}{p_0 - p_1} \frac{p - p_2}{p_0 - p_2} = \frac{p - 6}{5 - 6} \frac{p - 7}{5 - 7} = \frac{1}{2}(p - 6)(p - 7) , \\ \ell_1(p) &= \frac{p - p_0}{p_1 - p_0} \frac{p - p_2}{p_1 - p_2} = \frac{p - 5}{6 - 5} \frac{p - 7}{6 - 7} = -(p - 5)(p - 7) , \\ \ell_2(p) &= \frac{p - p_0}{p_2 - p_0} \frac{p - p_1}{p_2 - p_1} = \frac{p - 5}{7 - 5} \frac{p - 6}{7 - 6} = \frac{1}{2}(p - 5)(p - 6) , \end{aligned}$$

de manera que

$$\begin{aligned} P_2(p) &= 2,5 \times \frac{1}{2}(p - 6)(p - 7) - (p - 5)(p - 7) + 0,5 \times \frac{1}{2}(p - 5)(p - 6) \\ &= \frac{1}{2}(50 - 14p + p^2) . \end{aligned}$$

(ii) En primer lugar observamos que el parámetro γ no interviene de forma lineal en la ecuación $pv^\gamma = c$. Por consiguiente utilizaremos, en primer lugar, una estrategia de linearización de los datos. Para conseguirlo tomamos logaritmos neperianos en la anterior expresión y obtenemos $\ln c - \gamma \ln v = \ln p$, de manera que los parámetros $\ln c$ y γ intervienen de forma lineal. En segundo lugar calculamos la nueva tabulación

$\ln p_k$	1.60944	1.79176	1.94591
$\ln v_k$	0.916291	0	-0.693147

A partir de la linearización de los datos realizada se tiene el sistema lineal sobredeterminado $Ax = b$ siguiente

$$\begin{pmatrix} 1 & -0,916291 \\ 1 & 0 \\ 1 & 0,693147 \end{pmatrix} \begin{pmatrix} \ln c \\ \gamma \end{pmatrix} = \begin{pmatrix} 1,60944 \\ 1,79176 \\ 1,94591 \end{pmatrix} .$$

De la teoría de aproximación por mínimos cuadrados se obtienen las ecuaciones normales $A^T Ax = A^T b$, es decir,

$$\begin{pmatrix} 3 & -0,223144 \\ -0,223144 & 1,32004 \end{pmatrix} \begin{pmatrix} \ln c \\ \gamma \end{pmatrix} = \begin{pmatrix} 5,34711 \\ -0,125914 \end{pmatrix} .$$

La solución de las anteriores ecuaciones normales es

$$(\ln c, \gamma) = (1,79788, 0,208533) ,$$

de lo que se deduce los valores

$$c = \exp(1,79788) = 6,03684 , \quad \gamma = 0,208533 .$$

Concluimos que la mejor aproximación por mínimos cuadrados de la tabla presión-volumen del proceso adiabático del enunciado viene dada por

$$pv^{0,208533} = 6,03684 .$$

(iii) Cuando se realiza la aproximación $v(p) \approx P_n(p)$ para cualquier valor $p \in [3, 5]$, el error absoluto Δ_n cometido es

$$\Delta_n := |v(p) - P_n(p)| = \left| \frac{v^{(n+1)}(\xi_p)}{(n+1)!} \prod_{k=0}^n (p - p_k) \right| ,$$

siendo $\xi_p \in (3, 5)$. En primer lugar se ha de calcular la derivada de orden $n+1$ de la función $v(p)$. Puesto que se supone válida la aproximación de gas ideal, se tiene $v(p) = NRT/p$, que con los datos del enunciado resulta $v(p) = C/p$, siendo la constante $C = NRT = 37413$. Para esta función $v(p)$ resulta fácil calcular por inducción la derivada de orden $n+1$. En concreto, $v'(p) = -Cp^{-2}$, $v''(p) = 2Cp^{-3}$, $v'''(p) = -6Cp^{-4}$, $v^{(4)}(p) = 24Cp^{-5}, \dots, v^{(n)}(p) = (-1)^n n! Cp^{-(n+1)}$. Sustituyendo en la expresión del error absoluto Δ_n se tiene

$$\Delta_n = \left| \frac{(-1)^{n+1} C}{\xi_p^{n+2}} \prod_{k=0}^n (p - p_k) \right| = \frac{37413}{\xi_p^{n+2}} \prod_{k=0}^n |p - p_k| .$$

Veamos como acotar esta expresión. En primer lugar, puesto que $\xi_p \in (3, 5)$, es evidente que $1/\xi_p^{n+2} < 1/3^{n+2}$ para cualquier $n \in \mathbb{N}$. En segundo lugar, como $p \in [3, 5]$ y $p_k \in [3, 5]$ para cualquier $k = 0, 1, \dots, n$, se deduce que $|p - p_k| \leq 5 - 3 = 2$ para $k = 0, 1, \dots, n$ y por lo tanto $\prod_{k=0}^n |p - p_k| \leq 2^{n+1}$. En definitiva, hemos probado la siguiente cadena de desigualdades

$$\Delta_n = \frac{37413}{\xi_p^{n+2}} \prod_{k=0}^n |p - p_k| < \frac{37413}{3^{n+2}} \prod_{k=0}^n |p - p_k| \leq \frac{37413}{3^{n+2}} 2^{n+1} .$$

Una vez se tiene la acotación de Δ_n para cualquier n , es directo ver que si $n = 3$ entonces $\Delta_3 < 66512/27 = 2463,41$. Además, puesto que $\lim_{n \rightarrow \infty} (2/3)^{n+1} = 0$, a partir de la acotación de Δ_n es obvio que

$$\lim_{n \rightarrow \infty} \Delta_n = 0 .$$

(iv.a) Multiplicando la ecuación de estado de van der Waals por v^2 se obtiene $(pv^2 + N^2 a)(v - Nb) = NRTv^2$. Así, pasando todos los términos de esta ecuación a un lado de la igualdad y agrupando por potencias de v se tiene

$$f(v) = pv^3 - N(bp + RT)v^2 + N^2 av - N^3 ab = 0 .$$

(iv.b) Sustituyendo los datos del enunciado se tiene que

$$f(v) = -14375,5 + 30915v - 38343v^2 + 2000v^3 = 0 .$$

Puesto que f es un polinomio, en particular $f \in C[17, 20]$. Además $f(17) = -743947 < 0$, $f(20) = 1,26672 \times 10^6 > 0$ de lo que se deduce, a través del Teorema de Bolzano, que existe al menos un $v^* \in (17, 20)$ tal que $f(v^*) = 0$. Obviamente este cálculo no ha “demostrado” que el volumen v^* ocupado por el Oxígeno está en el intervalo $[17, 20]$, pero una comparación con el resultado que proporciona la aproximación de gas ideal muestra que el rango de valores es el adecuado físicamente.

Usaremos el algoritmo de bisección para aproximar el cero v^* de la función f que pertenece al intervalo $[a_0, b_0] = [17, 20]$. Sabemos que el error absoluto cometido en n iteraciones del algoritmo de bisección, es decir aproximando $v^* \approx c_n$, está acotado por

$$|v^* - c_n| \leq \frac{|b_0 - a_0|}{2^{n+1}} = \frac{3}{2^{n+1}} .$$

Si imponemos un error menor que 10^{-5} tenemos

$$|v^* - c_n| \leq \frac{3}{2^{n+1}} < 10^{-5} ,$$

de lo que se deduce $2^{n+1} > 3 \times 10^5$ o equivalentemente

$$n > \log_2(3 \times 10^5) - 1 = \frac{\log(3 \times 10^5)}{\log 2} - 1 = \frac{5 + \log 3}{\log 2} - 1 = 17,1946 .$$

Podemos concluir pues que son necesarias al menos 18 iteraciones.

(iv.c) El método de la secante para aproximar las soluciones de la ecuación $f(v) = 0$ consiste en calcular la sucesión de números reales $\{v_n\}_{n \in \mathbb{N}}$ generada a partir de la recurrencia siguiente

$$v_{n+1} = v_n - f(v_n) \frac{v_n - v_{n-1}}{f(v_n) - f(v_{n-1})} .$$

Partiendo de los valores iniciales $v_0 = 17$ y $v_1 = 20$ se obtiene la siguiente tabla

n	v_n
0	17
1	20
2	18.3111
3	18.3516

Concluimos que el volumen v^* que ocupa el Oxígeno es aproximadamente $v^* \approx v_3 = 18,3516$.

Problema 9.6 Consideremos una masa m restringida a moverse con fricción despreciable y gravedad g constante en una circunferencia vertical de radio ℓ . Supongamos además que la masa está sujeta a la posición más alta de la circunferencia mediante un muelle de longitud natural ℓ y constante elástica k . Sea $\theta(t)$ el ángulo que forma la recta que une la masa y el centro de la circunferencia con la vertical en función del tiempo.

- (i) Es fácil mostrar que la ecuación diferencial del movimiento de la masa m viene dada por

$$m\ell\ddot{\theta} = -mg \sin \theta + k\ell \left(2 \cos \frac{\theta}{2} - 1 \right) \sin \frac{\theta}{2}.$$

La estabilidad del punto de equilibrio $(\theta, \dot{\theta}) = (0, 0)$ depende de dos contribuciones: intuitivamente, el sistema sugiere que si el muelle es blando (k pequeña) y la masa m es grande entonces el punto de equilibrio es estable y viceversa, es decir, k grande y m pequeña convertirá el punto de equilibrio en inestable.

Si $\ell = k = 1$, $m = 100$ y la masa se encuentra inicialmente en reposo y formando un ángulo $\theta(0) = \pi/4$, usar un paso del método RK4 con una longitud de paso $h = 1/5$ para verificar las hipótesis de estabilidad establecidas.

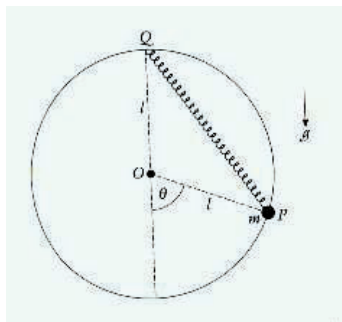
- (ii) Suponiendo que el muelle verifica la ley de Hooke, geometría elemental muestra que la fuerza F que ejerce el muelle es

$$F = k\ell \left(2 \cos \frac{\theta}{2} - 1 \right).$$

Si $\ell = 1$ y se han realizado las siguientes mediciones

F (N)	1.1	2.6	3.2	4.6
θ (rad.)	0.71	0.63	0.42	0.22

aproximar k por el método de mínimos cuadrados.



Solución. (i) Tomando $\ell = k = 1$, $m = 100$ y $g = 9,8$, la ecuación diferencial del movimiento de la masa es

$$100\ddot{\theta} = -100g \sin \theta + \left(2 \cos \frac{\theta}{2} - 1\right) \sin \frac{\theta}{2}.$$

Utilizando la velocidad angular $w = \dot{\theta}$ de la masa como variable auxiliar, convertimos la ecuación diferencial de segundo orden en el siguiente sistema de primer orden

$$\begin{aligned}\dot{\theta} &= f(t, \theta, w) = w, \\ \dot{w} &= g(t, \theta, w) = -9,8 \sin \theta + \frac{1}{100} \left(2 \cos \frac{\theta}{2} - 1\right) \sin \frac{\theta}{2},\end{aligned}$$

con la condición inicial $(\theta(0), w(0)) = (\pi/4, 0)$. Puesto que $f = f(w)$ sólo depende de w y $g = g(\theta)$ sólo de θ , el método RK4 reduce a

$$\begin{pmatrix} \theta_{i+1} \\ w_{i+1} \end{pmatrix} = \begin{pmatrix} \theta_i \\ w_i \end{pmatrix} + \frac{h}{6} \begin{pmatrix} r_1 + 2r_2 + 2r_3 + r_4 \\ k_1 + 2k_2 + 2k_3 + k_4 \end{pmatrix},$$

siendo

$$\begin{aligned}r_1 &= f(w_i), \quad k_1 = g(\theta_i), \\ r_2 &= f\left(w_i + \frac{h}{2}k_1\right), \quad k_2 = g\left(\theta_i + \frac{h}{2}r_1\right), \\ r_3 &= f\left(w_i + \frac{h}{2}k_2\right), \quad k_3 = g\left(\theta_i + \frac{h}{2}r_2\right), \\ r_4 &= f(w_i + hk_3), \quad k_4 = g(\theta_i + hr_3).\end{aligned}$$

Tomando la longitud de paso $h = 1/5$ y la condición inicial $(\theta(0), w(0)) = (\theta_0, w_0) = (\pi/4, 0)$ se tiene, para la primera iteración,

$$\begin{aligned}r_1 &= f(0) = 0, \quad k_1 = g(\pi/4) = -6,9264, \\ r_2 &= f\left(0 + \frac{1}{10}k_1\right) = -0,69264, \quad k_2 = g\left(\pi/4 + \frac{1}{10}r_1\right) = -6,9264, \\ r_3 &= f\left(0 + \frac{1}{10}k_2\right) = -0,69264, \quad k_3 = g\left(\pi/4 + \frac{1}{10}r_2\right) = -6,43038, \\ r_4 &= f\left(0 + \frac{1}{5}k_3\right) = -1,28608, \quad k_4 = g\left(\pi/4 + \frac{1}{5}r_3\right) = -5,90353,\end{aligned}$$

de manera que

$$\begin{pmatrix} \theta_1 \\ w_1 \end{pmatrix} = \begin{pmatrix} \pi/4 \\ 0 \end{pmatrix} + \frac{1}{30} \begin{pmatrix} -4,05664 \\ -39,5435 \end{pmatrix} = \begin{pmatrix} 0,650177 \\ -1,31812 \end{pmatrix}.$$

Notar que $\theta_1 < \theta_0$ y $w_1 < 0$, lo cual confirma numéricamente que la masa se desplaza hacia el punto de equilibrio estable.

(ii) Como $\ell = 1$, la relación entre F y θ es

$$F = k \left(2 \cos \frac{\theta}{2} - 1 \right) .$$

Imponiendo la tabulación

F (N)	1.1	2.6	3.2	4.6
θ (rad.)	0.71	0.63	0.42	0.22

a dicha relación obtenemos el sistema lineal sobredeterminado

$$\begin{pmatrix} 0,875293 \\ 0,901593 \\ 0,956062 \\ 0,987912 \end{pmatrix} k = \begin{pmatrix} 1,1 \\ 2,6 \\ 3,2 \\ 4,6 \end{pmatrix} .$$

Multiplicamos por la izquierda ambos miembros de esta ecuación por la matriz traspuesta de la matriz de coeficientes del sistema obteniendo de esta forma la ecuación normal

$$\begin{pmatrix} 0,875293 \\ 0,901593 \\ 0,956062 \\ 0,987912 \end{pmatrix}^T \begin{pmatrix} 0,516724 \\ 0,616055 \\ 0,826178 \\ 0,951795 \end{pmatrix} k = \begin{pmatrix} 0,875293 \\ 0,901593 \\ 0,956062 \\ 0,987912 \end{pmatrix}^T \begin{pmatrix} 1,1 \\ 2,6 \\ 3,2 \\ 4,6 \end{pmatrix} ,$$

es decir,

$$3,469k = 10,9108 ,$$

con lo que $k = 3,14523$.

Problema 9.7 Consideremos la aproximación de la función $f(x) = x/2$ definida en el intervalo $[-\pi, \pi]$ por una expresión del tipo

$$f^*(x) = c_1 + c_2 \cos(x) + c_3 \sin(x) .$$

Calcular los valores de los coeficientes reales c_j de modo que la aproximación se realice dentro del contexto de mínimos cuadrados.

Solución. Deseamos realizar una aproximación $f(x) \approx f^*(x)$ por mínimos cuadrados en el intervalo $[-\pi, \pi]$. En primer lugar notamos que la expresión de la función aproximadora es de la forma $f^*(x) = \sum_{i=1}^3 c_i \varphi_i(x)$, siendo las funciones $\varphi_1(x) = 1$, $\varphi_2(x) = \cos x$ y $\varphi_3(x) = \sin(x)$ linealmente independientes en el intervalo $[-\pi, \pi]$. Las ecuaciones normales (sistema lineal de ecuaciones que verifican los coeficientes c_i) serán de la forma

$$\begin{pmatrix} \langle \varphi_1, \varphi_1 \rangle & \langle \varphi_1, \varphi_2 \rangle & \langle \varphi_1, \varphi_3 \rangle \\ \langle \varphi_2, \varphi_1 \rangle & \langle \varphi_2, \varphi_2 \rangle & \langle \varphi_2, \varphi_3 \rangle \\ \langle \varphi_3, \varphi_1 \rangle & \langle \varphi_3, \varphi_2 \rangle & \langle \varphi_3, \varphi_3 \rangle \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} \langle f, \varphi_1 \rangle \\ \langle f, \varphi_2 \rangle \\ \langle f, \varphi_3 \rangle \end{pmatrix} ,$$

donde \langle, \rangle denota el producto escalar euclideo continuo en el intervalo $[-\pi, \pi]$. Se tiene pues que

$$\begin{aligned}\langle \varphi_1, \varphi_1 \rangle &= \int_{-\pi}^{\pi} dx = 2\pi, \\ \langle \varphi_2, \varphi_2 \rangle &= \int_{-\pi}^{\pi} \cos^2(x) dx = \pi \\ \langle \varphi_3, \varphi_3 \rangle &= \int_{-\pi}^{\pi} \sin^2(x) dx = \pi \\ \langle \varphi_1, \varphi_2 \rangle &= \int_{-\pi}^{\pi} \cos(x) dx = 0 \\ \langle \varphi_1, \varphi_3 \rangle &= \int_{-\pi}^{\pi} \sin(x) dx = 0 \\ \langle \varphi_2, \varphi_3 \rangle &= \int_{-\pi}^{\pi} \sin(x) \cos(x) dx = 0 \\ \langle f, \varphi_1 \rangle &= 1/2 \int_{-\pi}^{\pi} x dx = 0 \\ \langle f, \varphi_2 \rangle &= 1/2 \int_{-\pi}^{\pi} x \cos(x) dx = 0 \\ \langle f, \varphi_3 \rangle &= 1/2 \int_{-\pi}^{\pi} x \sin(x) dx = \pi.\end{aligned}$$

Las integrales que aparecen en los anteriores productos escalares son, o bien inmediatas, o bien (segunda y tercera) se calculan con las identidades trigonométricas

$$\sin^2(x) = \frac{1 - \cos(2x)}{2}, \quad \cos^2(x) = \frac{1 + \cos(2x)}{2},$$

excepto la última que se calcula mediante la fórmula de integración por partes. De hecho hay muchas de esas integrales (desde la quinta hasta la octava incluidas) que valen cero puesto que la función integrando es impar y el intervalo de integración es simétrico respecto del origen.

Finalmente, como el producto escalar verifica la propiedad conmutativa, la matriz de Gram es diagonal y las ecuaciones normales tienen la forma

$$\begin{pmatrix} 2\pi & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & \pi \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \pi \end{pmatrix}.$$

La solución del sistema es $(c_1, c_2, c_3) = (0, 0, 1)$, de modo que la aproximación es

$$\frac{x}{2} \approx \sin(x).$$

Problema 9.8 La siguiente tabla muestra la evolución de una enfermedad contagiosa en función del tiempo t . Aquí, c indica el número de contagiados.

t	0	3	5	7
c	1	20	22	23

- (i) Aproximar, mediante interpolación polinómica con el método de Lagrange el valor de $c(4)$.
- (ii) En la tabulación dada de $c(t)$ se observa que dicha función parece saturarse a medida que pasa el tiempo, es decir, presenta una asíntota horizontal. Explica entonces por qué el polinomio obtenido en el apartado anterior no resulta adecuado para aproximar valores de c en tiempos superiores a 7.
- (iii) Una forma de resolver el problema del apartado (ii) consiste en ajustar los datos de la tabulación a un modelo de crecimiento logístico

$$c(t) = \frac{1}{\alpha + \beta \exp(-3t)} .$$

Calcula el valor de los parámetros α y β mediante el método de mínimos cuadrados.

Solución. (i) Aproximaremos $c(t) \approx P_3(t)$ siendo $P_3(t)$ el polinomio de grado menor o igual que 3 que interpola a los datos de la tabla adjunta. Denotando los nodos de interpolación por $\{t_0, t_1, t_2, t_3\} = \{0, 3, 5, 7\}$ y usando la fórmula de interpolación de Lagrange se tiene que

$$P_3(t) = \sum_{k=0}^3 c_k \ell_k(t) ,$$

siendo $c_k = c(t_k)$ y los polinomios básicos de Lagrange

$$\ell_k(t) = \prod_{\substack{j=0 \\ j \neq k}}^n \frac{t - t_j}{t_k - t_j}, \quad k = 0, 1, 2, 3.$$

En concreto, se tiene que

$$\begin{aligned} \ell_0(t) &= \frac{t-3}{0-3} \frac{t-5}{3-5} \frac{t-7}{5-7} , \\ \ell_1(t) &= \frac{t-0}{3-0} \frac{t-5}{3-5} \frac{t-7}{5-7} , \\ \ell_2(t) &= \frac{t-0}{5-0} \frac{t-3}{5-3} \frac{t-7}{5-7} , \\ \ell_3(t) &= \frac{t-0}{7-0} \frac{t-3}{7-3} \frac{t-5}{7-5} . \end{aligned}$$

De este modo obtenemos la siguiente aproximación

$$c(4) \approx P_3(4) = c_0 \ell_0(4) + c_1 \ell_1(4) + c_2 \ell_2(4) + c_3 \ell_3(4)$$

$$\begin{aligned}
&= \ell_0(4) + 20\ell_1(4) + 22\ell_2(4) + 23\ell_3(4) \\
&= -\frac{1}{35} + 20 \times \frac{1}{2} + 22 \times \frac{3}{5} - 23 \times \frac{1}{14} = \frac{1507}{70} = 21,5286 .
\end{aligned}$$

(ii) Es bien sabido que los polinomios no pueden tener asíntotas horizontales, de modo que crecen o decrecen indefinidamente cuando t es suficientemente grande. Este es el motivo por el cual la aproximación $c(t) \approx P_3(t)$ no es adecuada para t grandes.

(iii) Consideremos el modelo de crecimiento logístico

$$c(t) = \frac{1}{\alpha + \beta \exp(-3t)} .$$

Como los parámetros α y β intervienen de forma no lineal en la expresión de $c(t)$, antes de utilizar el método de mínimos cuadrados para hallarlos realizaremos una estrategia de linearización de datos. En el caso que nos ocupa basta con invertir la ecuación logística de la forma

$$\frac{1}{c(t)} = \alpha + \beta \exp(-3t) .$$

Realizamos la nueva tabulación

$\exp(-3t)$	1	$\exp(-9)$	$\exp(-15)$	$\exp(-21)$
$1/c$	1	1/20	1/22	1/23

e imponemos dicha tabulación sobre nuestro modelo linearizado para llegar al siguiente sistema lineal sobredeterminado

$$\begin{pmatrix} 1 & 1 \\ 1 & \exp(-9) \\ 1 & \exp(-15) \\ 1 & \exp(-21) \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ 1/20 \\ 1/22 \\ 1/23 \end{pmatrix} .$$

Escribimos el anterior sistema lineal en la forma vectorial $Ax = b$, siendo el vector incógnita $x = (\alpha, \beta)^T$ y

$$A = \begin{pmatrix} 1 & 1 \\ 1 & \exp(-9) \\ 1 & \exp(-15) \\ 1 & \exp(-21) \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1/20 \\ 1/22 \\ 1/23 \end{pmatrix} .$$

Entonces, el valor de x buscado se obtiene resolviendo las ecuaciones normales $A^T Ax = A^T b$, es decir,

$$\begin{pmatrix} 4 & 1,00012 \\ 1,00012 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1,13893 \\ 1,00001 \end{pmatrix} ,$$

cuya solución es

$$\alpha = 0,0462715, \beta = 0,953729.$$

En definitiva, la mejor aproximación viene dada por

$$c(t) = \frac{1}{0,0462715 + 0,953729 \exp(-3t)}.$$

Problema 9.9 Demostrar que la recta de regresión asociada a cualquier nube de puntos $(x_i, y_i) \in \mathbb{R}^2$ con $i = 1, \dots, m$, con $m \geq 2$, pasa por el punto $(\bar{x}, \bar{y}) \in \mathbb{R}^2$, siendo \bar{x} e \bar{y} los valores medios de $\{x_1, \dots, x_m\}$ y $\{y_1, \dots, y_m\}$ respectivamente.

Solución. Sea $y = f(x) = ax + b$ la recta de regresión asociada a la nube de puntos $(x_i, y_i) \in \mathbb{R}^2$ con $i = 1, \dots, m$. Calculemos en primer lugar los coeficientes $a, b \in \mathbb{R}$ de esta recta. Hemos de imponer que $f(x_i) = ax_i + b = y_i$ para todo $i = 1, \dots, m$. Obtenemos de este modo el sistema lineal sobredeterminado

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Escribimos las ecuaciones normales asociadas

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Multiplicando las matrices del sistema se tiene que (a, b) es la solución de

$$\begin{pmatrix} \sum_i x_i^2 & \sum_i x_i \\ \sum_i x_i & m \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_i x_i y_i \\ \sum_i y_i \end{pmatrix}, \quad (9.5)$$

donde en todos los sumatorios el subíndice i corre desde 1 hasta m . Finalmente, para demostrar que la recta de regresión $y = f(x) = ax + b$ pasa por el punto (\bar{x}, \bar{y}) , hemos de demostrar que $f(\bar{x}) = \bar{y}$, es decir, hemos de ver que $\bar{y} = a\bar{x} + b$ siendo (a, b) la solución del sistema lineal (9.5). La forma más simple de demostrarlo consiste en tomar la última ecuación del sistema (9.5), es decir, $a \sum_i x_i + mb = \sum_i y_i$. Entonces, dividiendo por m ambos miembros de esta ecuación y recordando que por definición de valor medio se tiene $\bar{x} = (\sum_i x_i)/m$, $\bar{y} = (\sum_i y_i)/m$, se obtiene el resultado deseado $a\bar{x} + b = \bar{y}$.

Capítulo 10

Introducción a la Programación

La presente sección está dedicada a la confección de pequeños programas en los lenguajes de programación FORTRAN y MATHEMATICA con el objetivo de iniciarse en sus peculiaridades.

10.1. Velocidad del Sonido en el Aire en Función de la Temperatura

En esta práctica veremos un ejemplo de tabulación de una función cuando su variable independiente recorre un determinado intervalo. En concreto se sabe que, bajo ciertas hipótesis, la velocidad del sonido en el aire en función de la temperatura viene expresada por

$$v(T) = 331\sqrt{1 + \frac{T}{273}} , \quad (10.1)$$

donde la temperatura T está en $^{\circ}C$ y la velocidad v en m/s .

Un código FORTRAN para obtener una tabulación de los pares (T, v) en el rango de temperaturas $20^{\circ}C \leq T \leq 35^{\circ}C$ con incrementos de $1^{\circ}C$ puede ser el dado en la Figura 10.1. Un código MATHEMATICA con el mismo objetivo es el siguiente.

```
v[T_] = 331*Sqrt[1 + T/273];  
  
Do[  
  Print["Para T = ", T, "la velocidad del sonido es ", v[T]],  
  {T, 20, 35, 1}  
]
```

```

PROGRAM VELOCIDAD
REAL T,DT,V
T=20.
DT=1.
10 V=331.*SQRT(1.+T/273.)
WRITE(*,*) ' Para T= ',T,' La velocidad del sonido es ', V
T=T+DT
IF(T .LT. 36.0) THEN
    GO TO 10
ELSE
    WRITE(*,*) 'Cálculo terminado'
    STOP
END IF
END

```

Figura 10.1: Tabulación de la velocidad del sonido para diferentes temperaturas.

10.2. Ecuación de Segundo Grado

Consideremos la ecuación de segundo grado

$$ax^2 + bx + c = 0 .$$

Sabemos que la naturaleza de las raíces depende del valor del discriminante $\delta = b^2 - 4ac$. En concreto tenemos

- (i) si $\delta > 0 \implies$ dos raíces reales y diferentes $x_{\pm} = \frac{1}{2a} (-b \pm \sqrt{\delta})$.
- (ii) si $\delta = 0 \implies$ una única raíz real doble $x_+ = x_- = -\frac{b}{2a}$.
- (iii) si $\delta < 0 \implies$ dos raíces complejas conjugadas $x_{\pm} = \frac{1}{2a} (-b \pm i\sqrt{-\delta})$.

El objetivo de esta práctica es realizar un programa para calcular la raíz real (si existe) de menor valor absoluto de una ecuación de segundo grado.

Observar que la raíz real ($\delta \geq 0$) de menor valor absoluto es

$$x_{\min} = \begin{cases} \frac{1}{2a} (-b - \sqrt{\delta}) & \text{si } b < 0 \\ \frac{1}{2a} (-b + \sqrt{\delta}) & \text{si } b \geq 0 \end{cases}$$

Un posible código FORTRAN para calcular la raíz real de menor valor absoluto de una ecuación de segundo grado puede ser el dado en la Figura 10.2. El análogo código MATHEMATICA es, por ejemplo, el siguiente.

```
a = 1; b = 8; c = 5;
```

```
discr = b^2 - 4*a*c;
```

```
If[discr < 0, Print["Raíces complejas "]; Interrupt[],
  T1 = -b/(2*a);
  T2 = Sqrt[discr]/(2*a);
  If[b >= 0, xmin = T1 + T2;, xmin = T1 - T2;]
]
```

```
Print["La raíz de menor valor absoluto es ", xmin]
```

```
PROGRAM CUADRATICO
REAL A,B,C,DISCR,DENOM,T1,T2,XMIN
WRITE(*,*) ' Introduce los coeficientes separados por comas '
READ(*,*) A,B,C
DISCR=B*B-4.*A*C
IF(DISCR .LT. 0) THEN
    WRITE(*,*) 'Raíces Complejas'
    STOP
ELSE
    DENOM=2.*A
    T1=-B/DENOM
    T2=SQRT(DISCR)/DENOM
    IF(B .GE. 0) THEN
        XMIN=T1+T2
    ELSE
        XMIN=T1-T2
    END IF
END IF
WRITE(*,*) ' La raíz real de menor valor absoluto es: ',XMIN
STOP
END
```

Figura 10.2: Raíz real de menor valor absoluto de una ecuación de segundo grado.

10.3. Método de Newton: Cálculo de Raíces Cuadradas

El siguiente procedimiento para calcular raíces cuadradas, utilizado por Isaac Newton¹, es un claro ejemplo de cálculo iterativo. Se puede demostrar que la sucesión de números reales $\{x_n\}$ con $n \in \mathbb{N}$ definida de forma recurrente mediante

$$x_{n+1} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right), \quad (10.2)$$

converge hacia \sqrt{a} , es decir $\lim_{n \rightarrow \infty} x_n = \sqrt{a}$, si tomamos el valor inicial x_0 próximo a la raíz cuadrada de a .

Como caso particular, vamos a demostrar que la sucesión

$$x_{n+1} = \frac{1}{2} \left(x_n + \frac{2}{x_n} \right), \quad x_0 = 2,$$

converge hacia $\sqrt{2}$. En primer lugar, observamos que $x_0 = 2$, $x_1 = 3/2 = 1,5$, $x_2 = 17/12 \approx 1,416$, de manera que parece que la sucesión es decreciente. Comprobémoslo de una manera formal, es decir, veamos que $x_n - x_{n+1} \geq 0$ para todo $n \in \mathbb{N}$. Notemos que, debido a la expresión de la fórmula de recurrencia y puesto que $x_0 = 2 > 0$, se tiene $x_n > 0$ para todo $n \in \mathbb{N}$. Entonces, como

$$x_n - x_{n+1} = x_n - \frac{1}{2} \left(x_n + \frac{2}{x_n} \right) = \frac{x_n^2 - 2}{2x_n},$$

es fácil ver que $x_n - x_{n+1} \geq 0$ si y sólo si $x_n^2 - 2 \geq 0$, es decir $x_n \geq \sqrt{2}$. Esta última desigualdad puede ser demostrada por *reducción al absurdo*² de la forma siguiente: Supongamos lo contrario, es decir supongamos que $x_n < \sqrt{2}$ para algún valor de n . Entonces, a partir de la recurrencia se tiene que $x_n = 1/2[x_{n-1} + 2/x_{n-1}] < \sqrt{2}$. Operando en esta desigualdad se llega a que $x_{n-1}^2 + 2 < 2\sqrt{2}x_{n-1}$, o de forma equivalente $[x_{n-1} - \sqrt{2}]^2 < 0$, lo que es imposible. Finalmente, concluimos que $x_n \geq \sqrt{2}$ para todo $n \in \mathbb{N}$ y por lo tanto la sucesión $\{x_n\}$ es decreciente.

Por otra parte, como la sucesión $\{x_n\}$ es decreciente y además está acotada inferiormente, puesto que $x_n \geq \sqrt{2}$, se concluye que dicha sucesión tiene límite,

¹Este método de aproximación de la raíz cuadrada es muy antiguo. Se atribuye generalmente a Heron de Alejandría, quien vivió en el siglo I, pero probablemente es anterior. Naturalmente, no fue obtenido a partir de lo que hoy se conoce como el método de Newton-Raphson (descrito en secciones posteriores), sino por un razonamiento que muestra bien la estructura del pensamiento matemático griego: si x_n es una aproximación por exceso de \sqrt{a} , entonces a/x_n es una aproximación por defecto y recíprocamente. Tomando la media aritmética entre una aproximación por defecto y una por exceso se obtiene una aproximación que es mejor que las anteriores.

²Recordemos que el método de reducción al absurdo consiste en suponer cierto lo contrario de lo que se quiere demostrar y comprobar que esta suposición implica algún tipo de absurdo.

es decir $\lim_{n \rightarrow \infty} x_n = \ell$. Dicho límite ℓ se puede calcular tomando el paso al límite cuando $n \rightarrow \infty$ en la relación de recurrencia, de manera que se obtiene

$$\ell = \frac{1}{2} \left(\ell + \frac{2}{\ell} \right),$$

o de forma equivalente $\ell = \sqrt{2}$.

La idea del programa que lleve a cabo el algoritmo de Newton (10.2) ha de partir de un número a del cual queramos calcular su raíz cuadrada y de un número x_0 próximo al valor \sqrt{a} . Entonces, el resultado de la primera iteración x_1 es una mejor aproximación al valor \sqrt{a} . Si el valor x_1 no es lo suficientemente bueno (es decir $x_1^2 - a$ no es pequeño) realizamos una nueva iteración y calculamos x_2 y así sucesivamente hasta obtener un resultado aceptable o hasta que se supere un número de iteraciones máximo preestablecido.

NOTA: Todos los programas basados en un método iterativo deben estar preparados para hacer frente a una serie de problemas específicos de estos algoritmos, entre los más comunes se encuentran:

- El proceso diverge. Es decir, en lugar de acercarnos a la solución nos alejamos de ella.
- Aunque exista convergencia, ésta es demasiado lenta. Es decir, después de efectuar un cierto número de iteraciones prefijado, la solución no es lo suficientemente aproximada.

Supongamos que se permite una *tolerancia* de 10^{-6} , es decir paramos las iteraciones si se verifica la condición $|x_n^2 - a| \leq 10^{-6}$. Tomando como primera aproximación al número \sqrt{a} el valor $x_0 = (a + 1)/2$ y permitiendo un máximo de 100 iteraciones, un posible código FORTRAN es el dado en la Figura 10.3. El código MATHEMATICA es el siguiente.

```
a = 2;

x = (a + 1)/2;

For[i = 1, i <= 100, i++,
  x = (x + a/x)/2;
  If[Abs[x^2 - a] < 10^(-6), Print["La raíz cuadrada de ", a, " es
", x]; Interrupt[]]
]

Print["No hay convergencia en 100 iteraciones"]
```

```

PROGRAM NEWTON
REAL A,X
INTEGER I
WRITE(*,*) ' Introduce el número cuya raíz queremos hallar '
READ(*,*) A
I=0
X=(A+1.)/2.
1  X=0.5*(X+A/X)
  I=I+1
  IF(I. GT. 100) THEN
      WRITE(*,*) ' No hay convergencia en 100 iteraciones '
      STOP
  ELSE IF (ABS(X*X-A) .LT. 1.E-6) THEN
      WRITE(*,*) ' La raíz cuadrada de ', A, ' es ', X
      STOP
  ELSE
      GO TO 1
  END IF
END

```

Figura 10.3: Algoritmo de Newton para el cálculo de la raíz cuadrada de a .

10.4. Algoritmo de Arquímedes y el Número π

Uno de los primeros que calculó el número π fue Arquímedes quien comparó la longitud de una circunferencia con el perímetro de un polígono regular inscrito en dicha circunferencia.

Supongamos un círculo de radio unidad en el cual inscribimos un cuadrado (polígono regular de 2^2 lados). Entonces, el lado del cuadrado vale

$$\ell_2 = 2 \sin \left(\frac{\pi}{4} \right) = \sqrt{2} ,$$

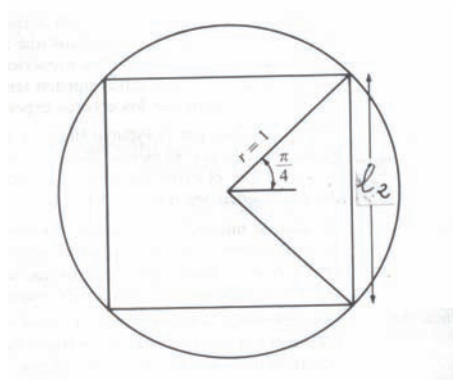
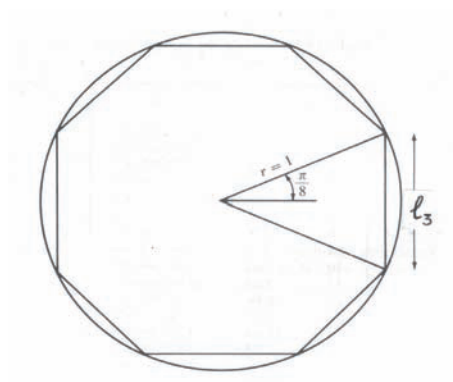
de manera que el perímetro del cuadrado es

$$P_2 = 2^2 \ell_2 = 4\sqrt{2} ,$$

que es la primera aproximación a la longitud de la circunferencia cuyo valor es 2π .

Decimos que una aproximación p_k de orden k -ésimo al número π es la mitad del perímetro de un polígono regular de 2^k lados inscrito en la circunferencia unidad, es decir

$$p_k := \frac{1}{2} P_k = \frac{1}{2} 2^k \ell_k = \frac{1}{2} 2^k 2 \sin \theta_k = 2^k \sin \theta_k , \quad (10.3)$$

Figura 10.4: Aproximación con $k = 2$.Figura 10.5: Aproximación con $k = 3$.

siendo el ángulo $\theta_k = \frac{\pi}{2^k}$, ver las Figuras 10.4 y 10.5. De la ecuación (10.3) observamos que si doblamos el número de lados, es decir k pasa a $k+1$, se tiene

$$p_{k+1} = 2^{k+1} \sin\left(\frac{\theta_k}{2}\right). \quad (10.4)$$

Realizando ahora un poco de trigonometría

$$\sin^2\left(\frac{\theta_k}{2}\right) = \frac{1}{2}(1 - \cos \theta_k) = \frac{1}{2}(1 - \sqrt{1 - \sin^2 \theta_k}),$$

de manera que introduciendo esta igualdad trigonométrica en la ecuación (10.4) se obtiene

$$p_{k+1} = 2^{k+1} \sqrt{\frac{1}{2} \left(1 - \sqrt{1 - \sin^2 \theta_k}\right)}.$$

Finalmente, despejando $\sin \theta_k$ de la ecuación (10.3) e introduciéndolo en la anterior ecuación obtenemos una relación de recurrencia para las aproximaciones p_k al número π de la forma

$$p_{k+1} = \sqrt{2}2^k \sqrt{1 - \sqrt{1 - \left(\frac{p_k}{2^k}\right)^2}}. \quad (10.5)$$

La práctica consiste en:

1. Programar la iteración (10.5) para distintos valores de $k = 2, 3, 4, \dots$.
2. Observar el comportamiento del algoritmo según va aumentando el valor k .

El código FORTRAN de la Figura 10.6 implementa el algoritmo de Arquímedes (10.5) para aproximar al número π . El subíndice k toma los valores $k = 2, 3, \dots, 15$. La salida del algoritmo son la parejas (k, p_k) . De forma análoga, el código MATHEMATICA es el siguiente.

```
k = 2;

P = 2*Sqrt[2];

While[k <= 15,
  Print[k, " ", " ", N[P]]
  P = Sqrt[2]*2^k*Sqrt[1 - Sqrt[1 - (P/2^k)^2]];
  k = k + 1;
]
```

```
PROGRAM PI
REAL P
INTEGER K
K=2
P=2*SQRT(2)
DO WHILE(K .LE. 15)
    WRITE(*,*) K,P
    P=SQRT(2)*2**K*SQRT(1-SQRT(1-(P/2**K)**2))
    K=K+1
END DO
END
```

Figura 10.6: Algoritmo de Arquímedes.

Una vez ejecutado el programa, obtenemos el resultado siguiente.

k	p_k
2	2.828427
3	3.061467
4	3.121445
5	3.136549
6	3.140331
7	3.141277
8	3.141514
9	3.141573
10	3.141588
11	3.141592
12	3.141593
13	3.141593
14	3.141593
15	3.141593

10.5. El problema $3x + 1$

Vamos a describir un problema no resuelto en matemáticas conocido como el problema $3x + 1$. A partir de un número natural $x \in \mathbb{N}$ cualquiera, se construye una sucesión de números naturales siguiendo el siguiente procedimiento:

1. Si x es par, el siguiente término de la sucesión es $x/2$.
2. Si x es impar, el siguiente término de la sucesión es $3x + 1$.

Por ejemplo, si la sucesión comienza en $x = 6$, se obtiene la secuencia: 6, 3, 10, 5, 16, 8, 4, 2, 1, 4, 2, 1,.... Observar que, a partir de ahora los números 4, 2, 1 se repiten indefinidamente. Hasta el momento, todos los que han probado construir esta sucesión a partir de algún número natural han observado que siempre se acaba en 4, 2, 1. A veces se puede tardar un poco en llegar (tomar por ejemplo el valor inicial $x = 27$). Entonces, se ha conjeturado que siempre ocurrirá el comportamiento descrito anteriormente independientemente de qué número natural inicial x sea tomado. La conjetura $3x + 1$ es muy fácil de enunciar, pero hasta la actualidad no se ha podido verificar si es cierta o falsa a pesar de haber sido atacada por eminentes matemáticos. Parece ser que el primero en formular la conjetura fue el matemático Lothar Collatz de la Universidad de Hamburg en el año 1930. Actualmente hay más de 150 artículos publicados sobre la conjetura $3x + 1$, ver por ejemplo la página web <http://www.cecm.sfu.ca/organics/papers/lagarias/>. El año 1999, Oliveira y Silva prueban que la conjetura es cierta para números naturales x menores o iguales que los del orden de 10^{16} .

El siguiente programa en OCTAVE genera la secuencia del problema $3x + 1$ comenzando en un número $x \in \mathbb{N}$ especificado. Si se alcanza un término de

la sucesión con valor 1, el programa finaliza. La sucesión se almacena en las componentes del vector s .

```
function s=sucesion(x)
s=x;
while(x>1)
if(rem(x,2)==0)
x=x/2;
else
x=3*x+1;
endif
s=[s, x];
endwhile
endfunction
```

Capítulo 11

El Periodo del Péndulo Simple

11.1. Introducción

Esta sección está dedicada a la presentación de una práctica en concreto con el objetivo de que el alumno disponga de un material preliminar.

Un problema de física elemental es el movimiento de un péndulo simple de longitud ℓ y masa m . Utilizando las leyes de Newton y suponiendo que no hay rozamiento, es fácil ver que la ecuación del movimiento viene dada por $m\ell\ddot{\theta} = -mg\sin\theta$, siendo θ el ángulo que forma el péndulo respecto de la vertical y $g = 9,806 \text{ ms}^{-2}$ un valor promediado de la aceleración de la gravedad en la superficie terrestre.

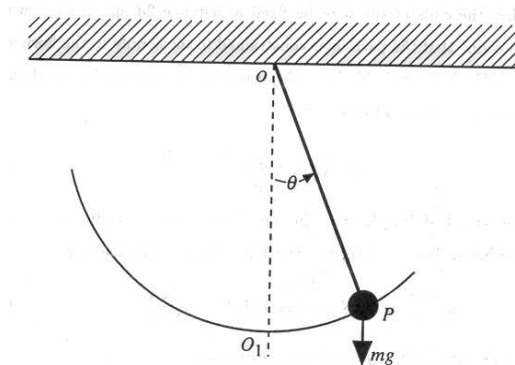


Figura 11.1: El péndulo simple.

Notemos que la ecuación del movimiento

$$\ddot{\theta} = -\frac{g}{\ell} \sin \theta ,$$

es independiente de la masa m del péndulo.

Puesto que la fuerza de la gravedad mg es conservativa y la tensión de la cuerda es perpendicular a la trayectoria de la partícula, se sabe que la energía total E del péndulo es constante. Teniendo en cuenta que $E = K + U$ es la suma de la energía cinética K y la potencial gravitatoria U se tiene fácilmente que $E = \frac{1}{2}m\ell^2\dot{\theta}^2 + mg\ell(1 - \cos\theta)$. Concluimos con la existencia de la integral primera (constante del movimiento)

$$\frac{1}{2}\dot{\theta}^2 + \frac{g}{\ell}(1 - \cos\theta) = C ,$$

siendo C una constante que se determina con las condiciones iniciales del movimiento del péndulo. Tomando $\dot{\theta} = 0$ para tiempo $t = 0$ en el cual el péndulo forma un ángulo $\theta = \theta_0$ se tiene, despejando de la anterior ecuación

$$\frac{d\theta}{dt} = \dot{\theta} = \sqrt{\frac{2g}{\ell}} \sqrt{\cos\theta - \cos\theta_0} .$$

Recordando que el periodo T es cuatro veces el tiempo que tarda el péndulo en ir de $\theta = 0$ hasta $\theta = \theta_0$, separando variables e integrando se tiene

$$T = 4\sqrt{\frac{\ell}{2g}} \int_0^{\theta_0} \frac{d\theta}{\sqrt{\cos\theta - \cos\theta_0}} .$$

Finalmente, teniendo en cuenta la identidad trigonométrica

$$\cos\theta - \cos\theta_0 = 2 \left(\sin^2 \frac{\theta_0}{2} - \sin^2 \frac{\theta}{2} \right) ,$$

y realizando el cambio de variable

$$\sin \xi = \frac{\sin(\theta/2)}{\sin(\theta_0/2)} ,$$

se obtiene que el periodo es

$$T = 4\sqrt{\frac{\ell}{g}} K(\sin(\theta_0/2)) ,$$

siendo

$$K(k) := \int_0^{\pi/2} \frac{d\xi}{\sqrt{1 - k^2 \sin^2 \xi}} , \quad (11.1)$$

la *integral elíptica completa de primera especie*.

11.2. Algoritmo de Integración Gaussiana con 5 Nodos

Si la función de peso es $w(x) = 1$, el intervalo de integración es $[-1, 1]$ y tomamos 5 nodos x_0, x_1, x_2, x_3, x_4 , las raíces del polinomio de Legendre de quinto grado

$$P_5(x) = \frac{1}{5!2^5} \frac{d^5}{dx^5} (x^2 - 1)^5 = \frac{1}{8} x(15 - 70x^2 + 63x^4) ,$$

son

$$\begin{aligned} -x_0 &= x_4 = \frac{1}{3} \sqrt{5 + 2\sqrt{10/7}} \approx 0,906179845938664 , \\ -x_1 &= x_3 = \frac{1}{3} \sqrt{5 - 2\sqrt{10/7}} \approx 0,538469310105683 , \\ x_2 &= 0 . \end{aligned}$$

Además, los coeficientes α_k de la fórmula de integración gaussiana

$$\int_{-1}^1 f(x) dx \approx \sum_{k=0}^4 \alpha_k f(x_k) ,$$

son

$$\alpha_k = \int_{-1}^1 \ell_k(x) dx = \int_{-1}^1 \prod_{\substack{j=0 \\ j \neq k}}^4 \frac{x - x_j}{x_k - x_j} dx .$$

Calculando estas integrales se obtiene

$$\begin{aligned} \alpha_0 &= \alpha_4 = 0,3 \frac{0,7 + 5\sqrt{0,7}}{2 + 5\sqrt{0,7}} \approx 0,236926885056189 , \\ \alpha_1 &= \alpha_3 = 0,3 \frac{-0,7 + 5\sqrt{0,7}}{-2 + 5\sqrt{0,7}} \approx 0,478628670499366 , \\ \alpha_2 &= \frac{128}{225} \approx 0,568888888888889 . \end{aligned}$$

A continuación presentamos un programa FORTRAN para aproximar los valores de la integral

$$I = \int_a^b f(\xi) d\xi ,$$

mediante el algoritmo gaussiano presentado anteriormente. El programa realiza el cambio lineal de variables $\xi(x) = \frac{1}{2}[(b-a)x + (a+b)]$ en la anterior integral, de manera que transforma el intervalo de integración $[a, b]$ en $[-1, 1]$. De esta forma se tiene

$$I = \frac{b-a}{2} \int_{-1}^1 f(\xi(x)) dx \approx \frac{b-a}{2} \sum_{k=0}^4 \alpha_k f([(b-a)x_k + (a+b)]/2) .$$

Las características del programa son las siguientes:

- Define en la sentencia `PARAMETER` la constante π y la aceleración de la gravedad g .
- Las constantes $X(I)$ y $ALPHA(I)$ son los nodos x_i y los coeficientes α_i de la fórmula de integración gaussiana. El algoritmo aprovecha las simetrías de estas constantes en la fórmula gaussiana con 5 puntos.
- Las constantes A y B son los límites de integración, a y b que en nuestro caso valen $a = 0$ y $b = \pi/2$.
- Las variables L y $THETA$ denotan la longitud ℓ y el ángulo inicial θ_0 respecto de la vertical del péndulo. Ambos son los únicos datos de entrada del algoritmo.
- La variable K denota la integral elíptica completa de primera especie (11.1) que es aproximada mediante la fórmula gaussiana con 5 puntos. El integrando de la integral elíptica es la función F definida al final del programa.
- La salida del algoritmo es el valor aproximado de la integral elíptica completa de primera especie K y el periodo del péndulo T .

El programa FORTRAN es el siguiente:

```

PROGRAM INTGAUSS5
REAL L, K
COMMON THETA0
DIMENSION X(0:1), ALPHA(0:2)

PARAMETER(PI=3.141592, G=9.806)

X(0)= -0.906179
X(1)= -0.538469
ALPHA(0)= 0.236926
ALPHA(1)= 0.478628
ALPHA(2)= 0.568888

A=0.
B=PI*0.5

WRITE(*,*)'Introduce la longitud del pendulo'
READ(*,*) L

WRITE(*,*)'Introduce EN GRADOS el angulo inicial'
READ(*,*) THETA0
THETA0=THETA0*PI/180.

K=ALPHA(2)*F((A+B)*0.5)

```

```

DO I=0,1
U=((B-A)*X(I)+A+B)*0.5
V=(-(B-A)*X(I)+A+B)*0.5
K=K+ALPHA(I)*(F(U)+F(V))
END DO
K=0.5*(B-A)*K

T=4.*SQRT(L/G)*K

WRITE(*,*)'La integral eliptica vale', K
WRITE(*,*)'El periodo del pendulo es T=', T

STOP
END

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

FUNCTION F(XI)
COMMON THETA0
F=1./SQRT(1.-(SIN(THETA0*0.5)*SIN(XI))**2)
RETURN
END

```

11.3. Resultados Numéricos

Si se realiza la aproximación del péndulo de pequeñas oscilaciones, es decir si aproximamos $\sin \theta \approx \theta$ en las ecuaciones del movimiento del péndulo se obtiene la ecuación diferencial lineal $\ddot{\theta} = -\frac{g}{\ell} \theta$ cuya solución es $\theta(t) = \theta_0 \cos[\sqrt{g/\ell} t]$. De aquí se deduce fácilmente que el periodo del péndulo de pequeñas oscilaciones es $T_{\text{aprox}} = 2\pi\sqrt{\ell/g}$.

Tomando la longitud del péndulo $\ell = 1$, nótese en la tabla adjunta al final de esta sección que la diferencia entre la aproximación obtenida por nuestro algoritmo para el periodo real del péndulo T y el periodo $T_{\text{aprox}} = 2,007089$ aumenta con el ángulo θ_0 . Dicho de otro modo, los resultados obtenidos muestran que el periodo de un péndulo simple es una función monótona creciente de su amplitud.

La salida del programa $K(k)$ para diferentes ángulos iniciales θ_0 es comparada con las tablas de integrales elípticas completas de primera especie $K(k)$ dadas en la página 610 del libro *HANDBOOK OF MATHEMATICAL FUNCTIONS*, M. Abramowitz y I.A. Stegun, Dover Publications, New York. Observamos que, a medida que aumenta el ángulo inicial θ_0 se van perdiendo en nuestros cálculos dígitos significativos hasta llegar a un resultado totalmente absurdo para $\theta_0 = 180^\circ$. Notemos que, para este ángulo inicial se tiene que la integral elíptica completa de primera especie $K(1)$ es una integral impropia divergente puesto

que

$$\begin{aligned}
 K(1) &= \int_0^{\pi/2} \frac{d\xi}{\sqrt{1 - \sin^2 \xi}} = \lim_{\epsilon \rightarrow 0^+} \int_0^{\frac{\pi}{2} - \epsilon} \frac{d\xi}{\cos \xi} \\
 &= \lim_{\epsilon \rightarrow 0^+} \left[\ln(\sec \xi + \tan \xi) \right]_0^{\frac{\pi}{2} - \epsilon} = \infty .
 \end{aligned}$$

$\theta_0 = 2 \arcsin k$	$K(k)$	$K(k)$ en ABRAMOWITZ	T para $\ell = 1$
20°	1.582839	1.582842	2.021859
40°	1.620022	1.620025	2.069355
60°	1.685747	1.685750	2.153309
80°	1.786768	1.786769	2.282350
100°	1.935589	1.935581	2.472448
120°	2.156442	2.156515	2.754558
140°	2.503612	2.504550	3.198020
160°	3.163038	3.153385	4.040345
180°	4.808182	∞	6.141790

Capítulo 12

Sistemas Lineales de Ecuaciones: Eliminación Gaussiana

12.1. Realización de la Práctica

En esta práctica, el alumno ha de realizar un programa con el objetivo de resolver sistemas de ecuaciones lineales. El método utilizado será el *algoritmo de eliminación Gaussiana* sin pivotación que mostramos en la Figura 12.1.

```
para k=1, n-1
  para i=k+1, n      (CALC. MULTIP.)
     $m_{ik} = a_{ik}/a_{kk}$ 
  finpara

  para i=k+1, n      (REDUC. COEF. SIST.)
    para j=k+1, n
       $a_{ij} = a_{ij} - m_{ik}a_{kj}$ 
    finpara
  finpara

  para i=k+1, n      (REDUC. TÉRM. INDEP.)
     $b_i = b_i - m_{ik}b_k$ 
  finpara
finpara
```

Figura 12.1: Algoritmo de *eliminación Gaussiana*.

Una vez hemos triangularizado la matriz de los coeficientes del sistema con el

algoritmo anterior, el segundo paso consiste en resolverlo por sustitución hacia atrás, ver el algoritmo de la Figura 12.2.

Se puede realizar un único programa que realice los dos pasos (eliminación Gaussiana y sustitución hacia atrás) o bien partirlo en subrutinas.

<pre> para i=n, 1, -1 $x_i = b_i/u_{ii}$ para j=i+1, n $x_i = x_i - u_{ij}x_j/u_{ii}$ finpara finpara </pre>
--

Figura 12.2: Algoritmo de *sustitución hacia atrás*.

- Queremos resolver el sistema lineal compatible y determinado $Ax = b$, de manera que los datos de entrada han de ser los coeficientes de la matriz $A = (a_{ij})$ y el vector de términos independientes $b = (b_1, b_2, \dots, b_n)$. Estos datos se pueden introducir manualmente por el teclado o bien ser leídos por un fichero.
- El programa nos debe devolver en la salida la solución del sistema lineal, es decir el vector $x = (x_1, x_2, \dots, x_n)$. Esta salida puede ser por pantalla o bien escrita en un fichero.
- Un segundo programa mejorado ha de incluir la estrategia de *pivotación parcial*.

12.1.1. Solución en FORTRAN

Un programa FORTRAN que implementa el algoritmo de eliminación Gaussiana es el siguiente:

```

PROGRAM GAUSS
IMPLICIT REAL*8(A-H,O-Z)
REAL*8 M
PARAMETER(NMAX=100)
DIMENSION A(NMAX,NMAX), M(NMAX,NMAX), B(NMAX), X(NMAX)

10 WRITE(*,*) ' Introduce el numero de ecuaciones '
   READ(*,*) N
   IF (N .GE. NMAX) GO TO 10

20 WRITE(*,*) ' Sistema por (1) TECLADO, (2) FICHERO '
   READ(*,*) L

```

```
IF ((L .NE. 1) .AND. (L .NE. 2)) GO TO 20
IF(L .EQ. 2) GO TO 40
```

C-----ENTRADAR EL SISTEMA POR TECLADO

```
WRITE(*,*)' Introduce los coeficientes del sistema:'
DO I=1,N
DO J=1,N
WRITE(*,*)'fila',I,'columna',J,'='
READ(*,*) A(I,J)
END DO
END DO
WRITE(*,*)' Introduce los terminos independientes:'
DO I=1,N
WRITE(*,*)'B(',I,')='
READ(*,*) B(I)
END DO
GO TO 50
```

C-----ENTRADA DEL SISTEMA POR FICHERO

```
40 OPEN(UNIT=1,FILE='SISTEMA.DAT',STATUS='UNKNOWN')
DO I=1,N
READ(1,*,ERR=100) (A(I,J),J=1,N), B(I)
END DO
```

C-----TRIANGULARIZACION

```
50 DO K=1,N-1
IF(A(K,K) .EQ. 0.) THEN
WRITE(*,*)'Hay un pivote nulo: MATRIZ SINGULAR'
STOP
END IF
```

C-----CALCULO DE LOS MULTIPLICADORES

```
DO I=K+1,N
M(I,K)=A(I,K)/A(K,K)
END DO
```

C-----ELIMINACION GAUSSIANA EN LA MATRIZ DEL SISTEMA

```
DO I=K+1,N
DO J=K+1,N
A(I,J)=A(I,J)-M(I,K)*A(K,J)
END DO
END DO
```

C-----ELIMINACION GAUSSIANA EN LOS TERMINOS INDEPENDIENTES

```
DO I=K+1,N
B(I)=B(I)-M(I,K)*B(K)
```

```

END DO
END DO

C-----RESOLUCION POR SUSTITUCION HACIA ATRAS
DO I=N,1,-1
  X(I)=B(I)/A(I,I)
  DO J=I+1,N
    X(I)=X(I)-A(I,J)*X(J)/A(I,I)
  END DO
END DO

C-----OPCIONES DE SALIDA
60 WRITE(*,*)' Solucion por (1) PANTALLA, (2) FICHERO'
  READ(*,*) L
  IF ((L .NE. 1) .AND. (L .NE. 2)) GO TO 60
  IF(L .EQ. 1) GO TO 70

C-----IMPRIMIR SOLUCION EN FICHERO
  OPEN(UNIT=2,FILE='SOLUCION.DAT',STATUS='UNKNOWN')
  DO I=1,N
    WRITE(2,*,ERR=200) X(I)
  END DO
  STOP

C-----IMPRIMIR SOLUCION EN PANTALLA
70 DO I=1,N
  WRITE(*,*) X(I)
END DO
STOP

C-----ADVERTENCIAS
100 WRITE(*,*)'Error de lectura en el fichero de datos'
  STOP
200 WRITE(*,*)'Error de lectura en el fichero de datos'
  STOP
END

```

12.1.2. Solución en MATHEMATICA

Un programa MATHEMATICA que implementa el algoritmo de eliminación Gaussiana y además calcula la descomposición LU de la matriz de coeficientes del sistema es el siguiente:

```

A = ReadList[[' c:/matriz.dat ', Number,RecordLists ->True];
b = ReadList[[' c:/vector.dat ', Number,RecordLists ->True][[1]];

```

```

n = Dimensions[A][[1]];

M = IdentityMatrix[n];

U = Table[0, {i, n}, {j, n}]; x = Table[0, {i, n}];

For[k = 1, k <= n - 1, k++, (* k-esimo paso de eliminación *)
  For[i = k + 1, i <= n, i++, (* multip. con pivotes no nulos *)
    If[A[[k, k]] != 0, M[[i, k]] = A[[i, k]]/A[[k, k]],
    Print["pivote nulo"]; Interrupt[]
  ];

  For[i = k + 1, i <= n, i++, (* Eliminación *)
    For[j = k + 1, j <= n, j++,
      A[[i, j]] = A[[i, j]] - M[[i, k]]*A[[k, j]];
    ];
    For[i = k + 1, i <= n, i++, b[[i]] = b[[i]] - M[[i, k]]*b[[k]]
  ];
];

For[i = n, i >= 1, i--, (* Sustitución hacia atrás *)
  x[[i]] = b[[i]]/A[[i, i]];
  For[j = i + 1, j <= n, j++,
    x[[i]] = x[[i]] - A[[i, j]]*x[[j]]/A[[i, i]]
  ];
];

L = M; (* Matrices L y U *)
For[i = 1, i <= n, i++,
  For[j = 1, j <= n, j++,
    If[i > j, U[[i, j]] = 0, U[[i, j]] = A[[i, j]]
  ]
];

Print["L=", MatrixForm[L]] (*salida*)
Print["U=", MatrixForm[U]]
Print["x=", MatrixForm[x]]
Print["FIN"]

```

Es posible comprobar si los cálculos efectuados son correctos mediante comandos propios de Mathematica para la resolución de sistemas lineales de ecuaciones.

```
AA = ReadList["c:/matriz.dat", Number, RecordLists ->True];
```

```
bb = ReadList["c:/vector.dat", Number, RecordLists ->True][[1]];

MatrixForm[AA]
MatrixForm[L]
MatrixForm[U]
MatrixForm[AA - L.U]

x
LinearSolve[AA, bb]
LinearSolve[U, b]
```

12.2. Aplicación al Cálculo de Intensidades en Redes Eléctricas

Finalizamos esta sección con una aplicación del programa de eliminación Gaussiana desarrollado anteriormente al cálculo de intensidades en redes eléctricas.

Es conocido que las *Leyes de Kirchoff* de los circuitos eléctricos establecen que

1. *Conservación de la carga*: el flujo neto de la corriente que pasa por las uniones de un circuito (*nudos*) es cero.
2. *Conservación de la energía*: la caída de voltaje neto alrededor de las partes cerradas de un circuito (*mallas*) es cero.

Cualquier problema de redes eléctricas se puede resolver de una forma sistemática por medio de las Leyes de Kirchoff. Recordemos que, para aplicar las Leyes de Kirchoff, es necesaria la *Ley de Ohm* que establece que la caída de potencial v_j en una resistencia R_j por la que circula una intensidad de corriente eléctrica i_j es $v_j = i_j R_j$. De esta forma, la primera y segunda ley de Kirchoff adoptan la forma siguiente

$$\sum_j i_j = 0, \quad \sum_j v_j + \sum_j i_j R_j = 0.$$

Supóngase que aplicamos un potencial $V = 5,5$ voltios entre los puntos A y G del circuito de la Figura 12.3. Aplicar las Leyes de Kirchoff para obtener el sistema lineal de ecuaciones que verifican las intensidades i_j con $j = 1, 2, 3, 4, 5$, y resolverlo con el programa de eliminación gaussiana realizado anteriormente.

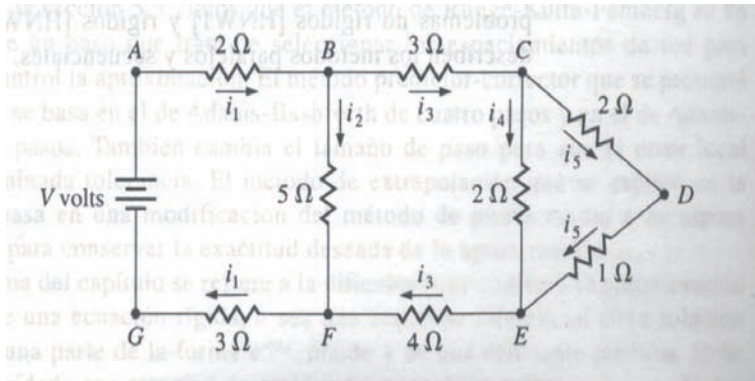


Figura 12.3: Las intensidades de la red eléctrica verifican un sistema lineal de ecuaciones.

Capítulo 13

La Ecuación del Calor Unidimensional

13.1. Introducción

Consideremos un material de coeficiente de *conductividad térmica* κ , *calor específico* c y *densidad* ρ tal que su anchura es despreciable respecto de su longitud L (por ejemplo un alambre). Es fácil demostrar que, si el material está aislado lateralmente y se verifica la ley de Fourier, entonces la ecuación que gobierna la evolución de la temperatura $T(x, t)$ en un punto de posición x para tiempo t es la llamada *ecuación del calor*

$$\frac{\partial T}{\partial t}(x, t) = \sigma^2 \frac{\partial^2 T}{\partial x^2}(x, t) , \quad (13.1)$$

donde $\sigma^2 = \frac{\kappa}{c\rho}$ es el llamado *coeficiente de difusividad*. Este es un ejemplo de ecuación en derivadas parciales parabólica. Para hallar una única solución $T(x, t)$ se necesita una condición inicial, la distribución inicial de temperatura en el material

$$T(x, 0) = f(x) , \quad 0 \leq x \leq L , \quad (13.2)$$

y una condición de contorno, temperatura en los extremos del material

$$T(0, t) = T_0 , \quad T(L, t) = T_L , \quad 0 \leq t < \infty . \quad (13.3)$$

Las constantes T_0 y T_L así como la función $f(x)$ son datos del problema.

13.2. Método de Crank-Nicholson

Vamos a dar un método numérico para aproximar la solución $T(x, t)$ mediante un esquema de diferencias divididas implícito ideado por John Crank y Phyllis Nicholson.

Tomemos el origen de coordenada espacial x en un extremo de la barra del material. Realizemos una discretización del espacio y del tiempo con longitudes de paso h y k constantes, de manera que $x_i = ih$ y $t_j = jk$ para $i = 0, 1, \dots, n$, $j = 0, 1, \dots, m$. De esta forma hemos formado una malla de puntos

$$(x_i, t_j) \in \{(x, t) : 0 \leq x \leq L, 0 \leq t \leq \tau\}.$$

Obviamente, se verifica $T(x_i, t_0) = f(x_i)$ y se pretende obtener las aproximaciones $T_i^j \approx T(x_i, t_j)$ de la solución exacta para $1 \leq j \leq m$.

Usando la aproximación de la derivada parcial respecto del tiempo que se obtiene a partir de la fórmula de diferencias centradas

$$\frac{\partial T}{\partial t}(x, t + k/2) = \frac{T(x, t + k) - T(x, t)}{k} + O(k^2), \quad (13.4)$$

y, en vez de utilizar $\partial^2 T / \partial x^2(x, t) = [T(x - h, t) - 2T(x, t) + T(x + h, t)]/h^2 + O(h^2)$, usamos como aproximación el valor medio de las aproximaciones en las cantidades $\partial^2 T / \partial x^2(x, t)$ y $\partial^2 T / \partial x^2(x, t + k)$, es decir,

$$\begin{aligned} \frac{\partial^2 T}{\partial x^2}(x, t + k/2) &= \frac{1}{2h^2} [T(x - h, t + k) - 2T(x, t + k) + T(x + h, t + k) \\ &\quad + T(x - h, t) - 2T(x, t) + T(x + h, t)] + O(h^2). \end{aligned} \quad (13.5)$$

Sustituyendo (13.4) y (13.5) en la ecuación del calor (13.1) y despreciando los términos $O(h^2)$ y $O(k^2)$ se tiene

$$\frac{T_i^{j+1} - T_i^j}{k} = \sigma^2 \frac{T_{i-1}^{j+1} - 2T_i^{j+1} + T_{i+1}^{j+1} + T_{i-1}^j - 2T_i^j + T_{i+1}^j}{2h^2}$$

Definiendo $r = \sigma^2 k / h^2$ y despejando los valores de T^{j+1} se obtiene la ecuación implícita

$$-rT_{i-1}^{j+1} + (2 + 2r)T_i^{j+1} - rT_{i+1}^{j+1} = (2 - 2r)T_i^j + r(T_{i-1}^j + T_{i+1}^j),$$

para $i = 1, 2, \dots, n - 1$. Es habitual tomar $r = 1$ en las anteriores ecuaciones, de manera que se simplifican notablemente dando lugar a

$$-T_{i-1}^{j+1} + 4T_i^{j+1} - T_{i+1}^{j+1} = T_{i-1}^j + T_{i+1}^j, \quad i = 1, 2, \dots, n - 1, \quad (13.6)$$

siendo la longitud de paso en el tiempo

$$k = \frac{h^2}{\sigma^2}. \quad (13.7)$$

Conocidos los T^j se calculan a partir del sistema (13.6) los T^{j+1} . Además, hay que tener en cuenta las condiciones de contorno $T_0^j = T_0^{j+1} = T_0$ y $T_n^j = T_n^{j+1} = T_L$.

Desarrollando en forma matricial las ecuaciones (13.6) se tiene el siguiente sistema lineal de ecuaciones donde la matriz del sistema es tridiagonal y diagonal dominante por filas

$$\begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & -1 & 4 & -1 & \\ & & \ddots & \ddots & \ddots \\ & & & -1 & 4 & -1 \\ & & & & -1 & 4 \end{pmatrix} \begin{pmatrix} T_1^{j+1} \\ T_2^{j+1} \\ T_3^{j+1} \\ \vdots \\ T_{n-2}^{j+1} \\ T_{n-1}^{j+1} \end{pmatrix} = \begin{pmatrix} 2T_0 + T_2^j \\ T_1^j + T_3^j \\ T_2^j + T_4^j \\ \vdots \\ T_{n-3}^j + T_{n-1}^j \\ T_{n-2}^j + 2T_L \end{pmatrix}. \quad (13.8)$$

Se han de resolver m sistemas lineales (13.8) correspondientes a los distintos valores $j = 0, 1, \dots, m-1$. Obsérvese que, cada vez que se resuelve un sistema lineal (13.8) se está aproximando la solución de la temperatura en la barra en una unidad de tiempo, es decir, conocidos los términos independientes que no son más que los T_i^j , se obtienen los T_i^{j+1} .

13.3. Separación de Variables: Series de Fourier

Es posible calcular la solución exacta $T(x, t)$ de la ecuación (13.1) sujeta a las condiciones (13.2) y (13.3) mediante el uso de series de Fourier. Para conseguirlo se buscan soluciones de la ecuación del calor (13.1) con las variables separadas, es decir, de la forma $T(x, t) = \xi(x)\eta(t)$. La idea general del método de separación de variables consiste en que es posible obtener un número infinito de este tipo de soluciones que verifican al mismo tiempo la condición inicial (13.2) y las condiciones de contorno (13.3). Este número infinito de soluciones serán denotadas por $T_n(x, t) = \xi_n(x)\eta_n(t)$ con $n = 1, 2, \dots, \infty$ y son llamadas *soluciones fundamentales*. Puesto que la ecuación del calor es una ecuación en derivadas parciales lineal se verifica el principio de superposición, de manera que la serie

$$T(x, t) = \sum_{n=1}^{\infty} \alpha_n T_n(x, t) = \sum_{n=1}^{\infty} \alpha_n \xi_n(x) \eta_n(t), \quad (13.9)$$

con coeficientes $\alpha_n \in \mathbb{N}$, es la solución de la ecuación del calor en el supuesto de que la serie sea convergente.

13.3.1. Búsqueda de las soluciones fundamentales $T_n(x, t)$ con condiciones de contorno homogéneas

Supongamos que las condiciones de contorno (13.3) son homogéneas, es decir, $T_0 = T_L = 0$. Entonces, las condiciones de contorno (13.3) reducen a

$$T(0, t) = 0, \quad T(L, t) = 0, \quad 0 \leq t < \infty. \quad (13.10)$$

Impongamos que $T(x, t) = \xi(x)\eta(t)$ sea una solución de la ecuación del calor (13.1). Entonces se obtiene que

$$\xi(x)\dot{\eta}(t) = \sigma^2 \xi''(x)\eta(t) .$$

Separando en esta ecuación los términos espaciales y temporales llegamos a

$$\frac{\dot{\eta}(t)}{\sigma^2 \eta(t)} = \frac{\xi''(x)}{\xi(x)} .$$

Puesto que las variables x y t son independientes, cada miembro de la anterior ecuación sólo puede ser una constante que, por conveniencia, llamamos $-\lambda^2$. Se llega de esta forma a que la ecuación del calor (13.1) en derivadas parciales reduce a un sistema de dos ecuaciones diferenciales ordinarias lineales y no acopladas

$$\begin{aligned} 0 &= \dot{\eta}(t) + \sigma^2 \lambda^2 \eta(t) , \\ 0 &= \xi''(x) + \lambda^2 \xi(x) , \end{aligned}$$

ambas fáciles de resolver. En concreto $\xi(x)$ verifica la ecuación de un oscilador armónico y $\eta(t)$ verifica una ecuación lineal de primer orden de variables separables. En definitiva se tiene

$$\begin{aligned} \xi(x) &= A \sin(\lambda x) + B \cos(\lambda x) , \\ \eta(t) &= C \exp(-\lambda^2 \sigma^2 t) , \end{aligned}$$

siendo $A, B, C \in \mathbb{R}$ constantes arbitrarias. De momento se tiene que las funciones del tipo

$$\begin{aligned} T(x, t) &= \xi(x)\eta(t) = [A \sin(\lambda x) + B \cos(\lambda x)]C \exp(-\lambda^2 \sigma^2 t) \\ &= [A_1 \sin(\lambda x) + A_2 \cos(\lambda x)] \exp(-\lambda^2 \sigma^2 t) , \end{aligned}$$

con $A_1, A_2, \lambda \in \mathbb{R}$ constantes arbitrarias verifican la ecuación del calor (13.1). Calculemos dichas constantes imponiendo ahora las condiciones de contorno homogéneas (13.10). De $T(0, t) = 0$ para todo $t > 0$ se deduce que $A_2 = 0$. Finalmente, de $T(L, t) = 0$ para todo $t > 0$ se obtiene $A_1 \sin(\lambda L) = 0$. Como no queremos hallar la solución trivial $T(x, t) \equiv 0$ no anularemos de la anterior ecuación A_1 sino que tomaremos $\sin(\lambda L) = 0$ o equivalentemente $\lambda := \lambda_n$ con

$$\lambda_n = \frac{n\pi}{L} , \quad n \in \mathbb{N} . \quad (13.11)$$

Hemos obtenido pues las soluciones fundamentales

$$T_n(x, t) = \xi_n(x)\eta_n(t) = \sin(\lambda_n x) \exp(-\lambda_n^2 \sigma^2 t) . \quad (13.12)$$

Por lo tanto, según (13.15), la solución exacta $T(x, t)$ viene dada por la serie

$$T(x, t) = \sum_{n=1}^{\infty} \alpha_n \sin\left(\frac{n\pi}{L}x\right) \exp\left[-\left(\frac{n\pi}{L}\right)^2 \sigma^2 t\right] . \quad (13.13)$$

Sólo falta hallar los coeficientes α_n , que vendrán determinados por la distribución inicial de temperaturas. En concreto, puesto que

$$T(x, 0) = f(x) = \sum_{n=1}^{\infty} \alpha_n \sin\left(\frac{n\pi}{L}x\right),$$

es una serie de Fourier, los coeficientes α_n vienen determinados por

$$\alpha_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{n\pi}{L}x\right) dx. \quad (13.14)$$

13.3.2. Realización de la práctica

Realizar un programa FORTRAN que implemente el esquema de diferencias divididas implícitas de Crank-Nicholson.

- (i) Los datos de entrada han de ser $L, \sigma, T_0, T_L, f(x), n, \tau$.
- (ii) Utilizar el procedimiento FUNCTION después del END del programa principal, para definir la distribución de temperaturas inicial $f(x)$. Por ejemplo, si $f(x) = \sin(x)$, se escribirá

```
FUNCTION F(X)
F=SIN(X)
RETURN
END
```

- (iii) Puesto que todos los sistemas lineales (13.8) tienen la misma matriz de coeficientes y lo único que varía son los términos independientes, un buen método para su resolución consiste en hallar la descomposición LU de la matriz de coeficientes del sistema (13.8). Además para hallar tal descomposición LU , la eliminación Gaussiana será numéricamente estable. Aprovechar además la estructura tridiagonal de la matriz de coeficientes de (13.8).
- (iv) Definir en una sentencia PARAMETER los valores máximos de n y m asignándoles, por ejemplo, los valores $NMAX = 100, MMAX = 100$. Dimensionar posteriormente las matrices $L, U \in \mathcal{M}_{n-1}(\mathbb{R})$. Los elementos $T(x_i, t_j) \approx T_i^j$ almacenarlos también en una matriz que se habrá dimensionado adecuadamente y que en el programa serán utilizados de la forma $T(I, J)$.
- (v) NOTA: Llamar en el programa LL a la longitud de la barra para no tener confusiones con la matriz triangular inferior L .
- (vi) La salida del programa será un fichero, por ejemplo *temperaturas.dat*, que contenga las parejas de valores (x_i, T_i^j) ordenados de la forma

$$x_0, T_0$$

$$\begin{array}{c}
x_1, T_1^0 \\
\vdots \\
x_{n-1}, T_{n-1}^0 \\
x_n, T_L \\
x_0, T_0 \\
x_1, T_1^1 \\
\vdots \\
x_{n-1}, T_{n-1}^1 \\
x_n, T_L \\
\vdots \\
x_0, T_0 \\
x_1, T_1^m \\
\vdots \\
x_{n-1}, T_{n-1}^m \\
x_n, T_L
\end{array}$$

Cuando el anterior programa funcione correctamente, aplicarlo al siguiente problema:

Problema: Considerar una barra de un cierto material cuya anchura es despreciable respecto de su longitud $L = 1$, que verifique la ley de Fourier y tal que su coeficiente de conductividad vale $\sigma = 0,5$. Los extremos de dicha barra se ponen en contacto con dos focos térmicos a temperatura cero. La barra está aislada térmicamente por los laterales y además se sabe que la distribución inicial de temperaturas viene dada por la función continua

$$T(x, 0) = f(x) := \begin{cases} x & \text{si } 0 \leq x \leq 1/2, \\ 1 - x & \text{si } 1/2 < x \leq 1. \end{cases}$$

Aproximar la evolución de la temperatura $T(x, t)$ en la barra en el intervalo de tiempo $t \in [0, \tau]$, es decir, hallar $T(x_i, t_j)$ para $x_i = ih$, $t_j = jk$ con $i = 0, 1, \dots, n$ y $j = 0, 1, \dots, m$, siendo $n = 10$ y $\tau = 2$.

13.3.3. Representación gráfica de los resultados

Se pretende realizar secciones a tiempo constante de la superficie $T(x, t)$, con el objetivo de hallar las curvas $T(x, t_j)$ con $x \in [0, L]$ y t_j fijado.

Para representar gráficamente los resultados obtenidos en el problema de la sección anterior, es decir, las aproximaciones de $T(x_i, t_j)$, se procede con el programa MATHEMATICA de la forma siguiente.

1. Definir los parámetros necesarios:

$n = 10$; $L = 1$; $\tau = 2$; $\sigma = 0,5$; $h = L/n$; $k = h^2/\sigma^2$; $m = \text{IntegerPart}[\tau/k]$;

2. Cargar en una *lista* que llamaremos *temp* el fichero *temperaturas.dat*:

`temp = ReadList["temperaturas.dat", Number, RecordLists \rightarrow True]`

3. Extraer de la lista *temp* las parejas de puntos (x_i, T_i^j) para cada valor de j , es decir, las posiciones y temperaturas de la barra par tiempo t_j con $j = 0, 1, \dots, m$ e incluirlos en sublistas llamadas $t[j]$:

```
For[j = 0, j <= m - 1, j++, lista = { };
For[i = (n + 1)*j + 1, i <= (n + 1)*(j + 1), i++,
AppendTo[lista, temp[[i]]];
t[j] = lista
]
```

4. Dibujar las gráficas de la temperatura de la barra para los tiempos t_j con $j = 0, 1, \dots, m$. En concreto, a cada una de estas gráficas las llamaremos *grafica[j]*:

```
For[j = 0, j <= m, j++,
grafica[j] = ListPlot[t[j], PlotJoined  $\rightarrow$  True, AxesLabel  $\rightarrow$  {"x", "T"}];
]
```

5. Dibujar superpuestas las gráficas de la temperatura de la barra para los tiempos t_j con $j = 0, 5, 10, 15$, es decir, superponer las graficas *grafica[0]*, *grafica[5]*, *grafica[10]*, *grafica[15]*:

`Show[grafica[0], grafica[5], grafica[10], grafica[15]]`

13.3.4. Comparación con el método de Fourier

Una forma de aproximar la solución exacta $T(x, t)$ dada por la serie (13.13) es tomar una suma parcial

$$T(x, t) \approx \sum_{n=1}^s \alpha_n \sin\left(\frac{n\pi}{L}x\right) \exp\left[-\left(\frac{n\pi}{L}\right)^2 \sigma^2 t\right], \quad (13.15)$$

con $s \in \mathbb{N}$ suficientemente grande y los coeficientes α_n dados por (13.14).

Tomar $s = 10$ y visualizar gráficamente los resultados con MATHEMATICA de la forma siguiente.

1. Definir los parámetros necesarios:

$$s = 10; L = 1; \text{sigma} = 0,5; h = L/10; k = h^2/\text{sigma}^2;$$

2. Definir las dos funciones que introducen la distribución inicial de temperaturas

$$\begin{aligned} f1[x_] &= x \\ f2[x_] &= 1-x \end{aligned}$$

3. Definir los coeficientes λ_n y α_n a partir de (13.11) y (13.14):

$$\text{lambda}[n_] = n * \text{Pi} / L$$

$$\text{alpha}[n_] := \text{alpha}[n] = 1/L * (\text{NIntegrate}[f1[x] * \text{Sin}[\text{lambda}[n] * x], \{x, 0, 1/2\}] + \text{NIntegrate}[f2[x] * \text{Sin}[\text{lambda}[n] * x], \{x, 1/2, 1\}])$$

4. Definir las soluciones fundamentales $T_n(x, t)$ dadas por (13.12):

$$T[x_, t_, n_] := T[x, t, n] = \text{alpha}[n] * \text{Sin}[\text{lambda}[n] * x] * \text{Exp}[-\text{sigma}^2 * \text{lambda}[n]^2 * t]$$

5. Obtener la expresión de la temperatura aproximada a partir de la suma parcial (13.15):

$$\text{Taprox}[x_, t_] = \text{Sum}[T[x, t, n], \{n, 1, s\}]$$

6. Dibujar las gráficas de la temperatura aproximada de la barra para los tiempos t_j con $j = 0, 5, 10, 15$ y almacenar dichas gráficas en $\text{tempaprox}[j]$:

$$\begin{aligned} &\text{For}[j = 0, j < = 15, j += 5, \\ &t[j] = j * k; \\ &\text{tempaprox}[j] = \text{Plot}[\text{Taprox}[x, t[j]], \{x, 0, L\}, \text{AxesLabel} \rightarrow \{ "x", "T" \}] \\ &] \end{aligned}$$

7. Dibujar superpuestas las anteriores gráficas:

$$\text{Show}[\text{tempaprox}[0], \text{tempaprox}[5], \text{tempaprox}[10], \text{tempaprox}[15]]$$

Capítulo 14

Interpolación de Newton

14.1. Introducción

El problema de la interpolación tiene una relevancia especial en muchas aplicaciones a la ingeniería. Los métodos de interpolación tratan de encontrar una función (polinómica, trigonométrica, exponencial,...) que satisfaga ciertas condiciones en un conjunto finito de puntos. Estas condiciones pueden darse sobre los valores de la función, y también sobre sus derivadas. En diferentes disciplinas científicas y técnicas se plantea muy a menudo este problema. De hecho, son muchas y muy distintas las situaciones en las que aparecen series de datos o resultados de mediciones experimentales de los cuales sólo se conoce cómo se comportan en una cierta cantidad finita de puntos y para los cuales se necesita encontrar una “ley general”. Dicha ley no es mas que una función que tome unos valores predeterminados. Éste es, precisamente, el cometido de la interpolación. La función interpoladora servirá para sustituir a la función desconocida, tanto para evaluarla en puntos en los que no se conoce su valor (interpolación en sentido estricto), como para conocer su tasa de variación (diferenciación numérica) o su distribución acumulativa (integración numérica). Además, la interpolación sirve para fundamentar una amplia gama de métodos para el tratamiento numérico de ecuaciones diferenciales.

La función interpoladora debe ser, por tanto, fácil de evaluar, derivar e integrar.

Nos limitaremos a la interpolación polinómica que consiste en lo siguiente:

“Dada una función $f : [a, b] \rightarrow \mathbb{R}$ y un conjunto de abscisas $\Sigma = \{x_0, x_1, \dots, x_n\} \subset [a, b]$ con $x_i \neq x_j$ si $i \neq j$, encontrar un polinomio P_n verificando $P_n(x_i) = f(x_i)$ para $i = 0, 1, \dots, n$ ”.

En esta práctica el alumno ha de realizar un programa que implemente el método interpolatorio de las *diferencias divididas* de Newton. En una segunda parte de la práctica, aplicaremos el programa realizado para verificar de forma

numérica el *fenómeno Runge*, es decir comprobar que en general

$$\lim_{n \rightarrow \infty} P_n(x) \neq f(x) .$$

14.2. Diferencias Divididas de Newton

Sabemos que el polinomio que interpola a las parejas (x_i, f_i) con $i = 0, 1, \dots, n$ puede escribirse de la forma

$$P_n(x) = \sum_{i=0}^n c_i \varphi_i(x) ,$$

donde

$$\varphi_0(x) := 1 , \quad \varphi_i(x) = \prod_{k=0}^{i-1} (x - x_k) \quad \text{para } i = 1, 2, \dots, n.$$

Los coeficientes c_i se conocen con el nombre de *diferencias divididas* y su notación histórica es

$$c_i := f[x_0, x_1, \dots, x_i] .$$

Se puede demostrar que las *diferencias divididas* vienen definidas de la forma siguiente

$$\begin{aligned} f[x_i] &= f_i , \\ f[x_i, x_{i+1}, \dots, x_{i+j}] &= \frac{f[x_{i+1}, x_{i+2}, \dots, x_{i+j}] - f[x_i, x_{i+1}, \dots, x_{i+j-1}]}{x_{i+j} - x_i} . \end{aligned}$$

Un ejemplo de como esquematizar el proceso descrito para el cálculo de las diferencias divididas en el caso $n = 3$ es el siguiente

x_0	$f[x_0]$			
		$f[x_0, x_1]$		
x_1	$f[x_1]$		$f[x_0, x_1, x_2]$	
		$f[x_1, x_2]$		$f[x_0, x_1, x_2, x_3]$
x_2	$f[x_2]$		$f[x_1, x_2, x_3]$	
		$f[x_2, x_3]$		
x_3	$f[x_3]$			

A continuación, ver Figura 14.1, damos un algoritmo para el cálculo de las diferencias divididas.

Este algoritmo permite el cálculo de los coeficientes c_i que no son más que $f[x_0, x_1, \dots, x_i]$. Una forma de ahorrar memoria en la computadora consiste en dimensionar un vector $c = (c_0, c_1, \dots, c_n)$. Al introducir los datos de entrada f_0, f_1, \dots, f_n , se almacenan en el vector c , de manera que c_0 ya es el primer coeficiente del polinomio interpolador. Después se computa la primera columna del esquema de diferencias divididas y colocamos estos valores en c_1, c_2, \dots, c_n , de manera que c_0 conserva su valor y el nuevo c_1 es el segundo coeficiente del polinomio interpolador y así sucesivamente.

```

para i=0,n
     $c_i = f_i$ 
finpara

para j=1,n
    para i=n,j
         $c_i = (c_i - c_{i-1}) / (x_i - x_{i-j})$ 
    finpara
finpara

```

Figura 14.1: Algoritmo de diferencias divididas de Newton.

14.3. Fenómeno Runge

Para $n = 5, 10$ y 15 , calcular utilizando el programa de la sección anterior con abscisas equiespaciadas, el polinomio de interpolación de Newton $P_n(x)$ correspondiente a la función racional

$$f(x) = \frac{1}{1+x^2} \quad \text{definida en el intervalo } [-5, 5].$$

Finalmente, computar para cada valor de n el error cometido $|f(x_j) - P_n(x_j)|$ en 31 puntos distribuidos a iguales distancias en el intervalo $[-5, 5]$, es decir en $x_j = -5 + j/3$ para $j = 0, 1, \dots, 30$. El objetivo de este cálculo es evidenciar de forma numérica la divergencia de P_n respecto de f que es lo que se conoce como fenómeno Runge.

Observemos que para realizar esta parte de la práctica es necesario evaluar el polinomio interpolador en 31 puntos, es decir $P_n(x_j)$ para $j = 0, 1, \dots, 30$.

Una vez se conocen las diferencias divididas c_i para $i = 0, 1, \dots, n$, ver el anterior apartado, una forma de evaluar el polinomio interpolador P_n en un valor x viene dada en el pseudocódigo de la Figura 14.2. Dicho código utiliza la forma de Newton¹ del polinomio P_n , es decir

$$\begin{aligned}
 P_n(x) &= \sum_{i=0}^n c_i \varphi_i(x) \\
 &= c_0 + c_1(x - x_0) + \dots + c_n(x - x_0) \cdots (x - x_{n-1}).
 \end{aligned}$$

Como resultado final de esta práctica se tiene que apreciar los *efectos de borde*: el polinomio interpolador tiende a oscilar en los extremos del intervalo de interpolación $[-5, 5]$.

¹En esta práctica tomamos esta forma del polinomio aunque recordemos que, la forma más conveniente desde el punto de vista del número de operaciones necesarias para evaluar un polinomio es la de *Horner*. Por ejemplo el polinomio de cuarto grado $P_4(x) = \sum_{i=0}^4 a_i x^i$ adopta la forma de Horner $P_4(x) = a_0 + x(a_1 + x(a_2 + x(a_3 + aa_4)))$.

```

inicializar  $x$ 
valor= $c_0$ 
 $p = 1$ 
para  $i=0, n-1$ 
     $p = p(x - x_i)$ 
    valor=valor+ $c_i p$ 
finpara
salida  $x, valor$ 

```

Figura 14.2: Dando un valor x y conocidas las diferencias divididas c_i , el algoritmo devuelve la pareja $(x, P_n(x))$.

14.4. Minimización del Error de Interpolación

Se sabe que, si $f \in C^{n+1}[a, b]$ y P_n es el polinomio que interpola a f en $n+1$ puntos distintos $\{x_0, x_1, \dots, x_n\} \subset [a, b]$, para cada $x \in [a, b]$ existe $\xi_x \in [a, b]$ tal que el error Δ_x cometido en la interpolación en el punto x viene dado por

$$\Delta_x := |f(x) - P_n(x)| = \left| \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \Pi_n(x) \right| ,$$

siendo

$$\Pi_n(x) = \prod_{i=0}^n (x - x_i) .$$

Fijadas la función f y el grado n del polinomio interpolador P_n , se pretende encontrar los puntos de interpolación $\{x_0, x_1, \dots, x_n\}$ que hagan mínimo el error Δ_x con la norma del infinito o del máximo. En concreto se ha de minimizar $\|\Pi_n\|_\infty$ en el intervalo $[a, b]$, es decir

$$\|\Pi_n\|_\infty := \max_{a \leq x \leq b} |\Pi_n(x)| .$$

P.L. Chebychev fue el primero en hallar la solución al problema planteado.

Teorema 14.1 (Chebychev) *Dada una función $f \in C^{n+1}[-1, 1]$, el menor error en la norma del máximo en la interpolación de f se comete cuando se toman las siguientes abscisas de interpolación $\{x_0, x_1, \dots, x_n\} \subset [-1, 1]$*

$$x_j = \cos \frac{(2j+1)\pi}{2(n+1)} , \quad j = 0, 1, \dots, n .$$

NOTAS:

- Los puntos $\{x_0, x_1, \dots, x_n\}$ dados por el Teorema 14.1 no son equiespaciados, sino que se agrupan cerca de los extremos del intervalo.

- Aunque el Teorema 14.1 hace únicamente referencia a una interpolación en el intervalo $[-1, 1]$ puede extenderse fácilmente a un intervalo arbitrario $[a, b]$. Para ello basta con realizar el cambio lineal de variables $z = (2x - a - b)/(b - a)$ que lleva el intervalo $[a, b]$ al $[-1, 1]$. De esta forma, la nueva función a interpolar $F(z) : [-1, 1] \rightarrow \mathbb{R}$ está definida como

$$F(z) = f(x) = f\left(\frac{a+b}{2} + \frac{b-a}{2}z\right).$$

Se puede concluir que, las abscisas de interpolación

$$x_j = \frac{a+b}{2} + \frac{b-a}{2} \cos \frac{(2j+1)\pi}{2(n+1)}, \quad j = 0, 1, \dots, n, \quad (14.1)$$

minimizan el error de la interpolación de la función f en el intervalo $[a, b]$.

La parte final de la práctica consiste en repetir los cálculos de la sección anterior pero ahora con las abscisas de interpolación dadas por (14.1) con $[a, b] = [-5, 5]$ y comprobar como el error $|f(x) - P_n(x)|$ ha disminuido sensiblemente.

Capítulo 15

Interpolación por Splines Cúbicos Naturales

15.1. Introducción

El fenómeno Runge mostrado en la práctica anterior nos está indicando que la interpolación polinomial sobre un número grande de abscisas no es siempre adecuada debido a la tendencia a oscilar que tienen los polinomios. Una solución adecuada para este problema es la interpolación mediante funciones *spline*¹.

Matemáticamente, una función spline en un intervalo $[x_0, x_n]$ está formada por polinomios definidos en subintervalos de $[x_0, x_n]$ obedeciendo a ciertas condiciones de regularidad. Más concretamente, consideremos un conjunto de $n + 1$ pares (x_i, f_i) para $i = 0, 1, \dots, n$ sobre los que queremos construir un spline cúbico S que los interpole. Sea S_i el polinomio cúbico que representa a S en el intervalo $x \in [x_i, x_{i+1}]$, es decir

$$S(x) = \begin{cases} S_0(x) & x \in [x_0, x_1] \\ S_1(x) & x \in [x_1, x_2] \\ \vdots & \vdots \\ S_{n-1}(x) & x \in [x_{n-1}, x_n] \end{cases} \quad (15.1)$$

tal que se verifique la continuidad de S , S' y S'' .

Se puede demostrar que los splines cúbicos naturales se calculan de la forma

$$S_i(x) = f_i + (x - x_i) \left[C_i + (x - x_i) [B_i + (x - x_i) A_i] \right], \quad (15.2)$$

¹La palabra inglesa “spline” denota un instrumento flexible usado en dibujo técnico que sirve para trazar curvas suaves. Precisamente, la propiedad de adaptarse bien a formas dadas que tienen las funciones spline es lo que hace que se les dé tal nombre

donde hemos definido

$$A_i := \frac{1}{6h_i}(z_{i+1} - z_i), \quad B_i := \frac{z_i}{2}, \quad C_i := -\frac{h_i}{6}z_{i+1} - \frac{h_i}{3}z_i + \frac{1}{h_i}(f_{i+1} - f_i), \quad (15.3)$$

siendo

$$h_i := x_{i+1} - x_i, \quad z_i := S''(x_i). \quad (15.4)$$

Observar que lo único que nos queda por hacer para finalizar el problema de la interpolación de splines es calcular los valores z_0, z_1, \dots, z_n .

Tomando la elección $z_0 = z_n = 0$ (spline cúbico natural) se puede demostrar que las z_i verifican el sistema lineal, simétrico, tridiagonal y diagonal dominante por filas siguiente

$$\begin{pmatrix} \tilde{h}_1 & h_1 & & & \\ h_1 & \tilde{h}_2 & h_2 & & \\ & h_2 & \tilde{h}_3 & h_3 & \\ & & \ddots & \ddots & \ddots \\ & & & h_{n-3} & \tilde{h}_{n-2} & h_{n-2} \\ & & & & h_{n-2} & \tilde{h}_{n-1} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_{n-2} \\ z_{n-1} \end{pmatrix} = \begin{pmatrix} \tilde{f}_1 \\ \tilde{f}_2 \\ \tilde{f}_3 \\ \vdots \\ \tilde{f}_{n-2} \\ \tilde{f}_{n-1} \end{pmatrix} \quad (15.5)$$

donde hemos definido

$$\tilde{h}_i := 2(h_i + h_{i-1}), \quad \tilde{f}_i := b_i - b_{i-1}, \quad \text{siendo } b_i = \frac{6}{h_i}(f_{i+1} - f_i). \quad (15.6)$$

Para resolver numericamente el sistema lineal (15.5) se puede pensar en la eliminación Gaussiana sin pivotación. Las entradas del algoritmo son el número de nudos n y las parejas (x_i, f_i) desde $i = 0, 1, \dots, n$. La salida serán los valores de z_i para $i = 0, 1, \dots, n$. Observar que el algoritmo utiliza divisiones por \tilde{h}_i , pero no existe ningún problema ya que es fácil comprobar que $\tilde{h}_i > 0$.

Una vez se han determinado todos los valores de z_i se procedería al cálculo de la función spline $S(x)$ definida a partir de las ecuaciones (15.1), (15.4), (15.2) y (15.3). Una forma de calcular numéricamente el subintervalo $[x_i, x_{i+1}]$ al que pertenece x es mediante la evaluación ordenada de las diferencias

$$x - x_{n-1}, x - x_{n-2}, \dots, x - x_1,$$

de manera que si denotamos por $x - x_i$ a la primera de estas diferencias que es positiva, entonces $x \in [x_i, x_{i+1})$.

Un código para resolver el sistema lineal (15.5) puede ser el dado en la Figura 15.1.


```

entrada  $n, x_i, f_i$ 

para  $i=0, n-1$ 
     $h_i = x_{i+1} - x_i$ 
     $b_i = 6(f_{i+1} - f_i)/h_i$ 
finpara

 $th_1 = 2(h_0 + h_1)$ 
 $tf_1 = b_1 - b_0$ 
para  $i=2, n-1$ 
     $th_i = 2(h_i + h_{i-1}) - h_{i-1}^2/th_{i-1}$ 
     $tf_i = b_i - b_{i-1} - h_{i-1}tf_{i-1}/th_{i-1}$ 
finpara

 $z_n = 0$ 
para  $i=n-1, 1, -1$ 
     $z_i = (tf_i - h_i z_{i+1})/th_i$ 
finpara
 $z_0 = 0$ 

salida  $z_i$ 

```

Figura 15.1: Algoritmo para generar las segundas derivadas de splines cúbicos naturales calculadas en los nodos.

15.2. Realización de la Práctica

- Hacer un pequeño programa FORTRAN para tabular, con nudos equiespaciados, la función \sqrt{x} en el intervalo $[0, 1]$ utilizando $n = 10$ y escribiendo las parejas $(x_i, \sqrt{x_i})$ en un fichero.
- Realizar un programa FORTRAN que lea de un fichero los datos (x_i, f_i) para $i = 0, 1, \dots, n$ y que implemente un algoritmo de splines cúbicos naturales.
- Utilizar el programa anterior para aproximar en nudos equiespaciados la función \sqrt{x} en el intervalo $[0, 1]$ utilizando $n = 100$ y escribiendo las parejas $(x_i, S(x_i))$ en otro fichero.
- Mostrar (con un software adecuado) las gráficas superpuestas de la función \sqrt{x} y de la función spline $S(x)$ obtenida en el apartado anterior. Por ejemplo, con el software MATHEMATICA:

(i) Cargar en una *lista* el fichero guardado en el apartado anterior

datos=ReadList["fichero.dat", Number, RecordLists→True]

- (ii) Dibujar una gráfica de la función Spline utilizando los puntos de fichero.dat

```
spline=ListPlot[datos, PlotJoined→True]
```

- (iii) Dibujar una gráfica de la función \sqrt{x} en el intervalo $[0, 1]$

```
raiz=Plot[ Sqrt[x], { x,0,1 } ]
```

- (iv) Mostrar las dos gráficas superpuestas

```
Show[ spline, raiz ]
```

15.3. Solución en FORTRAN

Un posible código FORTRAN que implemente el cálculo de los splines cúbicos naturales es el siguiente.

```
PROGRAM SPLINE
IMPLICIT REAL*8(A-H,O-Z)
PARAMETER(NMAX=100)
DIMENSION X(0:NMAX), F(0:NMAX), Z(0:NMAX)
DIMENSION FT(0:NMAX), H(0:NMAX), HT(0:NMAX), B(0:NMAX)

10 WRITE(*,*)' Introduce numero de nodos N'
   READ(*,*) N
   IF (N .GE. NMAX) GO TO 10

20 WRITE(*,*)' Introducir puntos por (1) TECLADO, (2) FICHERO'
   READ(*,*) M
   IF ((M .NE. 1) .AND. (M .NE. 2)) GO TO 20
   IF(M .EQ. 2) GO TO 40

C-----ENTRADA DE PUNTOS POR TECLADO
   WRITE(*,*)' Introduce los puntos:'
   DO I=0,N
      READ(*,*) X(I),F(I)
   END DO
   GO TO 50

C-----ENTRADA DE PUNTOS POR FICHERO
40 OPEN(UNIT=1,FILE='INTERP.DAT',STATUS='UNKNOWN')
   DO I=0,N
      READ(1,*,ERR=100) X(I), F(I)
   END DO

C-----INICIALIZACION DE PARAMETROS
```

```

50 DO I=0,N-1
    H(I)=X(I+1)-X(I)
    B(I)=6*(F(I+1)-F(I))/H(I)
    END DO

    DO I=1,N-1
        HT(I)=2*(H(I)+H(I-1))
        FT(I)=B(I)-B(I-1)
    END DO

C-----INICIALIZACION DE PARAMETROS
    DO I=2,N-1
        HT(I)=HT(I)-H(I-1)**2/HT(I-1)
        FT(I)=FT(I)-H(I-1)*FT(I-1)/HT(I-1)
    END DO

C-----SPLINES NATURALES
    Z(0)=0.
    Z(N)=0.

C-----SUSTITUCION HACIA ATRAS
    DO I=N-1,1,-1
        Z(I)=(FT(I)-H(I)*Z(I+1))/HT(I)
    END DO

C--INTRODUCIR EL VALOR X DONDE SE REALIZARA LA APROXIMACION
C--AVERIGUAR A QUE SUBINTERVALO DE LA FUNCION SPLINE PERTENECE X
    WRITE(*,*) ' Introduce X (valor donde se realiza la aproximacion)'
    READ(*,*) XX
    DO I=N-1, 1,-1

        IF(XX-X(I) .GT. 0.) THEN
            GO TO 55
        END IF
    END DO

C-----CALCULO DE LA APROXIMACION
55 AA=(Z(I+1)-Z(I))/(6.*H(I))
    BB=Z(I)/2.
    CC=-H(I)*Z(I+1)/6.-H(I)*Z(I)/3.+(F(I+1)-F(I))/H(I)
    S=F(I)+(XX-X(I))*(CC+(XX-X(I))*(BB+(XX-X(I))*AA))
    WRITE(*,*) ' La aproximacion calculada vale', S
    STOP

C-----ADVERTENCIAS

```

```

100 WRITE(*,*)'Error de lectura en el fichero de datos'
STOP
END

```

15.4. Solución en MATHEMATICA

Un programa MATHEMATICA para el cálculo de la función Spline cúbica natural a partir de una tabulación dada en un fichero es el siguiente.

```

tabul = ReadList['c:/tabulacion.dat', Number, RecordLists ->True];

n = Dimensions[tabul][[1]];

x = Table[tabul[[i, 1]], i, 1, n];

y = Table[tabul[[i, 2]], i, 1, n];

For[i = 1, i <= n - 1, i++, (* Inicializo variables *)
  h[i] = x[[i + 1]] - x[[i]];
  b[i] = 6*(y[[i + 1]] - y[[i]])/h[i];
]

For[i = 2, i <= n - 1, i++,
  ht[i] = 2 *(h[i] + h[i - 1]);
  ft[i] = b[i] - b[i - 1];
]

A = IdentityMatrix[n - 2]; (* Matriz de coeficientes *)
For[i = 1, i <= n - 2, i++, A[[i, i]] = ht[i + 1] ];
For[i = 1, i <= n - 3, i++,
  A[[i, i + 1]] = A[[i + 1, i]] = h[i + 1];
]

bb = Table[0, {i, n - 2}]; (* términos independientes *)
For[i = 1, i <= n - 2, i++, bb[[i]] = ft[i + 1]];

z = LinearSolve[A, bb];
(* resuelvo sistema lineal e impongo splines naturales *)
z = AppendTo[z, 0];
z = PrependTo[z, 0];

For[i = 1, i <= n - 1, i++,
  AA[i] = (z[[i + 1]] - z[[i]])/(6*h[i]);
  BB[i] = z[[i]]/2;
  CC[i] = -h[i]*z[[i + 1]]/6 - h[i]*z[[i]]/3 +

```

```

    (y[[i + 1]] - y[[i]])/h[i];
]

For[i = 1, i <= n - 1, i++, (*Defino el spline*)
    s[i,t]=y[[i]]+(t-x[[i]])*(CC[i]+(t-x[[i]])*(BB[i]+
    (t-x[[i]])*AA[i]));
];

Print["La funcin spline S(t) vale:"] (*salida*)
For[i = 1, i <= n - 1, i++,
    Print["Si ", x[[i]], " <= t <=", x[[i + 1]], " S(t)=",
    Expand[s[i, t]]
]

For[i = 1, i <= n - 1, i++, (* gráficas *)
    graf[i] = Plot[s[i, t], {t, x[[i]], x[[i + 1]] },
    DisplayFunction ->Identity]
]

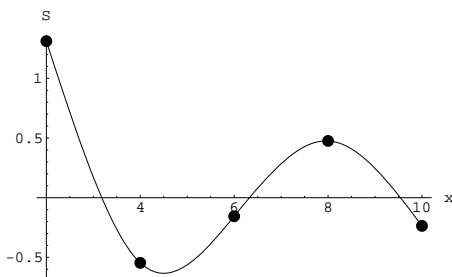
grafspline = { };
For[i = 1, i <= n - 1, i++, AppendTo[grafspline, graf[i]]];
grafspline = Show[grafspline,
DisplayFunction ->Identity];

graftabul = ListPlot[tabul, PlotStyle ->{PointSize[0.03]},
DisplayFunction ->Identity];

Show[graftabul, grafspline, AxesLabel ->{"x", "S"},
DisplayFunction ->{$DisplayFunction}]

```

Este programa, además de dar la expresión explícita de la función spline $S(x)$ muestra una gráfica, ver Figura 15.4, donde se superponen los puntos de la tabulación inicial y la gráfica de $S(x)$.



Capítulo 16

Derivación Numérica

16.1. Introducción

Para la aproximación de la derivada de una función $f(x)$ en un punto x_0 ; los paquetes de cálculo simbólico, por ejemplo el comando $D[f[x],x]$ de MATHEMATICA, proporcionan la expresión de la derivada $f'(x)$ y, por lo tanto, sólo queda evaluar dicha función en el punto dado para obtener $f'(x_0)$. Sin embargo, en el caso en que tan sólo se conocen los valores de la función en algunos puntos, aquí las fórmulas de derivación aproximada sí serán de gran utilidad.

En esta práctica se pretende programar un método numérico sencillo para evaluar la derivada de una función tabulada en un cierto conjunto de abscisas equiespaciadas.

Una forma de obtener expresiones en diferencias finitas para la derivada consiste (suponiendo que la función es suficientemente derivable) en combinar diferentes desarrollos de Taylor en un entorno de los puntos donde conocemos la función. Supongamos que tenemos la función $f(x)$ discretizada, es decir conocemos los valores $f_i := f(x_i)$ para las abscisas equiespaciadas $x_i = x_0 + ih$ con $i = 0, 1, \dots, n$ y siendo $h > 0$ la longitud de paso. Realizemos ahora los siguientes desarrollos de Taylor

$$f(x + 2h) = f(x) + 2hf'(x) + 2h^2f''(x) + \frac{4h^3}{3}f'''(x) + \frac{2h^4}{3}f^{(iv)}(x) + O(h^5), \quad (16.1)$$

$$f(x + h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{6}f'''(x) + \frac{h^4}{24}f^{(iv)}(x) + O(h^5), \quad (16.2)$$

$$f(x - h) = f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{6}f'''(x) + \frac{h^4}{24}f^{(iv)}(x) + O(h^5), \quad (16.3)$$

$$f(x - 2h) = f(x) - 2hf'(x) + 2h^2f''(x) - \frac{4h^3}{3}f'''(x) + \frac{2h^4}{3}f^{(iv)}(x) + O(h^5). \quad (16.4)$$

De la ecuación (16.2) se obtiene

$$f'(x) = \frac{f(x + h) - f(x)}{h} + O(h), \quad (16.5)$$

que se conoce con el nombre de *fórmula de diferencias hacia adelante*. De forma totalmente análoga, de la ecuación (16.3) podemos obtener la llamada *fórmula de diferencias hacia atrás*

$$f'(x) = \frac{f(x) - f(x-h)}{h} + O(h) . \quad (16.6)$$

Restando la ecuación (16.2) de la (16.3) se deduce la *fórmula de diferencias centradas*

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + O(h^2) . \quad (16.7)$$

Combinando (16.1), (16.2), (16.3) y (16.4) podemos obtener

$$f'(x) = \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h} + O(h^4) . \quad (16.8)$$

Otra forma de obtener la fórmula (16.8) es mediante extrapolación de Richardson a partir de (16.7).

16.2. Realización de la Práctica

- Hacer un pequeño programa OCTAVE para discretizar, con abscisas equiespaciadas, la función $f(x) = \frac{\sin x}{x}$ en el intervalo $[-10, 10]$ tomando los pasos $h = 1$, $h = 0,5$ y $h = 0,01$. Guardar las parejas (x_i, f_i) en tres ficheros diferentes según la longitud de paso h utilizada. NOTA: cuidado con la función en $x = 0!!!$.
- Realizar un programa OCTAVE que lea de un fichero los datos (x_i, f_i) para $i = 0, 1, \dots, n$ y que aproxime la derivada de la función $f(x)$ combinando las fórmulas (16.5), (16.6), (16.7) y (16.8) de manera adecuada y con el objetivo de tener un error de truncamiento lo más pequeño posible. Mostrar las gráficas superpuestas de la función $f'(x)$ y los resultados numéricos obtenidos.
- Supongamos ahora que la función $f(x)$ es desconocida y que el tercer fichero de datos del primer apartado, es decir con $h = 0,01$, que contiene los puntos (x_i, f_i) proviene de un experimento. Para simular con la computadora los errores inherentes a las medidas efectuadas en el supuesto experimento añadiremos una incertidumbre a cada uno de los valores de la función. Esto lo conseguiremos sumando una perturbación aleatoria¹ de manera que

$$\tilde{f}(x_i) = f(x_i) + 0,01 \text{ RAND} . \quad (16.9)$$

- (i) Guardar la nueva discretización (x_i, \tilde{f}_i) en un archivo y mostrar, mediante una gráfica, que los nuevos puntos perturbados no se desvían excesivamente de la función sin perturbar.

¹Ver el apéndice.

- (ii) Calcular, a través del programa confeccionado en el segundo apartado, la derivada de la función perturbada y comparar, mediante una gráfica, los resultados con los obtenidos para la función sin perturbar. En este apartado se tiene que comprobar que unos pequeños errores en los datos experimentales pueden llevar a obtener una derivada que difiere significativamente de la que de forma teórica pretendíamos obtener. Explicar de forma teórica dicho comportamiento.

16.2.1. Apéndice: generación de números aleatorios

Octave: La función **rand** de OCTAVE genera números aleatorios uniformemente distribuidos en el intervalo $[0, 1]$. Así, siendo $n, m \in \mathbb{N}$, **rand(n)** genera una matriz cuadrada de orden n con elementos aleatorios en el intervalo $[0, 1]$. **rand(n,m)** genera una matriz de orden $n \times m$ con elementos aleatorios en el intervalo $[0, 1]$.

Mathematica: La instrucción **Random[]** de MATHEMATICA genera un número real pseudoaleatorio en el intervalo $[0, 1]$.

Fortran: Una forma de generar números pseudoaleatorios en un programa FORTRAN es siguiendo los pasos que mostramos a continuación.

- Al principio del programa, hacer la declaración de variables

```
DOUBLE PRECISION RND, RAND, CRND
RND=1.0D0
```

- Tomar un valor cualquiera al azar para la variable CRND

```
READ(*,*) CRND
```

- Para modificar los valores de la función, utilizar el siguiente bucle

```
DO I=0,N
  RND=RND*CRND
  RND=DMOD(1.6807D4*RND, 2.147483647D9)
  RAND=RND*4.65661287524579692D-10
  F(I)=F(I)+0.01*RAND
END DO
```

- De esta forma, la variable RAND toma valores aleatorios entre 0 y 1.

Capítulo 17

Integración Romberg

17.1. Introducción

Uno de los problemas matemáticos más antiguos es el del cálculo del área que encierra una curva. Este problema desembocó en lo que hoy se conoce como cálculo integral.

Las fórmulas de integración numéricas, también conocidas como fórmulas de cuadratura, tienen como objetivo aproximar el valor de la integral de una función en un intervalo $\int_a^b f(x) dx$. El cálculo simbólico mediante un manipulador algebraico es de gran ayuda, ver por ejemplo el comando `Integrate[f[x], {x,a,b}]` del programa MATHEMATICA. Sin embargo esta ayuda sólo es posible cuando se conoce la expresión explícita de $f(x)$ y además cuando dicha función admite una primitiva. En el resto de los casos, por ejemplo cuando sólo se conoce los valores de la función en algunos puntos o cuando la primitiva de la función no es posible expresarla mediante funciones elementales, habrá que acudir a las fórmulas de integración numérica.

En la *integración Romberg* se usa la regla de los Trapecios compuesta para obtener aproximaciones preliminares de la integral, y luego el proceso de *extrapolación de Richardson* para mejorar las sucesivas aproximaciones.

Sabemos que la *regla de los Trapecios compuesta* viene dada por

$$\int_a^b f(x) dx = T_h(f) + O(h^2) = h \left(\frac{f_0}{2} + f_1 + f_2 + \cdots + f_{n-1} + \frac{f_n}{2} \right) + O(h^2) .$$

Por otra parte, consideremos dos aproximaciones numéricas $R(2h, k-1)$ y $R(h, k-1)$ con longitud de paso $2h$ y h respectivamente para la integral $\int_a^b f(x) dx$. Si se verifican los siguientes desarrollos en potencias de h

$$\int_a^b f(x) dx = R(h, k-1) + a_1 h^{2k} + a_2 h^{2k+2} + \cdots ,$$

$$\int_a^b f(x) dx = R(2h, k-1) + a_1(2h)^{2k} + a_2(2h)^{2k+2} + \dots,$$

es fácil comprobar que, utilizando la *extrapolación de Richardson*, una mejor aproximación de la integral viene dada por

$$\int_a^b f(x) dx = R(2h, k) + O(h^{2k+2}) = \frac{4^k R(h, k-1) - R(2h, k-1)}{4^k - 1} + O(h^{2k+2}).$$

Si ahora definimos la regla trapezoidal compuesta

$$T(j, 1) := T_h(f), \quad \text{siendo el paso } h = \frac{b-a}{2^{j-1}}, j = 1, 2, \dots \quad (17.1)$$

podemos obtener mejores aproximaciones para la integral a partir de la llamada *integración Romberg*

$$T(j, k) = \frac{4^{k-1}T(j, k-1) - T(j-1, k-1)}{4^{k-1} - 1} \quad \text{para } j \geq k. \quad (17.2)$$

Se obtiene de esta forma la siguiente tabla (sólo se representan los primeros elementos $T(j, k)$).

$T(1, 1)$			
	$T(2, 2)$		
$T(2, 1)$		$T(3, 3)$	
	$T(3, 2)$		$T(4, 4)$
$T(3, 1)$		$T(4, 3)$	
	$T(4, 2)$		
$T(4, 1)$			

17.2. Realización de la Práctica

Realizar un programa que implemente el algoritmo de integración Romberg descrito en la sección anterior, de manera que obtengamos la aproximación

$$\int_a^b f(x) dx \approx T(j, j).$$

- (i) Si se programa en FORTRAN, dimensionar una matriz 100×100 para almacenar los valores de $T(j, k)$. En OCTAVE no es necesario.
- (ii) Las entradas del programa tienen que ser

- El intervalo de integración $[a, b]$.
- Dos criterios para finalizar las iteraciones: **(1)** El número máximo permitido para j (por ejemplo 20). **(2)** La tolerancia (por ejemplo 10^{-5}). Pararemos las iteraciones cuando se verifique la condición

$$|T(j, j) - T(j-1, j-1)| < \text{tolerancia}.$$

(iii) La salida del programa debe ser la escritura en un fichero de las parejas $j, T(j, j)$ para $j = 0, 1, 2, \dots$

(iv) Escribier la función integrando $f(x)$ en un fichero:

Octave: Si se programa en OCTAVE, y por ejemplo la función $f(x)$ que se desea integrar es $f(x) = e^x - x$, escribiremos¹

```
function y=f(x)
    y=exp(x)-x;
endfunction
```

Cuando en el programa principal llama a la función para evaluarla en un argumento, se escribe (suponiendo que el argumento es 3) de la forma $f(3)$.

Fortran: Si se programa en FORTRAN, utilizar el *procedimiento* FUNCTION (después del END del programa principal) para definir la función $f(x)$ que se desea integrar. Por ejemplo, si $f(x) = e^x - x$, escribiremos

```
FUNCTION F(X)
    F=EXP(X)-X
    RETURN
END
```

El programa principal llama a la función F para evaluarla en un argumento (5 por ejemplo), de la forma $F(5)$.

(v) Notemos que el numerador de la fórmula (17.2) es la resta de dos cantidades que pueden llegar a ser muy cercanas, de manera que es posible el fenómeno de *cancelación sustractiva*² con la consiguiente pérdida de cifras significativas en el cálculo. Para evitar dichos errores de redondeo, es recomendable utilizar en el programa la expresión

$$T(j, k) = T(j, k - 1) + CORRECTOR ,$$

en lugar de la fórmula (17.2). Averiguar la expresión que tiene el término CORRECTOR.

(vi) Con el objetivo de disminuir al máximo el tiempo de cómputo, es recomendable obtener una relación entre $T(j, 1)$ y $T(j - 1, 1)$ que nos permita evitar la evaluación de la función $f(x)$ en las abscisas donde ya ha sido evaluada.

¹Ver el último Capítulo dedicado a OCTAVE.

²La representación en los ordenadores de números de punto flotante puede contener errores capaces de propagarse en todo un proceso y siendo capaces de inutilizar una solución. Cuando en una resta el sustraendo es casi igual al minuendo, se puede producir en dicha operación un gran error relativo. Este fenómeno se llama cancelación sustractiva.

17.3. La Capacidad Calorífica de los Sólidos

El paso fundamental para resolver la discrepancia entre la teoría clásica y los resultados experimentales concernientes a la capacidad calorífica asociada con las vibraciones de la estructura de los sólidos cristalinos, fue dado por el famoso físico A. Einstein (1907) utilizando la teoría cuántica. En 1912 Debye mejoró dicha teoría introduciendo los modos normales de vibración del cristal completo. En esta teoría, la capacidad calorífica a volumen constante $C_v = (\partial E / \partial T)_v$ (es decir la variación de la energía de vibración E de los átomos respecto de la temperatura T a volumen constante) se calcula de la forma

$$C_v = \frac{9Nk}{x_D^3} \int_0^{x_D} \frac{x^4 e^x}{(e^x - 1)^2} dx, \quad (17.3)$$

siendo N el número de átomos del cristal y $k = 1,381 \times 10^{-23} J/^{\circ}K$ la *constante de Boltzmann*. La variable auxiliar x_D está definida como $x_D = \Theta_D/T$, siendo Θ_D la *temperatura de Debye* característica del material, ver por ejemplo la tabla adjunta en la Figura 17.1 donde se muestran las temperatura de Debye para diferentes elementos químicos.

	Pb	Na	Ag	Cu	Al	Be
Θ_D	$86^{\circ}K$	$160^{\circ}K$	$220^{\circ}K$	$310^{\circ}K$	$380^{\circ}K$	$980^{\circ}K$

Figura 17.1: Temperaturas de Debye Θ_D para algunos elementos químicos.

Observar que, el integrando en la expresión de C_v no está definido para $x = 0$. Se tiene una indeterminación $0/0$ que hay que resolver.

Problema 17.1 Suponiendo que la muestra de material es de un mol (con lo que N vale el número de Avogadro $N = 6,023 \times 10^{23}$ átomos) y que está a temperatura ambiente (tomar por ejemplo $T = 293^{\circ}K$), utilizar el algoritmo de Romberg para obtener una aproximación de la capacidad calorífica C_v en los materiales de la tabla de la Figura 17.1.

Capítulo 18

Ecuaciones Integrales: Transmisión de Calor

18.1. Introducción

Consideremos dos placas de longitud infinita, anchura ℓ y separadas una distancia d . Supongamos que las placas están a temperaturas constantes T_1 y T_2 , siendo sus *emisividades* ϵ_1 y ϵ_2 respectivamente. Tomemos dos sistemas de referencia unidimensionales x_1 y x_2 con el origen en un punto cualquiera del centro de las placas y dirigidos hacia sus extremos finitos respectivamente. Se puede demostrar que las *radiosidades* locales B_1 y B_2 , es decir, la suma de las cantidades de calor (radiación) emitida y reflejada que abandonan una unidad de área por unidad de tiempo, verifican el siguiente sistema de ecuaciones integrales

$$B_1(x_1) = \epsilon_1 \sigma T_1^4 + (1 - \epsilon_1) \int_{-\ell/2}^{\ell/2} B_2(x_2) K(x_1, x_2, d) dx_2, \quad (18.1)$$

$$B_2(x_2) = \epsilon_2 \sigma T_2^4 + (1 - \epsilon_2) \int_{-\ell/2}^{\ell/2} B_1(x_1) K(x_1, x_2, d) dx_1, \quad (18.2)$$

donde $\sigma = 5,67 \times 10^{-8} w.m^2.K^{-4}$ es la constante de Stefan-Boltzmann y el *núcleo* K de las ecuaciones integrales es

$$K(x_1, x_2, d) = \frac{d^2}{2[d^2 + (x_1 - x_2)^2]^{3/2}}. \quad (18.3)$$

La integral que aparece en la ecuación (18.1) se denota por

$$I_1(x_1) = (1 - \epsilon_1) \int_{-\ell/2}^{\ell/2} B_2(x_2) K(x_1, x_2, d) dx_2, \quad (18.4)$$

y es llamada *irradiosidad*. Representa a la cantidad de calor incidente por unidad de área en un punto x_1 por unidad de tiempo. Obviamente, esta integral aparece

debido a la radiación emitida por todos los puntos de la otra placa. De forma totalmente análoga se tiene la irradiancia en la otra placa

$$I_2(x_2) = (1 - \epsilon_2) \int_{-\ell/2}^{\ell/2} B_1(x_1) K(x_1, x_2, d) dx_1 . \quad (18.5)$$

Los caudales netos de calor por unidad de longitud Q_1 y Q_2 que se deben proporcionar a dichas placas con el objetivo de mantenerlas a temperatura constante son

$$Q_1 = \int_{-\ell/2}^{\ell/2} (B_1(x_1) - I_1(x_1)) dx_1 , \quad (18.6)$$

$$Q_2 = \int_{-\ell/2}^{\ell/2} (B_2(x_2) - I_2(x_2)) dx_2 . \quad (18.7)$$

18.2. Realización de la Práctica

Realizar un programa FORTRAN que calcule las radiosidades e irradiancias locales B_1 , B_2 , I_1 e I_2 en las abscisas de las placas equiespaciadas con longitud¹ de paso $h = \ell/(2n)$, así como la calor recibida Q_1 y Q_2 en cada placa. La forma de proceder es la siguiente:

- (i) Implementar la regla de Simpson compuesta en una subrutina o en una función. Dicha regla es

$$\int_a^b f(x) dx \approx \frac{h}{3} \left[f(a) + 4 \sum_{\substack{i=1 \\ \Delta i=2}}^{2n-1} f(z_i) + 2 \sum_{\substack{i=2 \\ \Delta i=2}}^{2n-2} f(z_i) + f(b) \right] ,$$

donde $h = (b - a)/(2n)$ y $z_i = a + ih$ y Δi indica el salto de la variable i de los sumatorios. Una llamada a dicha función podría ser `SIMPSON(A,B,N,F)` donde, según la fórmula anterior, $A = a$, $B = b$, $N = n$ y F es la función a integrar $f(x)$.

- (ii) Las entradas del programa tienen que ser T_1 , T_2 , ϵ_1 , ϵ_2 , d , ℓ y n , además de dos constantes de control `ITMAX` y `TOL` (ver su definición en el apartado siguiente).
- (iii) **ALGORITMO DE SUSTITUCIONES SUCESIVAS:** El algoritmo iterativo que se utilizará para aproximar las funciones $B_1(x_1)$ y $B_2(x_2)$ solución del sistema de ecuaciones integrales (18.1) y (18.2) es el siguiente. En el primer paso se supone la aproximación de orden cero $B_1(x_1) \approx \epsilon_1 \sigma T_1^4$ y $B_2(x_2) \approx \epsilon_2 \sigma T_2^4$. Con estos valores, se calculan $I_1(x_1)$ e $I_2(x_2)$ dados por (18.4) y (18.5), lo

¹NOTA: La simetría de la geometría de las placas asociada a este problema nos permitirá suponer que $B_1(x_1)$ e $I_1(x_1)$ son funciones simétricas respecto del centro de la placa $x_1 = 0$ y, de forma análoga, $B_2(x_2)$ e $I_2(x_2)$ son simétricas respecto de $x_2 = 0$.

que nos permite estimar según (18.1) y (18.2) las aproximaciones de orden uno para $B_1(x_1)$ y $B_2(x_2)$. Estos valores son introducidos nuevamente en las integrales (18.4) y (18.5) para posteriormente obtener mediante (18.1) y (18.2) las aproximaciones de orden dos para $B_1(x_1)$ y $B_2(x_2)$ y así sucesivamente. El proceso iterativo terminará o bien cuando se llegue a un número máximo ITMAX de iteraciones permitido por el usuario, o bien cuando los valores de B_1 y B_2 no cambien más que un cierto valor de tolerancia TOL en dos iteraciones consecutivas.

- (iv) Cada vez que se aproximen, según el algoritmo descrito anteriormente, los valores de las integrales $I_1(x_1)$ e $I_2(x_2)$ dadas por (18.4) y (18.5) se realizará una llamada a la subrutina o función SIMPSON(A,B,N,F). Nótese que el número de aplicaciones de la regla de Simpson compuesta es n y que a su vez es también el número de abscisas en cada semiancho de las placas.

18.3. Existencia de Solución para Ecuaciones de Fredholm

En este apéndice se considerarán condiciones bajo las cuales existen soluciones $y(x)$ para la ecuación integral de tipo *Fredholm*

$$y(x) = f(x) + \alpha \int_a^b K(x, t, y(t)) dt, \quad (18.8)$$

siendo $\alpha \in \mathbb{R}$ un parámetro y K el núcleo. Para poder establecer un criterio de existencia de solución para la ecuación integral (18.8) supondremos lo siguiente:

1. La función $f(x)$ está acotada en el intervalo $[a, b]$, es decir, $|f(x)| < \bar{f}$ para todo $x \in [a, b]$.
2. El núcleo $K(x, y, z)$ es integrable y acotado de la forma $|K(x, y, z)| < \bar{K}$ en el dominio $D := \{(x, y, z) \in \mathbb{R}^3 : a \leq x, y \leq b, |z| < c\}$.
3. $K(x, y, z)$ satisface la condición de Lipschitz en D , es decir, $|K(x, y, z) - K(x, y, z')| < L|z - z'|$ con L constante.

Bajo las anteriores condiciones se asegura² que el *método de aproximaciones sucesivas* definido por

$$\begin{aligned} y_0(x) &= f(x) - f(a), \\ y_1(x) &= f(x) + \alpha \int_a^b K(x, t, y_0(t)) dt, \\ &\vdots \end{aligned}$$

²Ver por ejemplo el libro “*Introduction to nonlinear differential and integral equations*” de H.T. Davis, Dover Publications.

$$y_n(x) = f(x) + \alpha \int_a^b K(x, t, y_{n-1}(t)) dt ,$$

genera una sucesión de funciones $\{y_n(x)\}$ que converge uniformemente hacia una solución $y(x)$ de (18.8) cuando $n \rightarrow \infty$ para cualquier valor de α que verifique

$$|\alpha| < \frac{1}{M(b-a)} , \quad M := \max \left\{ L, \bar{K} \left(1 + \frac{\bar{f}}{|\alpha| \bar{K}(b-a)} \right) \right\} .$$

Algunos ejemplos de ecuaciones integrales y sus soluciones son:

1. La ecuación integral $y(x) = \int_0^1 (x-t)y^2(t) dt$ admite, además de la solución trivial $y(x) \equiv 0$, la solución $y(x) = 12x - 6$.
2. La ecuación integral $y(x) = -\frac{7}{3}x - \frac{3}{2} + \int_0^1 (x-t)y^2(t) dt$ admite la solución $y(x) = 2x - 3$.

Capítulo 19

Aproximación Trigonométrica y Coeficientes de Fourier

19.1. Introducción

Se define el espacio vectorial real \mathcal{T}_n como el generado por el conjunto de funciones

$$\left\{ \frac{1}{2}, \cos x, \sin x, \cos 2x, \sin 2x, \cos 3x, \sin 3x, \dots, \cos nx, \sin nx \right\}, \quad (19.1)$$

siendo $n \in \mathbb{N}$. \mathcal{T}_n es llamado el espacio de los *polinomios trigonométricos* de grado menor o igual que n . De esta forma, $g_n \in \mathcal{T}_n$ si existen $a_i, b_i \in \mathbb{R}$ llamados *coeficientes de Fourier*, tal que

$$g_n(x) = \frac{a_0}{2} + \sum_{m=1}^n a_m \cos mx + \sum_{m=1}^n b_m \sin mx, \quad (19.2)$$

Notemos que, para cualquier $n \in \mathbb{N}$, se verifica $g_n(x + 2\pi) = g_n(x)$ para todo x , es decir, g_n es una función periódica de periodo 2π .

El problema de la *aproximación trigonométrica* en el caso continuo se define de la forma siguiente: Dada una función $f : [0, 2\pi] \rightarrow \mathbb{R}$ continua y 2π -periódica, es decir $f(2\pi) = f(0)$, encontrar $f^* \in \mathcal{T}_n$ tal que

$$\|f - f^*\| = \min_{g_n \in \mathcal{T}_n} \|f - g_n\|_2,$$

siendo $\|\cdot\|_2$ la norma euclidea continua en el intervalo $[0, 2\pi]$. El problema de la aproximación trigonométrica es un problema de aproximación continua por

mínimos cuadrados.

Obsérvese que el conjunto de funciones (19.1) es ortogonal respecto del producto escalar euclideo continuo en el intervalo $[0, 2\pi]$. Ello es debido a que

$$\begin{aligned}
 \langle \sin \ell x, \sin mx \rangle &= \int_0^{2\pi} \sin \ell x \sin mx \, dx \\
 &= \frac{1}{2} \int_0^{2\pi} [\cos(\ell - m)x - \cos(\ell + m)x] \, dx , \\
 \langle \cos \ell x, \cos mx \rangle &= \int_0^{2\pi} \cos \ell x \cos mx \, dx \\
 &= \frac{1}{2} \int_0^{2\pi} [\cos(\ell - m)x + \cos(\ell + m)x] \, dx , \\
 \langle \sin \ell x, \cos mx \rangle &= \int_0^{2\pi} \sin \ell x \cos mx \, dx \\
 &= \frac{1}{2} \int_0^{2\pi} [\sin(\ell - m)x + \sin(\ell + m)x] \, dx .
 \end{aligned}$$

Por lo tanto

$$\langle \sin \ell x, \sin mx \rangle = \langle \cos \ell x, \cos mx \rangle = \langle \sin \ell x, \cos mx \rangle = \begin{cases} 0 & \text{si } \ell \neq m , \\ \pi & \text{si } \ell = m . \end{cases}$$

para todo $\ell, m \in \mathbb{N}$.

Sea

$$f^*(x) = \frac{a_0^*}{2} + \sum_{m=1}^n a_m^* \cos mx + \sum_{m=1}^n b_m^* \sin mx , \quad (19.3)$$

la aproximación trigonométrica de orden n de la función $f(x)$, también conocida como *desarrollo de Fourier* de orden n de f . Puesto que el conjunto de funciones (19.1) es ortogonal, las ecuaciones normales para los coeficientes de Fourier que resuelven el problema de la aproximación trigonométrica son diagonales, siendo su solución

$$\begin{aligned}
 a_0^* &= \frac{\langle 1/2, f \rangle}{\langle 1/2, 1/2 \rangle} = \frac{1}{\pi} \int_0^{2\pi} f(x) \, dx , \\
 a_m^* &= \frac{\langle \cos mx, f \rangle}{\langle \cos mx, \cos mx \rangle} = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos mx \, dx , \\
 b_m^* &= \frac{\langle \sin mx, f \rangle}{\langle \sin mx, \sin mx \rangle} = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin mx \, dx .
 \end{aligned}$$

19.2. Realización de la Práctica

Utilizar la regla de Simpson

$$\int_a^b g(x) \, dx \approx S_h(g) = \frac{h}{3}(g_0 + 4g_1 + 2g_2 + 4g_3 + \cdots + 4g_{n-1} + g_n)$$

con $h = (b - a)/n$ y $n = 20$ para hallar el desarrollo de Fourier $f^*(x)$ de orden 5 de la función $f(x) = x \exp(x)$ extendida a \mathbb{R} de forma 2π -periódica. Dar 6 gráficas donde se superpongan la función f y sus desarrollos de Fourier de orden $0, 1, \dots, 5$.

Capítulo 20

Distribución de Velocidades en un Fluido

20.1. Introducción

Se ha deducido¹ una expresión para la distribución de velocidades en un fluido en movimiento por una tubería circular lisa mediante una modificación de la llamada *longitud de mezcla de Prandtl*. En concreto, la velocidad adimensional axial u es una función de la distancia adimensional y desde la pared interior de la tubería dada por la expresión integral siguiente

$$u(y) = \int_0^y \frac{-1 + \sqrt{1 + 4c(y)d(y)}}{2c(y)} dy, \quad (20.1)$$

siendo

$$c(y) = \left[0,36y \left(1 - e^{-\phi \frac{y}{\bar{y}}} \right) \right]^2, \quad (20.2)$$

$$d(y) = 1 - \frac{y}{\bar{y}}, \quad (20.3)$$

donde \bar{y} es la distancia adimensional del interior de la tubería a su eje, cuya expresión viene dada por

$$\bar{y} = \frac{N_{Re}}{2} \sqrt{\frac{\bar{f}}{2}}, \quad (20.4)$$

y ϕ es una función de \bar{y} y de constantes empíricas dada por

$$\phi = \frac{\bar{y} - 60}{22}. \quad (20.5)$$

En la expresión (20.4), N_{Re} y \bar{f} representan el *número de Reynolds* y el *factor de fricción de Fanning* respectivamente, siendo estos parámetros adimensionales

¹W.N. GILL Y M. SCHEAR, *A modification of the momentum transport hypothesis*, A.I.Ch.E. Journal **7**, 61–65, (1961).

que dependen de las propiedades físicas y de la velocidad media del fluido, así como de la rugosidad superficial de las paredes de la tubería. La expresión del número de Reynolds es

$$N_{Re} = \frac{vD\rho}{\mu} , \quad (20.6)$$

siendo v , ρ y μ la velocidad media, densidad y viscosidad del fluido respectivamente y D el diámetro interior de la tubería. Para $2000 \leq N_{Re} \leq 50000$, \bar{f} satisface la *ecuación de Blasius*

$$\bar{f} = 0,079 N_{Re}^{-0,25} . \quad (20.7)$$

Se observa que el integrando de (20.1) es indeterminado en el límite inferior de integración $y = 0$. Una forma de evitar este problema consiste en multiplicar el numerador y el denominador de la función integrando en (20.1) por el conjugado del denominador $1 + \sqrt{1 + 4cd}$. De esta forma se obtiene la expresión equivalente

$$u(y) = \int_0^y \frac{2d(y)}{1 + \sqrt{1 + 4c(y)d(y)}} dy . \quad (20.8)$$

20.2. Integración Gauss-Chebichev

La fórmula de integración Gauss-Chebichev con $n+1$ puntos viene dada por

$$\int_{-1}^1 \frac{f(z)}{\sqrt{1-z^2}} dz \approx \frac{\pi}{n+1} \sum_{k=0}^n f(z_k) , \quad z_k = \cos \left(\frac{2k+1}{2n+2} \pi \right) . \quad (20.9)$$

Para poder utilizar la fórmula de integración Gauss-Chebichev (20.9) en la aproximación de la integral

$$\int_a^b g(x) dx , \quad (20.10)$$

la forma de proceder es la siguiente. Se transforma el intervalo de integración $a \leq x \leq b$ en el intervalo normalizado $-1 \leq z \leq 1$ mediante el cambio lineal de variable

$$x = \frac{b-a}{2} z + \frac{b+a}{2} .$$

De esta forma se tiene que

$$\int_a^b g(x) dx = \frac{b-a}{2} \int_{-1}^1 g \left(\frac{b-a}{2} z + \frac{b+a}{2} \right) dz = \frac{b-a}{2} \int_{-1}^1 \frac{f(z)}{\sqrt{1-z^2}} dz ,$$

siendo

$$f(z) = \sqrt{1-z^2} g \left(\frac{b-a}{2} z + \frac{b+a}{2} \right) .$$

20.3. Realización de la Práctica

Supongamos que el fluido tiene un número de Reynolds $N_{Re} = 10000$. Realizar un programa con el objetivo de aproximar $u(y_i)$ dada por (20.8) mediante integración de Gauss-Chebyshev (20.9) con $n = 15$ para los valores $y_i = 1 + ih$ con $i = 0, 1, \dots, 100$ y una longitud de paso $h = 10$.

Una forma de mejorar la exactitud numérica en las integraciones efectuadas es la siguiente. Denotando por u_i la aproximación obtenida para el valor de $u(y_i)$, se tiene que

$$u_i = u_{i-1} + \Delta u_i, \quad \Delta u_i = \int_{y_{i-1}}^{y_i} \frac{2d(y)}{1 + \sqrt{1 + 4c(y)d(y)}} dy.$$

Capítulo 21

La Condensación de Bose-Einstein

21.1. Introducción

El premio Nobel de Física del año 2001 fué otorgado a los científicos Carl E. Wieman (Universidad de Colorado, EEUU), Eric A. Cornell (National Institute of Standards & Technology, Colorado, EEUU) y Wolfgang Ketterle (Massachusetts Institute of Technology, Cambridge, Massachusetts, EEUU). Los galardonados crearon, de forma independiente, en el año 1995, por primera vez en el laboratorio un condensado de Bose-Einstein. De hecho, las relaciones existentes entre la condensación de Bose-Einstein y otros fenómenos de la física de bajas temperaturas como es la superfluidez (descubierta en el ^4He líquido por Kapitza en 1937) ha sido un tópico continuo en las investigaciones realizadas en los últimos 60 años.

En realidad, la descripción teórica del fenómeno físico de la condensación de Bose-Einstein se remonta al año 1924 cuando el físico Bose trató de deducir la fórmula de la radiación del cuerpo negro que Planck había obtenido en 1900, considerando la radiación como un gas de fotones. En el esquema utilizado por Bose está implícito el significativo hecho de que los fotones son partículas indistinguibles. Einstein, conocedor del trabajo de Bose, intuyó cómo extrapolar los resultados para fotones a otras partículas materiales de naturaleza indistinguible. De esta forma, en dos artículos publicados por Einstein en los años 1924 y 1925, utilizó lo que hoy en día se conoce como *estadística de Bose-Einstein* y predijo que, si un gas de este tipo de átomos se enfría hasta temperaturas suficientemente bajas, entonces todos sus átomos estarían en el estado de energía más baja (*estado fundamental*).

21.1.1. La física estadística cuántica

La formulación cuántica de la mecánica estadística¹, es decir, la visión microscópica de la termodinámica, proporciona la distribución de equilibrio en un sistema ideal de N *bosones* (partículas de spin entero). En concreto, el número medio de partículas n_ϵ en el estado de energía ϵ cuando el gas está sometido a una temperatura T viene dado por

$$n_\epsilon = \frac{1}{\exp \left[\frac{\epsilon - \mu}{K_B T} \right] - 1} , \quad (21.1)$$

donde $K_B = 1,381 \times 10^{-23} J/^\circ K$ es la constante de Boltzmann y μ el *potencial químico*, es decir, la energía que se le ha de dar al sistema para añadirle una partícula manteniendo su entropía y su volumen constantes.

Notemos que siempre es posible redefinir el potencial químico de forma que la energía del estado fundamental sea cero, de manera que, a partir de ahora supondremos esto. Observar que, puesto que n_ϵ es positivo y finito por definición, se tiene $\mu < 0$. Además, es habitual introducir z , la llamada *fugacidad* definida de la forma $z = \exp[\mu/(K_B T)]$, de manera que $0 \leq z < 1$. Así, en términos de la fugacidad, se puede escribir (21.1) como

$$n_\epsilon = \frac{1}{z^{-1} \exp \left[\frac{\epsilon}{K_B T} \right] - 1} . \quad (21.2)$$

Obsérvese que el estado fundamental corresponde al límite $z \rightarrow 1$ y aquí el miembro derecho de la ecuación (21.2) diverge, lo que indica que el número de partículas en ese estado es muy grande.

El número total N de partículas (bosones) de gas se escribe pues de la forma

$$N = N_0 + \sum_{\epsilon > 0} \frac{1}{z^{-1} \exp \left[\frac{\epsilon}{K_B T} \right] - 1} , \quad (21.3)$$

siendo $N_0 = z/(1 - z)$ en número de partículas en el estado fundamental y por lo tanto con energía $\epsilon = 0$. Existe una forma aproximada de calcular la suma (21.3) que se puede aplicar cuando el volumen V que ocupa el gas es macroscópico. El método consiste en pasar del discreto al continuo, es decir aproximar el sumatorio por una integral, introduciendo lo que se conoce como *densidad de estados* $\rho(\epsilon)$. De esta forma se tiene que

$$N = N_0 + \int_0^\infty \frac{\rho(\epsilon)}{z^{-1} \exp \left[\frac{\epsilon}{K_B T} \right] - 1} d\epsilon . \quad (21.4)$$

¹Ver por ejemplo K. HUANG, *Statistical Mechanics*, 2a ed. Wiley, New-York, 1987.

21.2. Un Gas de Bosones Ideal

Veamos cómo se aplica la teoría anterior en un caso particularmente sencillo. Supongamos que el gas de bosones es ideal y permanece en ausencia de fuerzas externas, de manera que su energía es puramente cinética. Se puede demostrar que, en este caso, la densidad de estados viene dada por

$$\rho(\epsilon) = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2}, \quad (21.5)$$

siendo m la masa de cada partícula y $\hbar = h/(2\pi)$ con $h = 6,626 \times 10^{-34} J.s.$ la *constante de Planck*. Introduciendo (21.5) en (21.4) se obtiene

$$N = N_0 + \frac{V}{\left(\frac{2\pi\hbar^2}{mK_B T} \right)^{3/2}} B_{3/2}(z), \quad (21.6)$$

siendo $B_{3/2}(z)$ la integral de Bose² de índice $3/2$.

Puede ocurrir que la temperatura T sea tan grande que $N_0 = 0$. Pero, por el contrario, si la temperatura es suficientemente baja entonces N_0 puede ser muy grande. En el caso extremo en que $z \rightarrow 1$ se puede definir una *temperatura crítica* T_c definida por

$$T_c = \frac{2\pi\hbar^2}{m} \left(\frac{N}{B_{3/2}(1)V} \right)^{2/3} \quad (21.7)$$

de manera que, si $T > T_c$ entonces no hay una ocupación del estado fundamental que sea macroscópicamente apreciable. Por el contrario, si $T < T_c$, un número macroscópico de partículas pueden ocupar el estado fundamental. Obsérvese que el fenómeno que se acaba de describir implica que el gas puede ser considerado como una mezcla de dos fases; una fase constituida por aquellas partículas que se encuentran en el estado fundamental y la otra fase por el resto de las partículas. Una forma interesante de ver el fenómeno consiste en reescribir la ecuación (21.6) de la forma

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_c} \right)^{3/2}, \quad (21.8)$$

con lo que la T_c puede interpretarse como la temperatura de transición entre las dos fases del gas: para $T \ll T_c$ prácticamente todas las partículas del gas están en el estado fundamental y se ha formado la llamada condensación de Bose-Einstein³.

²Ver el apéndice.

³Nótese las semejanzas con la clásica condensación gas-líquido. En realidad, la diferencia básica entre el cambio de fase de la clásica condensación gas-líquido y la condensación de Bose-Einstein reside en que la primera es debida a la interacción entre las partículas mientras que la segunda es un fenómeno puramente cuántico y se puede dar en un gas ideal sin ningún tipo de interacción entre partículas.

21.3. Realización de la Práctica

Calcular la temperatura crítica T_c a la cual se produce la transición Bose-Einstein para un gas ideal de $N = 10^{23}$ bosones cuyas moléculas tienen una masa $m = 10^{-19}$ Kg y ocupan un volumen $V = 0,01 \text{ m}^3$. Aproximar en dicho cálculo $B_{3/2}(1)$ mediante integración de Gauss-Hermite aumentando el número de puntos.

21.3.1. Apéndice: Integrales de Bose

Se define la *integral de Bose de índice n* como

$$B_n(z) = \frac{1}{\Gamma(n)} \int_0^\infty \frac{x^{n-1}}{z^{-1}e^x - 1} dx, \quad (21.9)$$

donde $\Gamma(n)$ es la *función Gamma de Euler* que está definida por

$$\Gamma(n) = \int_0^\infty t^{n-1} e^{-t} dt, \text{ con } n > 0.$$

Por ejemplo $\Gamma(1/2) = \sqrt{\pi}$. Se puede demostrar la siguiente fórmula de recurrencia

$$\Gamma(n+1) = n\Gamma(n).$$

Además, la función Gamma generaliza a la función factorial en el sentido de que

$$\Gamma(n+1) = n!, \text{ para todo } n \in \mathbb{N}.$$

Las integrales de Bose (21.9) han sido estudiadas por J.E. Robinson, Phys. Rev. 83, pag. 678 (1951).

Capítulo 22

Ecuaciones Diferenciales Ordinarias: Método RK4

22.1. Introducción

Una de las herramientas matemáticas más efectivas de que disponemos actualmente para modelizar y explicar la naturaleza son, sin lugar a dudas, las *ecuaciones diferenciales* (o más generalmente los *sistemas dinámicos*). Una resolución exacta de las ecuaciones diferenciales no es posible en general. De esta forma, son necesarias vías alternativas tales como los *métodos numéricos* o la *teoría cualitativa* para poder estudiar las ecuaciones diferenciales. Por lo que respecta a los métodos numéricos, el algoritmo que presentamos es uno de los más populares para aproximar la solución del problema de valor inicial

$$\begin{cases} \dot{x} &= f(t, x, y) , & x(t_0) = x_0 , \\ \dot{y} &= g(t, x, y) , & y(t_0) = y_0 , \end{cases} \quad (22.1)$$

donde el punto denota, como es usual, derivación respecto a la variable independiente t , es decir, $\dot{x} = dx/dt$, $\dot{y} = dy/dt$. Supondremos que las funciones f y g satisfacen las condiciones del Teorema de existencia y unicidad de la solución del problema de valor inicial (22.1).

Entenderemos por resolver de forma numérica el problema de Cauchy (22.1) en el intervalo $[t_0, t_n]$ la obtención de una tabla del tipo

t_0	t_1	t_2	\cdots	t_n
x_0	x_1	x_2	\cdots	x_n
y_0	y_1	y_2	\cdots	y_n

donde x_i e y_i sean una aproximación numérica de los valores exactos $x(t_i)$ e $y(t_i)$ respectivamente. Además, siempre consideraremos una discretización equiespaciada de la variable independiente $t_i = t_0 + ih$ para $i = 0, 1, \dots, n$, siendo $h > 0$

la longitud de paso.

El *método de Runge-Kutta* RK4 que utilizaremos para aproximar las soluciones de (22.1) viene dado por las siguientes iteraciones

$$\begin{cases} x_{i+1} &= x_i + \frac{h}{6}(r_1 + 2r_2 + 2r_3 + r_4) , \\ y_{i+1} &= y_i + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4) , \end{cases} \quad (22.2)$$

siendo

$$\begin{aligned} r_1 &= f(t_i, x_i, y_i) , \\ r_2 &= f\left(t_i + \frac{h}{2}, x_i + \frac{h}{2}r_1, y_i + \frac{h}{2}k_1\right) , \\ r_3 &= f\left(t_i + \frac{h}{2}, x_i + \frac{h}{2}r_2, y_i + \frac{h}{2}k_2\right) , \\ r_4 &= f(t_i + h, x_i + hr_3, y_i + hk_3) , \\ k_1 &= g(t_i, x_i, y_i) , \\ k_2 &= g\left(t_i + \frac{h}{2}, x_i + \frac{h}{2}r_1, y_i + \frac{h}{2}k_1\right) , \\ k_3 &= g\left(t_i + \frac{h}{2}, x_i + \frac{h}{2}r_2, y_i + \frac{h}{2}k_2\right) , \\ k_4 &= g(t_i + h, x_i + hr_3, y_i + hk_3) . \end{aligned}$$

22.2. El Método RK4 en \mathbb{R}^n

Por supuesto, el método RK4 puede ser generalizado utilizando notación vectorial a ecuaciones diferenciales ordinarias definidas en \mathbb{R}^n . La exposición de la sección anterior se ha restringido únicamente al caso $n = 2$.

De esta forma, consideremos el problema de Cauchy

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}) , \quad \mathbf{x}(t_0) = \mathbf{x}_0 . \quad (22.3)$$

Aquí, $t \in \mathbb{R}$ es la variable independiente y $\mathbf{x} = (x_1, \dots, x_n)^t \in \mathbb{R}^n$ es la variable vectorial dependiente y $\mathbf{f}(t, \mathbf{x}) = (f_1(t, \mathbf{x}), \dots, f_n(t, \mathbf{x}))^t \in \mathbb{R}^n$. Se tienen en definitiva las funciones incógnita $x_i(t)$, con $i = 1, \dots, n$. La condición inicial es el valor vectorial $\mathbf{x}_0 = (x_1^0, \dots, x_n^0)^t \in \mathbb{R}^n$. La iteración RK4 con longitud de paso h consiste en

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \frac{h}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4) , \quad (22.4)$$

siendo

$$\begin{aligned} \mathbf{k}_1 &= \mathbf{f}(t_i, \mathbf{x}_i) , \\ \mathbf{k}_2 &= \mathbf{f}\left(t_i + \frac{h}{2}, \mathbf{x}_i + \frac{h}{2}\mathbf{k}_1\right) , \end{aligned}$$

$$\begin{aligned} \mathbf{k}_3 &= \mathbf{f}\left(t_i + \frac{h}{2}, \mathbf{x}_i + \frac{h}{2}\mathbf{k}_2\right), \\ \mathbf{k}_4 &= \mathbf{f}(t_i + h, \mathbf{x}_i + h\mathbf{k}_3). \end{aligned}$$

Problema 22.1 Programar el método RK4 en \mathbb{R}^d dado en (22.4) mediante una función en OCTAVE cuya entrada sea:

- Los tiempos inicial t_0 y final t_n .
- El número de puntos n .
- El vector de condiciones iniciales $x_0 \in \mathbb{R}^d$.

La salida de la función debe ser:

- El vector fila de tiempos equiespaciados.
- La matriz $(n+1) \times d$ que contiene en cada fila las aproximaciones del vector de estado del sistema.

La función $f(t, x) \in \mathbb{R}^d$ que define el sistema de ecuaciones diferenciales se debe programar a parte en un fichero f.m que será utilizado por la anterior función RK4.

22.3. Ciclos Límite

Sea x una magnitud física que evoluciona con el tiempo t . Para fijar ideas podemos pensar en que $x(t)$ es la posición en función del tiempo para una partícula. Supongamos que la partícula realiza oscilaciones no lineales regidas por la siguiente ecuación diferencial de segundo orden

$$\frac{d^2x}{dt^2} + f(x)\frac{dx}{dt} + g(x) = 0. \quad (22.5)$$

Nótese que una partícula cuya ecuación del movimiento esté regida por la ecuación (22.5) no puede tener soluciones periódicas cuyas órbitas estén en la región en la que $f(x)$ sea de signo constante puesto que físicamente estas zonas corresponden a regiones de amortiguamiento positivo o negativo. Soluciones periódicas, es decir cuyas trayectorias formen curvas cerradas en el plano de las fases, desempeñan un papel importante en muchos fenómenos físicos. Una solución periódica a menudo representa una clase de “estado final”, hacia el cual todas las soluciones vecinas tienden. La existencia de estas soluciones periódicas (ciclos límite) sólo pueden explicarse por la consideración de términos no lineales en las ecuaciones que gobiernan el movimiento.

La ecuación diferencial de segundo orden (22.5) puede transformarse (mediante el cambio habitual $y = dx/dt$) en el siguiente sistema de ecuaciones

diferenciales de primer orden

$$\begin{cases} \frac{dx}{dt} = y, \\ \frac{dy}{dt} = -yf(x) - g(x). \end{cases} \quad (22.6)$$

El anterior sistema está definido en el plano de posiciones x y velocidades y que habitualmente se conoce tanto en matemáticas como en mecánica con el nombre de *plano de las fases*.

Un caso particular de la ecuación (22.5) es la conocida ecuación de *van der Pol*¹. Esta ecuación viene definida de la forma

$$\frac{d^2x}{dt^2} + \mu(x^2 - 1)\frac{dx}{dt} + x = 0, \quad (22.7)$$

siendo μ un parámetro real positivo por razones físicas.

Un estudio detallado de la ecuación (22.7) muestra la existencia de un único *ciclo límite* (órbita periódica y aislada) en el plano de las fases. Si denotamos el ciclo límite por C y su periodo por T , se tiene que

$$C := \{(x(t), y(t)) \in \mathbb{R}^2 : x(t+T) = x(t), y(t+T) = y(t)\}.$$

Además, puede demostrarse que el ciclo límite C es *estable*, es decir, todas las órbitas excepto el punto de equilibrio $(0,0)$ tienden asintóticamente hacia el ciclo límite C cuando $t \rightarrow \infty$, ver Figura 22.1.

Problema 22.2 Considerar la ecuación de van der Pol (22.7).

- Tomar la ecuación (22.7) con el parámetro $\mu = 0,1$ y escribirla como un sistema de primer orden en el plano de las fases $x - y$.
- Utilizar el algoritmo de Runge-Kutta de cuarto orden RK4 para evidenciar de forma numérica el ciclo límite. Experimentar con valores iniciales $(x(0), y(0)) = (x_0, y_0)$ diferentes y dibujar (con un software adecuado) las órbitas obtenidas (x_i, y_i) para $i = 0, 1, \dots$.
- Cuando el parámetro $\mu = 0,1$, el ciclo límite de van der Pol se puede aproximar por una circunferencia centrada en el origen, es decir, $C \approx \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = R^2\}$. Estimar el radio R de dicha circunferencia a partir de las gráficas realizadas en el apartado anterior.

¹Balthasar van der Pol (1889-1959), científico holandés especialista en aspectos teóricos de radio-ingeniería, inició el estudio de la ecuación que lleva su nombre en los años veinte, estimulando con ello a Liénard y otros a investigar la teoría matemática de las oscilaciones automantenidas en mecánica no lineal. En su ecuación, la variable x denota la intensidad de corriente eléctrica en un oscilador de triodo. Experimentalmente se observa que el flujo de corriente es periódico. Este flujo periódico sólo puede ser explicado mediante la consideración de una ecuación no lineal.

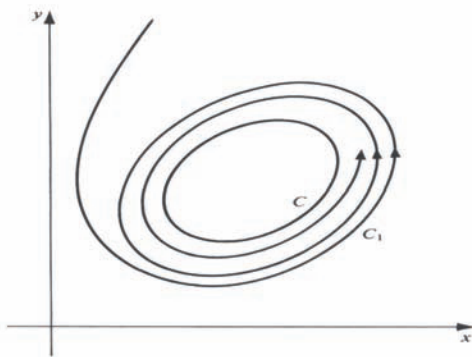


Figura 22.1: Ciclo límite C . Notar como la trayectoria C_1 se acerca espiralando hacia el ciclo límite.

22.4. El Problema del Centro-Foco

Consideremos una partícula de masa m que describe un movimiento *armónico simple* de frecuencia ω . Aplicando la segunda ley de Newton, la ecuación diferencial que rige la posición x de la partícula en función del tiempo t viene dada por

$$\frac{d^2x}{dt^2} = -\omega^2 x .$$

Tomando unidades adecuadas se puede conseguir que $\omega = 1$. Con estas unidades, el sistema diferencial que describe a un oscilador armónico simple es

$$\frac{dx}{dt} = -y , \quad \frac{dy}{dt} = x . \quad (22.8)$$

Este sistema es trivial de integrar puesto que $dy/dx = -x/y$. Entonces, separando las variables e integrando, obtenemos que las órbitas en el plano de las fases $x-y$ vienen dadas por circunferencias concéntricas $x^2 + y^2 = C^2$. Como el origen $(0, 0)$ posee un entorno totalmente relleno de órbitas cerradas del sistema diferencial, se dice que el punto crítico $(0, 0)$ es un *centro*.

Consideremos ahora una perturbación polinomial del sistema (22.8) definida de la forma

$$\frac{dx}{dt} = -y + f(x, y) , \quad \frac{dy}{dt} = x + g(x, y) , \quad (22.9)$$

siendo $f, g \in \mathbb{R}_n[x, y]$ polinomios reales en las variables x e y de grado $n \geq 2$.

Se sabe que, las órbitas en el plano de las fases del sistema perturbado (22.9) pueden seguir siendo periódicas, de manera que el punto crítico $(0, 0)$ continua siendo un centro, o bien espiralan alrededor del punto crítico en cuyo caso se

dice que el punto crítico $(0, 0)$ es un *foco*, ver las Figuras 22.2 y 22.3. Averiguar si el punto $(0, 0)$ es un centro o un foco para el sistema (22.9) es un problema clásico y difícil.

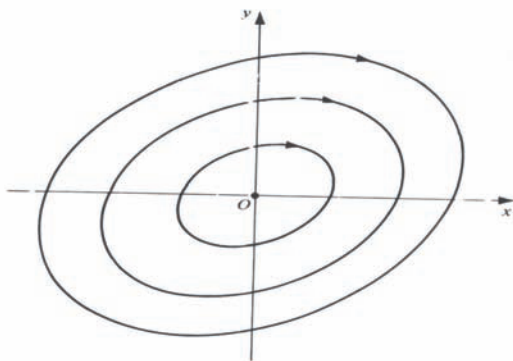


Figura 22.2: El origen de coordenadas $(0, 0)$ es un punto crítico de tipo centro.

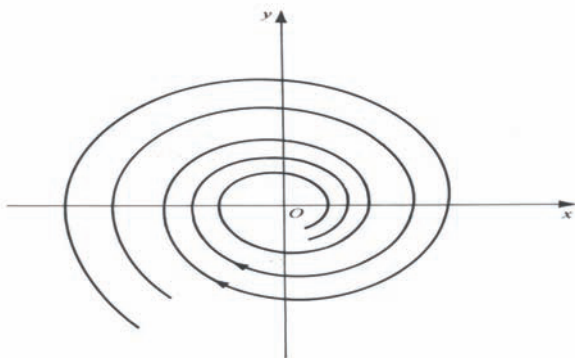


Figura 22.3: El origen de coordenadas $(0, 0)$ es un punto crítico de tipo foco.

Problema 22.3 Considerar los sistemas diferenciales siguientes

$$\frac{dx}{dt} = -y - x^2, \quad \frac{dy}{dt} = x, \quad (22.10)$$

$$\frac{dx}{dt} = -y - x^3, \quad \frac{dy}{dt} = x, \quad (22.11)$$

- Utilizando el algoritmo de Runge-Kutta de cuarto orden *RK4* con diferentes condiciones iniciales (x_0, y_0) , averiguar de forma numérica si el origen $(0, 0)$ es un centro o un foco para los sistemas (22.10) y (22.11). Una forma gráfica de averiguarlo es mediante el dibujo (con un software adecuado) de las órbitas computadas de forma numérica.
- Es posible que al experimentar con valores iniciales $(x(0), y(0)) = (x_0, y_0)$ diferentes se obtenga en el programa un error de *OVERFLOW*. Dar una explicación para este error.

22.5. El Péndulo de Wilberforce

El péndulo de Wilberforce² es un popular instrumento de demostración en los laboratorios de física y mecánica, ver Figura 22.4.

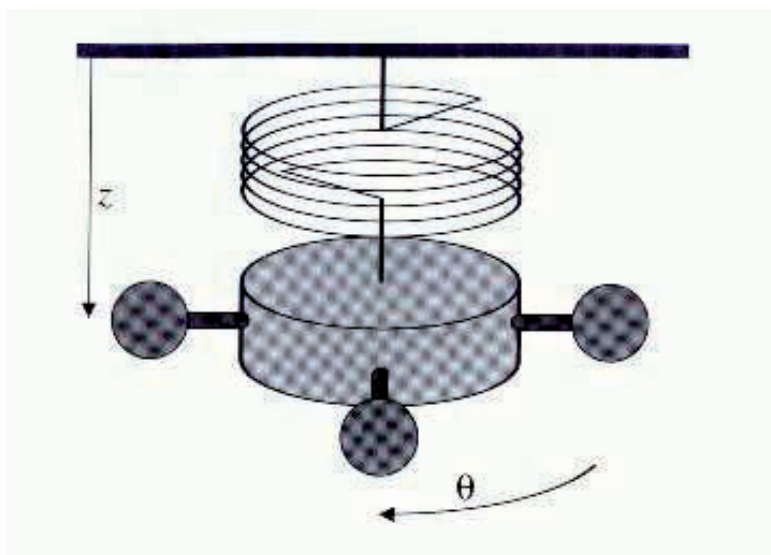


Figura 22.4: El péndulo de Wilberforce con sus dos modos de oscilación.

²R.E. Berg and T. S. Marshall, *Wilberforce pendulum oscillations and normal modes*, Am. J. Phys. **59**, 32–38 (1991).

Para simplificar el sistema, supondremos que el movimiento se caracteriza por una fuerza lineal restauradora tanto en la dirección lineal z como en la angular θ . Añadiendo un factor de acoplamiento ϵ entre los dos modos de oscilación vertical y de torsión, las ecuaciones del movimiento del péndulo de Wilberforce vienen dadas por dos ecuaciones diferenciales ordinarias de segundo orden lineales y acopladas para los dos grados de libertad del péndulo:

$$m\ddot{z} + kz + \frac{1}{2}\epsilon\theta = 0, \quad I\ddot{\theta} + \delta\theta + \frac{1}{2}\epsilon z = 0, \quad (22.12)$$

siendo m la masa del péndulo, I el momento de inercia, k y δ las constantes elásticas longitudinales y de torsión del muelle.

Problema 22.4 *Demostrar de forma numérica y utilizando el algoritmo RK4 que cuando se entra en resonancia, es decir, cuando las frecuencias longitudinales $w_z = \sqrt{k/m}$ y de torsión $w_\theta = \sqrt{\delta/I}$ del péndulo de Wilberforce son iguales, entonces el movimiento del péndulo alterna de forma periódica entre movimientos únicamente longitudinales y únicamente de torsión. Para conseguirlo, mostrar de forma superpuesta las gráficas de $z(t)$ y de $\theta(t)$, tomando, por ejemplo, los valores siguientes: $m = 0,5 \text{ Kg}$, $k = 5 \text{ N/m}$, $I = 10^{-4} \text{ Kg.m}^2$, $\delta = 10^{-3} \text{ N.m}$, $\epsilon = 10^{-2} \text{ N}$. Además de ir probando distintas condiciones iniciales $z(0) = z_0$, $\theta(0) = \theta_0$, $\dot{z}(0) = \dot{z}_0$ y $\dot{\theta}(0) = \dot{\theta}_0$, utilizar también los tipos de condiciones iniciales siguientes: (i) $z(0) = 10 \text{ cm}$, $\theta(0) = 0$; (ii) $z(0) = 0 \text{ cm}$, $\theta(0) = 2\pi$.*

22.6. El Modelo de Lorenz: Convección de Fluidos

El modelo de Lorenz es un modelo áltamente simplificado del fenómeno de convección en fluidos. El modelo fue introducido en 1963 por el meteorólogo Edward Lorenz, quien estaba interesado en modelizar la convección en la atmósfera. Lo que Lorenz demostró fue que incluso un sistema muy sencillo de ecuaciones diferenciales puede tener soluciones cuyo comportamiento es esencialmente impredecible dando lugar a lo que hoy se conoce como teoría del caos. Una detallada derivación del modelo de Lorenz se puede ver, por ejemplo, en el libro de Robert C. Hilborn titulado *Chaos and Nonlinear Dynamics: an introduction for scientists and engineers*.

En términos físicos, Lorenz modeliza una capa horizontal de fluido que se calienta en el fondo y se enfría en la parte superior de la capa. La parte inferior y superior de la capa se mantiene a temperaturas constantes T_i y T_s respectivamente, con $T_i > T_s$. Este tipo de sistemas fue estudiado experimentalmente por Bénard en 1900. No fue hasta 1916 que Lord Rayleigh obtuvo algún resultado teórico. Hoy en día se conoce como sistema de fluido de Rayleigh-Bénard.

Si la diferencia de temperaturas $\Delta T = T_i - T_s$ no es muy grande, el fluido permanece estacionario y la transferencia de calor se produce desde abajo hacia

arriba por conducción. La tendencia del fluido (menos denso) en la parte inferior a elevarse se cancela debido a la pérdida de calor hacia el entorno y el rozamiento debido a la viscosidad. En estas condiciones estacionarias, la temperatura del fluido varía linealmente respecto de la posición vertical.

Por el contrario, si la diferencia de temperaturas ΔT aumenta suficientemente, las fuerzas flotantes llegan a ser suficientemente grandes como para vencer al rozamiento viscoso y se desarrollan corrientes estacionarias dentro del fluido. En estas condiciones, la transferencia de calor desde la parte inferior hacia la parte superior de la capa de fluido se realiza por convección, es decir, debida al movimiento de la masa del fluido. Como resultado se obtiene una circulación de fluido que es estable en el tiempo.

Con un aumento posterior de la diferencia de temperaturas ΔT , tanto las corrientes de fluido como las diferencias de temperaturas dentro del fluido dejan de ser estacionarias y comienzan a variar con el tiempo.

22.6.1. Las ecuaciones de Lorenz

El modelo de Lorenz está basado en una gran simplificación de las ecuaciones fundamentales de la dinámica de fluidos, es decir, de las ecuaciones de Navier-Stokes. En resumen, el movimiento del fluido y las diferencias de temperaturas resultantes en el fluido se pueden expresar en términos de tres variables $x(t)$, $y(t)$, $z(t)$ que dependen del tiempo t . La variable x está relacionada con la dependencia temporal de la llamada función de corriente cuya derivada respecto de las variables espaciales da las componentes del campo de vectores velocidad del fluido. Las variables y , z , están relacionadas con la dependencia temporal de las desviaciones de temperatura respecto del régimen lineal que se tiene en la situación no convectiva. En concreto, y es proporcional a la diferencia de temperaturas entre las partes del fluido que suben y las que bajan a una determinada altura y z es proporcional a la desviación de la linealidad en la temperatura del fluido como una función de la posición vertical.

Usando esas variables, las ecuaciones de Lorenz adoptan la forma de las siguientes tres ecuaciones diferenciales no lineales y acopladas

$$\dot{x} = p(y - x) , \quad \dot{y} = rx - y - xz , \quad \dot{z} = -bz + xy , \quad (22.13)$$

siendo p , r y b parámetros del modelo. En concreto, p es el llamado *número de Prandtl* que se define como el cociente entre la viscosidad cinemática del fluido y su coeficiente de difusión térmica. Por lo tanto p depende del fluido que se estudia. El parámetro r es proporcional al llamado *número de Rayleigh* que es una medida adimensional de la diferencia de temperaturas ΔT entre la parte inferior y superior de la capa de fluido. Cuando ΔT aumenta, r aumenta. Usualmente, r es utilizado como un parámetro de control. El parámetro b está relacionado con el cociente entre la altura de la capa de fluido y el tamaño de los rollos de convección que efectúa el fluido.

22.6.2. Realización de la práctica

A pesar de la sencillez del modelo de Lorenz (22.13), sus soluciones exhiben un comportamiento muy complejo (llamado caótico) bajo ciertas condiciones. No se conoce la solución general $(x(t), y(t), z(t))$ de (22.13) de modo que (22.13) se debe resolver de forma numérica. Utilizaremos el método RK4 en \mathbb{R}^3 y los valores de los parámetros $p = 10$, $b = 8/3$.

- El estado no convectivo del fluido viene determinado por la solución estacionaria $(x(t), y(t), z(t)) = (0, 0, 0)$ para todo t . Para valores pequeños del parámetro r , es decir, para ΔT pequeño, ¿predice el modelo de Lorenz el estado no convectivo? Responder suponiendo el valor de los parámetros $r = 0,5$ y que el sistema se encuentra en el estado inicial $(x(0), y(0), z(0)) = (0, 1, 0)$. Obtener las gráficas de las funciones $x(t)$, $y(t)$ y $z(t)$ en función del tiempo en el intervalo $0 \leq t \leq 10$ así como la trayectoria $(x(t), y(t), z(t))$ en el espacio de fases.
- Si $r < 1$, se puede demostrar que todas las soluciones de (22.13), independientemente de cual es su condición inicial, se aproximan en un tiempo suficiente al punto $(0, 0, 0)$. Mostrarlo numéricamente tomando varios valores de $r \in (0, 1)$ y diversas condiciones iniciales $(x(0), y(0), z(0))$.
- Si $r > 1$, pero no muy grande (por ejemplo menor que 13) se puede ver que toda solución que comience cercana al punto $(0, 0, 0)$ se aleja de este y se acerca (con un tiempo suficiente) a alguno de los puntos

$$A_{\pm} = (\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$$

dependiendo de cual es la condición inicial. La explicación en términos físicos consiste en que cuando r aumenta hacia un valor superior a 1 (es decir, incrementamos la diferencia de temperaturas ΔT entre la parte inferior y superior de la capa de fluido) el sistema ya no tiende al estado no convectivo representado por el punto $(0, 0, 0)$ sino que tiende hacia un estado convectivo estacionario representado por los puntos A_{\pm} donde el signo \pm indica una rotación de convección en el sentido horario o antihorario.

Mostrar este comportamiento de forma numérica usando el valor $r = 2$ y las dos condiciones iniciales $(x(0), y(0), z(0)) = (0, \pm 1, 0)$.

- Si continuamos aumentando el valor de r , se observa que no pasa nada dramático hasta que r alcanza un valor aproximadamente de $r \approx 13,93$ donde se observa que se desarrollan regiones entorno de los puntos A_{\pm} de modo que las soluciones con condiciones iniciales en dichas regiones se alejan de los puntos A_{\pm} . El comportamiento temporal de estas soluciones es bastante complicado. Obtener un ejemplo numérico con los valores del parámetro $r = 25$ y la condición inicial $(x(0), y(0), z(0)) = (0, -5, 15)$. Realizar la gráfica de la trayectoria $(x(t), y(t), z(t))$ en el espacio de fases. ¿Se acerca ahora la trayectoria a alguno de los puntos A_{\pm} ?

- Como el comportamiento de las soluciones en el apartado anterior ha sido muy complicado, seguimos aumentando el valor de r . Por ejemplo, tomamos $r = 160$ y la condición inicial $(x(0), y(0), z(0)) = (0, -5, 15)$ del apartado anterior. Mostrar mediante las gráficas de las funciones $x(t)$, $y(t)$ y $z(t)$ que, el comportamiento de dichas funciones es periódico. En concreto, $x(t)$ oscila de forma simétrica entorno del valor $x = 0$. ¿Qué significa físicamente?
- Nos embarcamos a continuación en la cuestión crucial de decidir si el modelo de Lorenz exhibe comportamiento caótico para valores grandes de r . Aquí, comportamiento caótico se traduce por sensibilidad del sistema a las condiciones iniciales. Nos preguntamos si, para valores grandes de r , dos soluciones con condiciones iniciales muy próximas divergen alejándose al pasar el tiempo. Si esto ocurre se dice que el sistema es caótico y, aunque proviene de un modelo determinista (sistema de ecuaciones diferenciales), resulta imposible predecir el estado futuro del sistema debido a la precisión limitada innata en los aparatos que miden físicamente las condiciones iniciales (este fenómeno se conoce popularmente como efecto mariposa).

Demostrar numéricamente que el sistema de Lorenz con el valor del parámetro $r = 143$ es caótico mostrando las gráficas de la función $z(t)$ para dos condiciones iniciales cercanas, por ejemplo, $(x(0), y(0), z(0)) = (20, 0, 163)$ y $(x(0), y(0), z(0)) = (20, 0, 166)$ y viendo que, tras unas pocas oscilaciones, las funciones son completamente diferentes.

Capítulo 23

Raíces de Ecuaciones: Método de Newton-Raphson

23.1. Introducción

Uno de los cálculos que más a menudo se realizan en el cálculo científico o técnico es el de encontrar las raíces de una ecuación. Son muchos los problemas en ciencia e ingeniería que se pueden modelar matemáticamente como una ecuación $f(x) = 0$, siendo f una función de la variable x . El algoritmo que presentamos en esta sección es un método iterativo. Partiremos de una aproximación inicial x_0 del cero exacto x^* de f y posteriormente construiremos una sucesión de números reales $\{x_n\}_{n=0}^{\infty}$ que converja hacia x^* cuando $n \rightarrow \infty$. En el *método de Newton-Raphson*, dicha sucesión, está generada a partir de la iteración siguiente

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \quad (23.1)$$

Notemos que, para poder utilizar el método de Newton-Raphson, hemos de suponer que la función $f(x)$ es derivable en un entorno de la raíz x^* .

Puesto que el método de Newton-Raphson genera una sucesión de números reales $\{x_n\}_{n=0}^{\infty}$, es claro que en su programación se deberá dar al menos un criterio de parada. En la programación que se pide en esta práctica se darán tres criterios de parada tal y como se detallan a continuación.

Problema 23.1 Programar el método de Newton-Raphson dado en (23.1) mediante una función en OCTAVE cuya entrada sea:

- El punto inicial de iteración $x_0 \in \mathbb{R}$.
- El número máximo de iteraciones **nmax**. De este modo, si el programa no finaliza, se calculará como mucho el valor x_{nmax} y se tomará $x^* \approx x_{nmax}$. Finalizar la ejecución del programa de esta forma indica alguna anomalía en su ejecución.
- La precisión **prec**. Así, si $|f(x_n)| < \text{prec}$, se parará el cálculo y se tomará $x^* \approx x_n$.
- La tolerancia **tol**. De este modo, si $|x_n - x_{n-1}| < \text{tol}$, se parará el cálculo y se tomará $x^* \approx x_n$.

La salida de la función debe ser:

- El vector fila de dimensión variable que contenga a los términos x_n de la sucesión de Newton-Raphson.

Las funciones $f(x)$ y $f'(x)$ se deben programar a parte en sendos fichero *f.m* y *fp.m* que serán utilizado por la anterior función principal.

23.2. Ecuaciones de Estado de los Gases

Es conocido que la ecuación de estado de los gases ideales es

$$P = \frac{nRT}{V}, \quad (23.2)$$

siendo n el número de moles del gas, T su temperatura y $R = 8,3143 JK^{-1}mol^{-1}$ la constante de los gases ideales.

Al tratar gases reales, debemos tomar en consideración las fuerzas intermoleculares y las dimensiones finitas de las moléculas. Las fuerzas intermoleculares son de corto alcance y disminuyen rápidamente con la distancia entre moléculas. Por consiguiente, la presión P de un gas real será tanto más cercana a la de un gas ideal, ecuación (23.2), cuanto menor sea el número de moléculas por unidad de volumen. Esto sugiere que expresemos la presión de un gas real en términos de la siguiente serie de potencias en n/V

$$P = \frac{nRT}{V} + A \frac{n^2}{V^2} + B \frac{n^3}{V^3} + C \frac{n^4}{V^4} + \dots \quad (23.3)$$

La ecuación (23.3) se puede considerar como la ecuación de estado de un gas real. A , B , C , ... son cantidades características de cada gas y son llamados segundo, tercer, cuarto, etc.. *coeficientes viriales*. Estos coeficientes son funciones de la temperatura T y dependen de la intensidad de las fuerzas moleculares.

Midiendo P a temperaturas y volúmenes diferentes se pueden obtener los coeficientes viriales $A(T)$, $B(T)$, $C(T)$, ... experimentalmente.

Los métodos de la mecánica estadística permiten un cálculo de los coeficientes viriales. En las secciones siguientes veremos algunos casos particulares de ecuaciones de estado relacionadas con desarrollos del virial.

23.2.1. Ecuación de Dieterici

Una ecuación empírica de estado para gases reales, propuesta por Dieterici, viene dada por

$$P(V - nb) \exp\left(\frac{na}{VRT}\right) = nRT, \quad (23.4)$$

siendo los coeficientes a y b característicos del gas.

Problema 23.2 Se tienen 10 moles de Nitrógeno gaseoso y suponemos que se verifica la ecuación de Dieterici (23.4) con $a = 140,4$ y $b = 0,039$. Aproximar, utilizando el método el Newton-Raphson, el volumen que ocupa sabiendo que está a temperatura $T = 300K$ y presión $P = 10^5 Pa$.

23.2.2. El Módulo de elasticidad en la ecuación de van der Waals

El *módulo de elasticidad* de volumen a temperatura constante de una sustancia se define de la forma

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_T.$$

Es fácil ver que un gas real que sigue la *ecuación de van der Waals*

$$\left(P + \frac{n^2 a}{V^2} \right) (V - nb) = nRT, \quad (23.5)$$

donde a y b son las *constantes de van der Waals* características del gas, tiene un módulo de elasticidad

$$\kappa_T = \frac{V^2(V - nb)^2}{nRTV^3 - 2n^2a(V - nb)^2}. \quad (23.6)$$

Problema 23.3 Demostrar la expresión de κ_T dada en (23.6).

Problema 23.4 Se sabe que el Oxígeno tiene unas constantes de van der Waals $a = 137,4$ y $b = 0,031$. Aproximar mediante el método el Newton-Raphson el volumen que ocupan 30 moles de Oxígeno a temperatura $T = 310K$ sabiendo que su módulo de elasticidad es $\kappa_T = 2,74 Pa^{-1}$ y el volumen ocupado es superior a $200000 m^3$.

NOTA: Tanto en el Problema 23.2 como en el Problema 23.4, al aplicar el método de Newton-Raphson es necesaria una estimación inicial del volumen V_0 para poder comenzar las iteraciones. Esta estimación se puede realizar o bien a través de un teorema de convergencia de la sucesión generada con la iteración de Newton-Raphson (ver Teorema 8.2) o bien tomando el volumen V_0 que predeciría la suposición de gas ideal.

23.3. Un Problema Sobre la Profundidad de Absorción

Consideremos un medio material poroso que absorbe un cierto gas. Tomemos el eje x perpendicular a la superficie plana del medio con sentido positivo hacia el interior del medio. Sea $\rho(x)$ la concentración de gas a una profundidad x del medio y supongamos que la concentración de gas $\rho(0)$ en la superficie del medio es constante. Para fijar ideas tomaremos $\rho(0) = 1$. Sea $a(x)$ la tasa de absorción de gas a profundidad x una función conocida. Obviamente, el gas se difunde únicamente hasta una profundidad ℓ y estamos interesados en hallarla. De esta forma, el modelo planteado es

$$\rho''(x) = -a(x), \quad 0 < x < \ell, \quad (23.7)$$

$$\rho(0) = 1, \quad (23.8)$$

$$\rho(\ell) = \rho'(\ell) = 0. \quad (23.9)$$

Es fácil probar que la única solución $\rho(x)$ del problema de valor final formado por las ecuaciones (23.7) y (23.9) viene dada por

$$\rho(x) = - \int_x^\ell (s - x)a(s) ds. \quad (23.10)$$

Sólo falta averiguar el valor ℓ de modo que se verifique la condición (23.8). Notemos que dicho valor de ℓ debe ser una raíz de la función

$$f(\ell) = 1 - \rho(0) = 1 + \int_0^\ell sa(s)ds. \quad (23.11)$$

23.3.1. Realización de la práctica

- ¿Es necesario algún método numérico para hallar la profundidad de absorción ℓ si la tasa de absorción de gas es $a(x) = 2 - x$? Hallar el valor exacto de ℓ en caso de ser posible.
- Supongamos ahora que la función $sa(s)$ no admite primitiva expresable mediante funciones elementales. Entonces, no es posible utilizar la ecuación (23.11) para hallar ℓ de forma exacta. En su defecto, se puede utilizar una estrategia numérica con el objetivo de aproximar ℓ como veremos a continuación. En primer lugar, supongamos que se dispone de una subrutina

de algún método numérico (por ejemplo el comando `NDSolve[]` de `MATHEMATICA` o el algoritmo `RK4`) que integra el problema reducido de valor final (23.7) y (23.9), es decir,

$$\rho''(x) = -a(x), \quad 0 < x < \ell_n, \quad \rho(\ell_n) = \rho'(\ell_n) = 0.$$

Por supuesto, al resolver el anterior problema, como comenzamos a integrar a partir de $x = \ell_n$ y llegamos hasta $x = 0$, si se utiliza `RK4` se hará con una longitud de paso negativa. A partir de ahora se asumirá que se dispone de la subrutina `RK4`(ℓ_n), cuyo argumento de entrada es ℓ_n y la salida es ρ_n , siendo ρ_n una aproximación al valor inicial $\rho(0)$, es decir, $\rho(0) \approx \rho_n$. El método numérico para hallar ℓ consiste en introducir dos valores distintos de inicio ℓ_0 y ℓ_1 que sean las primeras aproximaciones de ℓ y posteriormente generar una sucesión $\{\ell_n\}_{n \in \mathbb{N}}$ que tenga la propiedad de convergencia $\lim_{n \rightarrow \infty} \ell_n = \ell$ mediante el método de la secante aplicado a la función (23.11) dada por $f(\ell) = 1 - \rho(0)$. A la vez que se va calculando la sucesión $\{\ell_n\}_{n \in \mathbb{N}}$, necesitaremos computar también otra sucesión $\{\rho_n\}_{n \in \mathbb{N}}$ de números que aproximan a $\rho(0)$ tal que verifiquen $\lim_{n \rightarrow \infty} \rho_n = \rho(0)$.

El algoritmo funciona de la siguiente manera. En primer lugar se introducen los datos de entrada ℓ_0 y ℓ_1 . Posteriormente se calculan los valores $\rho_0 = \text{RK4}(\ell_0)$ y $\rho_1 = \text{RK4}(\ell_1)$. El resto de valores de las sucesiones $\{\ell_n\}_{n \in \mathbb{N}}$ y $\{\rho_n\}_{n \in \mathbb{N}}$ se calculan en un bucle. Ahí, los valores de ℓ_n se calculan con el método de la secante aplicado a la función $f(\ell) = 1 - \rho(0)$ donde $\rho(0)$ se ha sustituido por el último de los valores obtenidos de ρ_n . Los valores de ρ_n en el bucle se obtienen de la forma $\rho_n = \text{RK4}(\ell_n)$, siendo ℓ_n el valor calculado en el cálculo anterior.

Para que el algoritmo funcione bien sólo es necesario usar algún criterio de parada en el anterior bucle. Usaremos dos criterios de parada de modo que el bucle finaliza cuando $|1 - \rho_n| < \epsilon_1$ o bien $|\ell_n - \ell_{n-1}| < \epsilon_2$, siendo ϵ_1 y ϵ_2 las tolerancias introducidas a priori como datos de entrada del algoritmo y pueden ser, por ejemplo, del orden de 10^4 .

Comprobar que el algoritmo funciona bien utilizando la misma función $a(x)$ del apartado anterior.

23.4. Método de Newton para Sistemas de Ecuaciones

El método de Newton-Raphson para aproximar las soluciones de la ecuación $f(x) = 0$ se basa en el concepto de linearizar la función f . Se puede generalizar dicho método para aproximar por un método iterativo las soluciones de un sistema de ecuaciones.

Consideremos un sistema de n ecuaciones con n incógnitas x^1, x^2, \dots, x^n .

Dicho sistema puede ser escrito de la forma $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, siendo

$$\mathbf{x} = (x^1, x^2, \dots, x^n)^T.$$

La aplicación $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tiene las funciones componentes $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ para $i = 1, 2, \dots, n$, es decir,

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x}))^T.$$

Sea \mathbf{x}^* un cero de la función \mathbf{f} , es decir, $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$. Supongamos que \mathbf{x}_0 es un punto próximo a \mathbf{x}^* y que \mathbf{f} es diferenciable en \mathbf{x}_0 . Entonces, en primera aproximación se tiene

$$\mathbf{0} = \mathbf{f}(\mathbf{x}^*) \approx \mathbf{f}(\mathbf{x}_0) + D\mathbf{f}(\mathbf{x}_0)(\mathbf{x}^* - \mathbf{x}_0),$$

siendo $D\mathbf{f}(\mathbf{x}_0)$ la matriz Jacobiana de la función \mathbf{f} en el punto \mathbf{x}_0

$$D\mathbf{f}(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f_1(\mathbf{x}_0)}{\partial x^1} & \dots & \frac{\partial f_1(\mathbf{x}_0)}{\partial x^n} \\ \vdots & & \vdots \\ \frac{\partial f_n(\mathbf{x}_0)}{\partial x^1} & \dots & \frac{\partial f_n(\mathbf{x}_0)}{\partial x^n} \end{pmatrix}.$$

Si la matriz Jacobiana $D\mathbf{f}(\mathbf{x}_0)$ es no singular, es decir, si $\det D\mathbf{f}(\mathbf{x}_0) \neq 0$, entonces se puede despejar de la ecuación $\mathbf{0} = \mathbf{f}(\mathbf{x}_0) + D\mathbf{f}(\mathbf{x}_0)(\mathbf{x}_1 - \mathbf{x}_0)$ el vector \mathbf{x}_1 . De esta manera se obtiene

$$\mathbf{x}_1 = \mathbf{x}_0 - (D\mathbf{f}(\mathbf{x}_0))^{-1}\mathbf{f}(\mathbf{x}_0),$$

donde $(D\mathbf{f}(\mathbf{x}_0))^{-1}$ denota la matriz inversa de la Jacobiana. Generalizando el proceso de forma iterativa se tiene el *método de Newton-Raphson*

$$\mathbf{x}_{i+1} = \mathbf{x}_i - (D\mathbf{f}(\mathbf{x}_i))^{-1}\mathbf{f}(\mathbf{x}_i) \quad \text{con } i = 0, 1, \dots, \quad (23.12)$$

para aproximar las soluciones del sistema $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.

Nótese que, desde un punto de vista numérico y con el objetivo de no calcular matrices inversas, es más conveniente reescribir el método de Newton-Raphson de la forma

$$D\mathbf{f}(\mathbf{x}_i) \mathbf{x}_{i+1} = D\mathbf{f}(\mathbf{x}_i) \mathbf{x}_i - \mathbf{f}(\mathbf{x}_i) \quad \text{con } i = 0, 1, \dots. \quad (23.13)$$

Puesto que el método de Newton-Raphson genera una sucesión de vectores $\{\mathbf{x}_i\}_{i=0}^{\infty} \subset \mathbb{R}^n$, es claro que en su programación se deberá dar al menos un criterio de parada. En la programación que se pide en esta práctica se darán tres criterios de parada tal y como se detallan a continuación.

Problema 23.5 Programar el método de Newton-Raphson dado en (23.12) o en (23.13) mediante una función en OCTAVE cuya entrada sea:

- El punto inicial de iteración $x_0 \in \mathbb{R}^d$.
- El número máximo de iteraciones n_{max} . De este modo, si el programa no finaliza, se calculará como mucho el vector $x_{n_{max}}$ y se tomará $x^* \approx x_{n_{max}}$. Finalizar la ejecución del programa de esta forma indica alguna anomalía en su ejecución.
- La precisión **prec**. Así, si $\|f(x_n)\| < \text{prec}$, se parará el cálculo y se tomará $x^* \approx x_n$.
- La tolerancia **tol**. De este modo, si $\|x_n - x_{n-1}\| < \text{tol}$, se parará el cálculo y se tomará $x^* \approx x_n$.

La salida de la función debe ser:

- La matriz con d columnas y número de filas variable que contenga en cada fila a los vectores x_n de la sucesión de Newton-Raphson.

La función $f(x)$ y la matriz jacobiana $Df(x)$ se deben programar a parte en sendos fichero $f.m$ y $Df.m$ que serán utilizado por la anterior función principal.

23.4.1. Equilibrio en sistema masa-muelle

Consideremos el sistema bidimensional masa-muelle de la Figura 23.1.

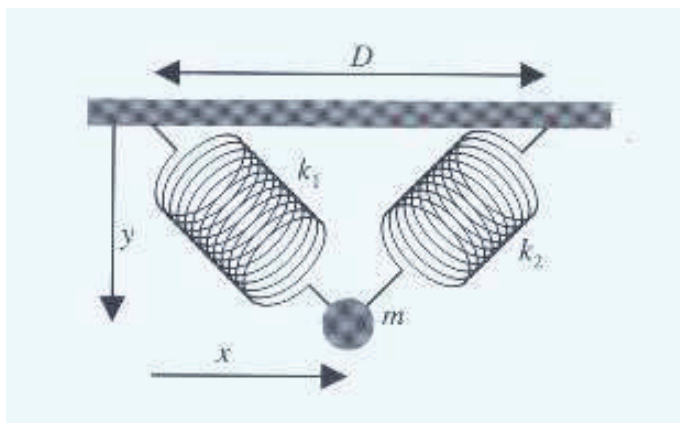


Figura 23.1: Sistema sistema bidimensional masa-muelle en equilibrio estable.

Es fácil demostrar que, tomando el sistema de referencia $x - y$ dado en la

Figura 23.1, la energía potencial U del sistema viene dada por

$$U(x, y) = \frac{1}{2}k_1(\sqrt{x^2 + y^2} - L_1)^2 + \frac{1}{2}k_2(\sqrt{(x - D)^2 + y^2} - L_2)^2 - mgy ,$$

siendo D la separación horizontal entre los puntos de enganche de los dos muelles, k_1 y k_2 las constantes elásticas de los muelles, L_1 y L_2 las longitudes naturales de los muelles, m la masa de la bola y $g = 9,8m/s^2$ la aceleración de la gravedad.

Puesto que la fuerza total \vec{F} que actúa sobre la masa m es conservativa, verifica $\vec{F} = -\vec{\nabla}U$, siendo $\vec{\nabla} = (\partial/\partial x, \partial/\partial y)$ el operador vectorial gradiente. De este modo, cuando el sistema se encuentre en equilibrio estático se verificará $\vec{\nabla}U = 0$. De forma equivalente, las coordenadas (x^*, y^*) de la masa m en equilibrio estático verificarán el sistema de ecuaciones

$$\frac{\partial U}{\partial x}(x^*, y^*) = 0 , \quad \frac{\partial U}{\partial y}(x^*, y^*) = 0 . \quad (23.14)$$

Problema 23.6 Supongamos que $k_1 = 10N/m$, $L_1 = L_2 = D = 0,1m$ y $m = 0,1Kg$. Obtener las gráficas de las funciones $x^*(k_2)$ y $y^*(k_2)$ para el rango de valores $0 \leq k_2 \leq 10$.

Para conseguirlo utilizaremos el método de Newton-Raphson para resolver los 101 sistemas (23.14) que se obtienen tomando los valores equiespaciados $k_2^{(i)} = i10/n$, con $i = 0, 1, \dots, n$, siendo $n = 100$. Además, puesto que los valores de $k_2^{(i)}$ consecutivos están cercanos, se tomará la condición inicial (x_0, y_0) para resolver un sistema como la aproximación (x^*, y^*) obtenida en la resolución del sistema anterior.

La forma de proceder en el Problema 23.6 se conoce como *Método de Continuación* y sirve para obtener una gráfica de $x(\xi)$ donde se verifica la ecuación $f(x; \xi) = 0$ siendo ξ un parámetro de la ecuación.

Capítulo 24

Método de los Elementos Finitos

24.1. Introducción: Modelo Matemático

Se considera la ecuación diferencial ordinaria lineal de segundo orden

$$a(x)\frac{d^2u}{dx^2} + b(x)\frac{du}{dx} + c(x)u(x) = g(x) \quad (24.1)$$

definida para x perteneciente al *dominio* $[0, L] \subset \mathbb{R}$. Supondremos que las funciones $a(x)$, $b(x)$, $c(x)$ y $g(x)$ son continuas en dicho dominio. Si la función $a(x)$ tiene signo constante en el intervalo $[0, L]$, es decir, (24.1) es una ecuación *elíptica*, entonces, (24.1) puede reescribirse como una ecuación diferencial *autoadjunta* de la forma

$$-\frac{d}{dx} \left(a_1(x) \frac{du}{dx} \right) + a_0(x)u = f(x) , \quad (24.2)$$

siendo

$$\begin{aligned} a_1(x) &:= \exp \left(\int_0^x \frac{b(t)}{a(t)} dt \right) , \\ a_0(x) &:= \frac{-c(x)a_1(x)}{a(x)} , \\ f(x) &:= \frac{-g(x)a_1(x)}{a(x)} . \end{aligned}$$

La forma autoadjunta (24.2) no sólo tiene ventajas desde el punto de vista numérico como se verá posteriormente, si no que también describe muchos procesos físicos. La ecuación (24.2) será acompañada de alguna de las siguientes *condiciones de contorno*

$$u(0) = \bar{U}_0 , \quad u(L) = \bar{U}_L , \quad (24.3)$$

$$u(0) = \bar{U}_0, \quad a_1(L) \frac{du}{dx}(L) = Q_L, \quad (24.4)$$

$$-a_1(0) \frac{du}{dx}(0) = Q_0, \quad a_1(L) \frac{du}{dx}(L) = Q_L, \quad (24.5)$$

$$u(0) = \bar{U}_0, \quad \left(a_1(L) \frac{du}{dx}(L) + \beta u(L) \right) = \gamma, \quad (24.6)$$

siendo \bar{U}_0 , \bar{U}_L , Q_0 , Q_L , β y γ valores fijados.

24.2. Solución Numérica

En esta sección buscaremos, mediante el método de los *elementos finitos*, aproximaciones de la solución del *problema de contorno* (24.2) con alguna de las condiciones (24.3), (24.4), (24.5) o (24.6).

24.2.1. Mallado del dominio

Se realizará un *mallado* del dominio $[0, L]$, es decir, una subdivisión de dicho dominio en *elementos*. Identificaremos algunos puntos de cada elementos del mallado y les llamaremos *nodos*¹. De esta forma, subdividimos el dominio $[0, L]$ en N elementos $\Omega_k \subset [0, L]$, con $k = 1, \dots, N$. Más concretamente, tomaremos $\Omega_k = [x_k, x_{k+1}]$, siendo x_k y x_{k+1} los nodos del elemento Ω_k y su anchura $h_k = x_{k+1} - x_k$. De esta forma sólo tomaremos dos nodos por cada elemento.

Posteriormente se realizará la *numeración global de los nodos*. En nuestro caso se tomará el nodo global 1 en $x = 0$ y, numerando los nodos en orden creciente y de manera consecutiva se llegará a que el nodo $x = L$ corresponderá al nodo global $N + 1$. Se tiene pues que el nodo global k -ésimo se encuentra en el punto de coordenada $x = x_k$.

24.2.2. Las ecuaciones de los elementos: método variacional

Consideremos el elemento $\Omega_k = [x_A, x_B]$, donde en nuestro caso se tiene $x_A = x_k$ y $x_B = x_{k+1}$ debido a que se han tomado dos nodos en cada elemento, pero esto no tenía por qué haber sido así.

Aproximaremos la solución exacta $u(x)$ de (24.2) en el elemento Ω_k por la función $U_k(x)$. Se tiene pues que

$$u(x) \approx U(x) := \begin{cases} U_1(x) & \text{si } x \in \Omega_1, \\ U_2(x) & \text{si } x \in \Omega_2, \\ \vdots & \\ U_N(x) & \text{si } x \in \Omega_N, \end{cases} \quad (24.7)$$

¹En función del grado de los polinomios que aproximarán a la solución $u(x)$ de (24.2) en cada elemento se distinguirán más o menos nodos dentro de cada elemento

donde $U_k(x)$ se tomará como una combinación lineal de un conjunto finito de n funciones de aproximación $\{\phi_{kj}(x)\}_{j=1}^n$, es decir,

$$U_k(x) := \sum_{j=1}^n u_{kj} \phi_{kj}(x) . \quad (24.8)$$

Habitualmente las funciones de aproximación $\phi_{kj}(x)$ se toman polinomiales de grado $n - 1$. Veremos posteriormente que las constantes u_{kj} serán los valores que toma la solución aproximada U_k en los nodos del elemento Ω_k . Además, se tomará n igual al número de nodos en el elemento Ω_k , que para nosotros será 2.

Definamos el *residuo* $R(u)$ de la ecuación diferencial (24.2) como

$$R(u) := -\frac{d}{dx} \left(a_1(x) \frac{du}{dx} \right) + a_0(x)u - f(x) . \quad (24.9)$$

Por supuesto, si $u(x)$ es la solución exacta de (24.2) entonces $R(u) = 0$. Sin embargo $R(U_k) \neq 0$ para $x \in \Omega_k$. Las relaciones que deben verificar las n constantes u_{kj} las obtendremos imponiendo que $U_k(x)$ aproxime a $u(x)$ en Ω_k de la “mejor manera posible”, donde esto se puede hacer de varias formas, dando lugar a diferentes métodos de elementos finitos. Aquí se utilizará un *método variacional*, que consiste en imponer que el residuo ponderado valga cero. En concreto se impondrá que $\int_{\Omega_k} w_j(x) R(u) dx = 0$ para $j = 1, \dots, n$, donde $w_j(x)$ son ciertas *funciones peso*. Las diferentes elecciones de las funciones peso w_j dan también lugar a diferentes métodos de elementos finitos. El método de *Galerkin* de elementos finitos consiste en tomar $w_j(x) = \phi_{kj}(x)$.

En definitiva, para hallar las constantes u_{kj} con $j = 1, \dots, n$, se impondrá la anulación de los residuos ponderados

$$\int_{x_A}^{x_B} w_j(x) \left[-\frac{d}{dx} \left(a_1(x) \frac{du}{dx} \right) + a_0(x)u - f(x) \right] dx = 0 , \quad j = 1, \dots, n .$$

Utilizando el método de integración por partes, es claro que

$$\int w_j(x) \left[-\frac{d}{dx} \left(a_1(x) \frac{du}{dx} \right) \right] dx = -w_j(x) a_1(x) \frac{du}{dx} + \int a_1(x) \frac{dw_j}{dx} \frac{du}{dx} dx ,$$

de manera que, pasaremos a la llamada *forma débil* de la anulación de residuos ponderados

$$\int_{x_A}^{x_B} \left(a_1(x) \frac{dw_j}{dx} \frac{du}{dx} + a_0(x)w_j(x)u - w_j(x)f(x) \right) dx - \left[w_j(x) a_1(x) \frac{du}{dx} \right]_{x_A}^{x_B} = 0 .$$

Definiendo

$$Q_A := -a_1(x_A) \frac{du}{dx}(x_A) , \quad Q_B := a_1(x_B) \frac{du}{dx}(x_B) , \quad (24.10)$$

donde el cambio de signo proviene de una interpretación física de Q_A y Q_B en términos de flujos, se tiene la forma débil

$$0 = \int_{x_A}^{x_B} \left(a_1(x) \frac{dw_j}{dx} \frac{du}{dx} + a_0(x) w_j(x) u - w_j(x) f(x) \right) dx - w_j(x_A) Q_A - w_j(x_B) Q_B . \quad (24.11)$$

24.2.3. Aproximación de la solución

Aproximaremos la solución exacta $u(x)$ de (24.2) en el elemento $\Omega_k = [x_A, x_B]$ por la función $U_k(x)$ dada en (24.2). Impondremos que $U_k(x)$ sea un polinomio de grado $n - 1$ que interpole a la función $u(x)$ en los nodos del elemento Ω_k . De esta forma se tendrá en particular que $U_k(x_A) = u(x_A) := u_A$ y $U_k(x_B) = u(x_B) := u_B$. Tomaremos el número de nodos $n = 2$ sobre cada elemento Ω_k , de manera que $U_k(x)$ será el polinomio interpolador de Lagrange

$$U_k(x) = \sum_{j=1}^2 u_{kj} \phi_{kj}(x) , \quad (24.12)$$

siendo las funciones aproximadoras $\phi_{kj}(x)$ los polinomios básicos de Lagrange

$$\phi_{k1}(x) = \frac{x - x_{k+1}}{x_k - x_{k+1}} , \quad \phi_{k2}(x) = \frac{x - x_k}{x_{k+1} - x_k} .$$

Observar que se verifica $U_k(x_k) = u_{k1}$ y $U_k(x_{k+1}) = u_{k2}$ para todo $k = 1, \dots, N$.

24.2.4. Método variacional de Galerkin

El método de *Galerkin* de elementos finitos consiste en tomar las funciones peso $w_j(x) = \phi_{kj}(x)$ en la formulación débil (24.11) con $j = 1, 2$, teniendo en cuenta (24.12). De esta forma se tiene

$$\begin{aligned} 0 &= \int_{x_A}^{x_B} \left[a_1(x) \frac{d\phi_{k1}}{dx} \left(\sum_{j=1}^2 u_{kj} \frac{d\phi_{kj}}{dx} \right) + a_0(x) \phi_{k1} \left(\sum_{j=1}^2 u_{kj} \phi_{kj}(x) \right) - \right. \\ &\quad \left. \phi_{k1}(x) f(x) \right] dx - \sum_{j=1}^2 \phi_{k1}(x_{kj}) Q_{kj} , \\ 0 &= \int_{x_A}^{x_B} \left[a_1(x) \frac{d\phi_{k2}}{dx} \left(\sum_{j=1}^2 u_{kj} \frac{d\phi_{kj}}{dx} \right) + a_0(x) \phi_{k2} \left(\sum_{j=1}^2 u_{kj} \phi_{kj}(x) \right) - \right. \\ &\quad \left. \phi_{k2}(x) f(x) \right] dx - \sum_{j=1}^2 \phi_{k2}(x_{kj}) Q_{kj} , \end{aligned}$$

para cada elemento Ω_k . Aquí se han denotado por x_{kj} , con $j = 1, 2$, los nodos del elemento Ω_k y por $Q_{k1} = Q_A$, $Q_{k2} = Q_B$.

Teniendo en cuenta que los polinomios básicos de Lagrange $\phi_{ki}(x)$ verifican la propiedad $\phi_{ki}(x_{kj}) = \delta_{ij}$, siendo δ_{ij} la delta de Kronecker, es decir, $\delta_{ij} = 1$ si $i = j$ y $\delta_{ij} = 0$ para $i \neq j$, podemos escribir la ecuación i -ésima del método de Galerkin mediante un reordenamiento como

$$0 = \sum_{j=1}^2 \left[\int_{x_A}^{x_B} a_1(x) \frac{d\phi_{ki}}{dx} \frac{d\phi_{kj}}{dx} dx \right] u_{kj} + \sum_{j=1}^2 \left[\int_{x_A}^{x_B} a_0(x) \phi_{ki} \phi_{kj} dx \right] u_{kj} - \int_{x_A}^{x_B} a_0(x) f(x) \phi_{ki} dx - Q_{ki}, \quad i = 1, 2.$$

Este sistema de ecuaciones, llamado *ecuación del elemento* $\Omega_k = [x_A, x_B]$ se escribe en forma matricial para el vector incógnita $u^{[k]} := (u_{k1}, u_{k2})^t$ como

$$K^{[k]} u^{[k]} = F^{[k]} + Q^{[k]}, \quad (24.13)$$

donde la matriz de coeficientes $K^{[k]}$, llamada *matriz de rigidez*, viene dada por

$$K^{[k]} = (K_{ij}^{[k]}) \in \mathcal{M}_2(\mathbb{R}), \quad K_{ij}^{[k]} = K_{ij}^{[k,1]} + K_{ij}^{[k,0]}, \quad (24.14)$$

$$K_{ij}^{[k,1]} = \int_{x_A}^{x_B} a_1(x) \frac{d\phi_{ki}}{dx} \frac{d\phi_{kj}}{dx} dx, \quad (24.15)$$

$$K_{ij}^{[k,0]} = \int_{x_A}^{x_B} a_0(x) \phi_{ki} \phi_{kj} dx, \quad (24.16)$$

y el vector de términos independientes tiene dos contribuciones, una dada por el *vector fuente* $F^{[k]} = (F_1^{[k]}, F_2^{[k]})^t$ de componentes

$$F_i^{[k]} = \int_{x_A}^{x_B} f(x) \phi_{ki} dx, \quad (24.17)$$

y la otra por el vector $Q^{[k]} = (Q_{k1}, Q_{k2})^t$. Observar que la matriz de rigidez $K^{[k]}$ es simétrica puesto que $K_{ij}^{[k]} = K_{ji}^{[k]}$.

Se tiene en definitiva, para cada elemento Ω_k , dos ecuaciones con cuatro incógnitas que son u_{k1} , u_{k2} y también Q_{k1} y Q_{k2} . Por lo tanto todavía nos faltan dos ecuaciones más para cada elemento Ω_k que se obtendrán posteriormente cuando realizemos el encadenamiento de los elementos finitos (recordar la ecuación (24.7)) y usemos las condiciones de contorno del problema.

24.2.5. Notas computacionales

Veamos como calcular de forma numérica las componentes $K_{ij}^{[k]}$ y $F_i^{[k]}$ de la matriz de rigidez y del vector fuente. Para ello realizaremos un cambio de variable lineal que pase el intervalo $x \in \Omega_k = [x_A, x_B]$ al intervalo $\xi \in [-1, 1]$, es decir,

$$x = \eta_k(\xi) := \frac{h_k}{2}(\xi + 1) + x_A, \quad (24.18)$$

siendo $h_k = x_B - x_A$ la longitud del elemento Ω_k . Con este cambio de variable, los polinomios básicos de Lagrange $\phi_{k1}(x)$ y $\phi_{k2}(x)$ se transforman en

$$\Phi_1(\xi) = \frac{1}{2}(1 - \xi) , \quad \Phi_2(\xi) = \frac{1}{2}(1 + \xi) . \quad (24.19)$$

En estas nuevas coordenadas, los elementos de la matriz de rigidez (24.16) adoptan la forma

$$K_{ij}^{[k,1]} = \frac{2}{h_k} \int_{-1}^1 a_1(\eta_k(\xi)) \frac{d\Phi_i}{d\xi} \frac{d\Phi_j}{d\xi} d\xi , \quad (24.20)$$

$$K_{ij}^{[k,0]} = \frac{h_k}{2} \int_{-1}^1 a_0(\eta_k(\xi)) \Phi_i(\xi) \Phi_j(\xi) d\xi , \quad (24.21)$$

y las componentes del vector fuente

$$F_i^{[k]} = \frac{h_k}{2} \int_{-1}^1 f(\eta_k(\xi)) \Phi_i(\xi) d\xi . \quad (24.22)$$

Para evaluar de forma numérica las integrales (24.20), (24.21) y (24.22) se utilizará, por ejemplo, la cuadratura de Gauss-Legendre con dos puntos

$$\int_{-1}^1 g(\xi) d\xi \approx g(-\sqrt{3}/3) + g(\sqrt{3}/3) . \quad (24.23)$$

24.2.6. Encadenamiento de los elementos

Una vez se tienen las ecuaciones para cada elemento, ver (24.13), para aproximar la solución exacta $u(x)$ en el dominio $[0, L]$ tal y como indica (24.7), se impondrá, en primer lugar, la continuidad de la función $U(x)$ definida a trozos (24.13). Para ello es necesario previamente una *renumeración global de los nodos*. Si numeramos de manera consecutiva a los nodos de todo el dominio $[0, L]$ comenzando por el nodo 1 que estará localizado en $x = 0$, y denotando por U_i , desde $i = 1, \dots, N + 1$ al valor de la función $U(x)$ en el nodo i -ésimo, es decir, $u_{ki} = U_{(n-1)(k-1)+i}$ en el supuesto de haber tomado n nodos en cada elemento Ω_k . Por supuesto, en nuestro caso $n = 2$ de manera que $u_{ki} = U_{k+i-1}$. De esta forma, la continuidad de la solución aproximada $U(x)$ se expresa como

$$\begin{aligned} U_1 &= u_{11} , \\ U_i &= u_{i-1,2} = u_{i1} , \text{ para } i = 2, \dots, N , \\ U_{N+1} &= u_{N2} . \end{aligned} \quad (24.24)$$

En segundo lugar se impondrá un balance de flujos de la forma

$$Q_{k2} + Q_{k+1,1} = \begin{cases} 0 , \\ Q_k , \end{cases}$$

dependiendo de si no hay fuente externa aplicada en el nodo que conecta el elemento Ω_k y Ω_{k+1} o por el contrario hay una fuente de valor Q_k . Observar

que, en el caso de que $Q_{k2} + Q_{k+1,1} = 0$, esta condición implica en dicho nodo de conexión que la derivada de la solución exacta $du(x)/dx$ es continua debido a la definición de Q_{ki} , ver (24.10).

La forma efectiva de realizar el balance de flujo consiste en sumar la última ecuación del elemento Ω_k con la primera del elemento Ω_{k+1} . Se tiene pues que, el sistema con numeración local

$$\begin{aligned} \Omega_1 & \begin{cases} K_{11}^{[1]}u_{11} + K_{12}^{[1]}u_{12} &= F_1^{[1]} + Q_{11} , \\ K_{21}^{[1]}u_{11} + K_{22}^{[1]}u_{12} &= F_2^{[1]} + Q_{12} , \end{cases} \\ \Omega_2 & \begin{cases} K_{11}^{[2]}u_{21} + K_{12}^{[2]}u_{22} &= F_1^{[2]} + Q_{21} , \\ K_{21}^{[2]}u_{21} + K_{22}^{[2]}u_{22} &= F_2^{[2]} + Q_{22} , \end{cases} \\ \Omega_3 & \begin{cases} K_{11}^{[3]}u_{31} + K_{12}^{[3]}u_{32} &= F_1^{[3]} + Q_{31} , \\ K_{21}^{[3]}u_{31} + K_{22}^{[3]}u_{32} &= F_2^{[3]} + Q_{32} , \end{cases} \\ & \vdots \\ \Omega_N & \begin{cases} K_{11}^{[N]}u_{N1} + K_{12}^{[N]}u_{N2} &= F_1^{[N]} + Q_{N1} , \\ K_{21}^{[N]}u_{N1} + K_{22}^{[N]}u_{N2} &= F_2^{[N]} + Q_{N2} , \end{cases} \end{aligned}$$

pasa al sistema con numeración global de nodos siguiente

$$\begin{aligned} & K_{11}^{[1]}U_1 + K_{12}^{[1]}U_2 &= F_1^{[1]} + Q_{11} , \\ K_{21}^{[1]}U_1 + (K_{22}^{[1]} + K_{11}^{[2]})U_2 + K_{12}^{[2]}U_3 &= F_2^{[1]} + F_1^{[2]} + Q_{12} + Q_{21} , \\ K_{21}^{[2]}U_2 + (K_{22}^{[2]} + K_{11}^{[3]})U_3 + K_{12}^{[3]}U_4 &= F_2^{[2]} + F_1^{[3]} + Q_{22} + Q_{31} , \\ & \vdots \\ K_{21}^{[N-1]}U_{N-1} + (K_{22}^{[N-1]} + K_{11}^{[N]})U_N + K_{12}^{[N]}U_{N+1} &= F_2^{[N-1]} + F_1^{[N]} + Q_{N-1,2} + Q_{N1} , \\ & K_{21}^{[N]}U_N + K_{22}^{[N]}U_{N+1} &= F_2^{[N]} + Q_{N2} . \end{aligned}$$

Estas ecuaciones lineales adoptan la forma de un sistema matricial tridiagonal simétrico que escribiremos como

$$KU = F , \quad (24.25)$$

y llamaremos *sistema encadenado*. Aquí, la matriz de coeficientes K , el vector incógnita U y el vector de términos independientes F son los dados de forma explícita en las siguientes ecuaciones.

$$K = \begin{pmatrix} K_{11}^{[1]} & K_{12}^{[1]} & & & \\ K_{21}^{[1]} & K_{22}^{[1]} + K_{11}^{[2]} & K_{12}^{[2]} & & \\ & K_{21}^{[2]} & K_{22}^{[2]} + K_{11}^{[3]} & K_{12}^{[3]} & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & \ddots \\ & & & & K_{21}^{[N-1]} & K_{22}^{[N-1]} + K_{11}^{[N]} & K_{12}^{[N]} \\ & & & & & K_{21}^{[N]} & K_{22}^{[N]} \end{pmatrix} ,$$

$$\begin{aligned}
 U &= \begin{pmatrix} U_1 \\ U_2 \\ U_3 \\ \vdots \\ \vdots \\ U_N \\ U_{N+1} \end{pmatrix} \\
 F &= \begin{pmatrix} F_1^{[1]} \\ F_2^{[1]} + F_1^{[2]} \\ F_2^{[2]} + F_1^{[3]} \\ \vdots \\ \vdots \\ F_2^{[N-1]} + F_1^{[N]} \\ F_2^{[N]} \end{pmatrix} + \begin{pmatrix} Q_{11} \\ Q_{12} + Q_{21} \\ Q_{22} + Q_{31} \\ \vdots \\ \vdots \\ Q_{N-1,2} + Q_{N1} \\ Q_{N2} \end{pmatrix}.
 \end{aligned}$$

24.2.7. La solución del sistema encadenado

En esta sección se verá cómo introducir los diferentes tipos de condiciones de contorno (24.3), (24.4), (24.5) o (24.6) así como técnicas prácticas que simplifiquen el tratamiento numérico para la resolución del sistema encadenado (24.25). En concreto se expondrá el *método de unos en la diagonal*. De esta forma, si las condiciones de contorno son, por ejemplo, $u(0) = \bar{U}_0$, $u(L) = \bar{U}_L$, entonces la primera y última fila de la matriz de coeficientes y el primero y último de los términos independientes del sistema encadenado (24.25) se modifican de forma que se verifique $U_1 = \bar{U}_0$, $U_{N+1} = \bar{U}_L$.

24.3. Desplazamientos Longitudinales Elásticos

Definamos las siguientes magnitudes:

- $u(x, t)$, el *desplazamiento longitudinal* respecto de la posición inicial de equilibrio del punto de coordenada x para el instante de tiempo t . UNIDADES: m .
- $\sigma(x, t)$, *tensión* del punto de coordenada x para el instante de tiempo t . UNIDADES: N/m^2 .
- $F(x, t)$, *fuerza interna* que actúa en el punto x para el tiempo t . Por ejemplo la gravedad. UNIDADES: N .
- $f(x, t) = \partial F / \partial x$. Son las fuerzas internas por unidad de longitud. UNIDADES: N/m .
- $E(x)$, *módulo de Young* del material. UNIDADES: N/m^2 .

- $A(x)$, área de la sección transversal del material. UNIDADES: m^2 .
- $\rho(x)$, densidad del material. UNIDADES: Kg/m^3 .

Teniendo en cuenta las fuerzas que actúan sobre el elemento de longitud dx y aplicando la segunda ley de Newton se tiene, para el instante de tiempo t ,

$$\sigma(x+dx, t)A(x+dx) - \sigma(x, t)A(x, t) + F(x+dx, t) - F(x, t) = \rho(x)A(x)dx \frac{\partial^2 u}{\partial t^2}.$$

Dividiendo por dx y pasando al límite cuando dx tiende a cero se tiene

$$\frac{\partial}{\partial x}(\sigma A) + f = \rho A \frac{\partial^2 u}{\partial t^2}.$$

Finalmente, teniendo en cuenta la *ley de Hooke*

$$\sigma = E \frac{\partial u}{\partial x},$$

se obtiene

$$\rho A \frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x} \left(EA \frac{\partial u}{\partial x} \right) = f$$

Para el caso estacionario, es decir, con independencia del tiempo, la anterior ecuación en derivadas parciales se convierte en la ecuación diferencial ordinaria de segundo orden autoadjunta

$$-\frac{d}{dx} \left(E(x)A(x) \frac{du}{dx} \right) = f(x), \quad (24.26)$$

que es un caso particular de la ecuación (24.2).

Problema 24.1 (Deformación de una columna de hormigón)

Consideremos una columna de hormigón de altura $H = 3m$ en forma de prisma rectangular truncado. Tomando un sistema de referencia con $x = 0$ en la base de la columna, la sección transversal $A(x)$ de la columna es un rectángulo donde uno de los lados siempre vale 0,5m mientras que el otro varía desde 1,5m en la base hasta 0,5m en la altura. Bajo estas condiciones es fácil demostrar que $A(x) = 3/4 - x/6$ metros cuadrados. La columna soporta en su parte superior una densidad de carga continua de $24 \times 10^3 N/m^2$ mientras que a una altura de dos metros tiene dos anclajes cargando $2 \times 10^3 N$ cada uno de ellos.

El hormigón tiene peso específico $w = 25 \times 10^3 N/m^3$ y módulo de elasticidad $E = 28 \times 10^9 N/m^2$. Las fuerzas internas $f(x)$ en el hormigón son debidas a su propio peso, de manera que $f(x) = -wA(x)$.

Queremos hallar los desplazamientos verticales que experimentan los puntos de la columna. Para conseguirlo usaremos 30 elementos de longitud constante $h = 0,1m$ y una numeración global consecutiva de los nodos comenzando en el nodo 1 en la base de la columna $x = 0$ y finalizando en el nodo 31 a la altura máxima de la columna $x = H$. Mostrar la gráfica de los desplazamientos U en función de la posición x .

24.4. Fluidos Viscosos en Canalización Cilíndrica

Consideremos un fluido viscoso laminar en una tubería cilíndrica de radio R . Debido a la geometría del problema, tomaremos coordenadas cilíndricas (r, θ, z) , siendo z el eje de simetría de la tubería. Supondremos que la velocidad v de las partículas de fluido sólo depende de la coordenada radial, es decir $v(r)$. Se puede demostrar en este caso que el perfil de velocidades verifica la ecuación diferencial

$$\frac{1}{r} \frac{d}{dr} \left(r \mu \frac{dv}{dr} \right) = \frac{dp}{dz}, \quad (24.27)$$

que es un caso particular de la ecuación (24.2). En esta ecuación diferencial, μ es la viscosidad dinámica del fluido cuyas unidades son Kg/ms y p es la presión en el interior del fluido con unidades N/m^2 . Supondremos que p depende únicamente de z y, por consideraciones físicas $dp/dz < 0$. Las condiciones de contorno asociadas a la ecuación (24.27) son $v(R) = 0$ y $\mu dv(0)/dr = 0$. La primera condición tiene su origen en que, debido a la viscosidad, las moléculas del fluido se adhieren a la pared de la tubería. La segunda condición es debida a la imposición de la anulación de tensiones de cizallamiento.

Observemos que, si la viscosidad μ es una constante independiente de r , entonces es fácil integrar respecto de r la ecuación (24.27) y comprobar que v depende de r a través de un polinomio de segundo grado, es el llamado perfil parabólico de velocidades.

Problema 24.2 (Perfil de velocidades con viscosidad variable)

Consideremos una solución de glicerina que se desplaza por una tubería de radio $R = 8cm$ y caída de presión de $100 N/m^2$, es decir, $dp/dz = -100 N/m^2$. Se sabe que la viscosidad de la solución de glicerina depende de la temperatura T . Además, la tubería está sometida a un gradiente radial de temperaturas según la tabulación siguiente.

$r(cm)$	0	2	4	6	8
$T(^{\circ}C)$	30	35	40	45	50
$\mu(Kg/ms)$	0.40	0.28	0.20	0.12	0.10

A partir de la primera y tercera fila de dicha tabulación podemos reconstruir la función $\mu(r)$ suponiendo que se comporta de forma lineal entre dos datos consecutivos.

Queremos hallar el perfil de velocidades $v(r)$ de la solución de glicerina. Para conseguirlo usaremos 16 elementos de longitud constante $h = 0,5m$ y una numeración global consecutiva de los nodos comenzando en el nodo 1 en el centro de la tubería ($r = 0$) y finalizando en el nodo 17 para $r = R = 8cm$. Mostrar la gráfica de la velocidad V en función de la posición r .

Capítulo 25

Problemas Variacionales con Fronteras Fijas

25.1. Introducción

Sea E un espacio de funciones $y(x)$ de clase $C^1[a, b]$ tomando valores fijos en los extremos, es decir, $y(a) = y_a$, $y(b) = y_b$. Consideremos el funcional $J : E \rightarrow \mathbb{R}$, definido de la forma

$$J(y) = \int_a^b F(x, y, y') \, dx . \quad (25.1)$$

El problema que nos planteamos consiste en hallar una función $y(x) \in E$ que sea un extremal (máximo o mínimo) del funcional $J(y)$ definido en (25.1). Por ejemplo, la función $y^*(x)$ minimiza al funcional J si se verifica $J(y^*) \leq J(y)$ para todo $y \in E$. Por supuesto, el principal problema con el que nos encontramos y que lo diferencia radicalmente de las técnicas utilizadas en los problemas clásicos de hallar máximos y mínimos de funciones de varias variables es el hecho de que la dimensión del espacio de funciones E es infinita.

25.2. Un método numérico

Sea $n \in \mathbb{N}$ y consideremos una partición equiespaciada del intervalo $[a, b]$ con longitud de paso h , es decir,

$$x_i = a + ih , \quad h = \frac{b-a}{n+1} , \quad i = 0, 1, \dots, n+1 .$$

Consideremos las aproximaciones $y(x_i) \approx y_i$, con $i = 1, \dots, n$, con $y(x_0) = y(a) = y_a$, $y(x_{n+1}) = y(b) = y_b$.

Realizamos la siguiente aproximación $F(x, y, y') \approx F_n(x, y_1, \dots, y_n)$ mediante una interpolación constante a trozos:

$$F_n(x, y_1, \dots, y_n) = F\left(x_i, y_i, \frac{y_i - y_{i-1}}{h}\right), \text{ si } x \in (x_i, x_{i+1}), \quad i = 1, \dots, n.$$

A continuación, aproximamos $J(y) \approx J_n(y_1, \dots, y_n)$, siendo la función $J_n : \mathbb{R}^n \rightarrow \mathbb{R}$ definida de la forma

$$J_n(y_1, \dots, y_n) = \int_a^b F_n(x, y_1, \dots, y_n) dx = \sum_{j=1}^{n+1} F\left(x_j, y_j, \frac{y_j - y_{j-1}}{h}\right) h.$$

Una condición necesaria para que la función J_n tenga un extremo en el punto crítico $y^* = (y_1^*, \dots, y_n^*)$ es $\nabla J_n(y^*) = 0$, siendo el operador gradiente $\nabla = (\partial/\partial y_1, \dots, \partial/\partial y_n)$.

Notemos que y_i sólo aparece 3 veces de manera explícita en la expresión de J_n . Más concretamente

$$J_n(y_1, \dots, y_n) = h \left[\dots + F\left(x_i, y_i, \frac{y_i - y_{i-1}}{h}\right) + F\left(x_{i+1}, y_{i+1}, \frac{y_{i+1} - y_i}{h}\right) \dots \right].$$

Entonces

$$\begin{aligned} \frac{\partial J_n}{\partial y_i} &= \frac{\partial F}{\partial y}\left(x_i, y_i, \frac{y_i - y_{i-1}}{h}\right) h + \frac{\partial F}{\partial y'}\left(x_i, y_i, \frac{y_i - y_{i-1}}{h}\right) \\ &\quad - \frac{\partial F}{\partial y'}\left(x_{i+1}, y_{i+1}, \frac{y_{i+1} - y_i}{h}\right) = 0, \quad i = 1, \dots, n. \end{aligned} \quad (25.2)$$

De forma más compacta, si definimos las funciones $f_i(y_1, \dots, y_n) := \partial J_n / \partial y_i$ según (25.2), la condición $\nabla J_n = 0$ es equivalente al sistema de n ecuaciones no lineal con n incógnitas $f(y_1, \dots, y_n) = 0$, siendo la función $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida de la forma $f = (f_1, \dots, f_n)$.

Buscaremos la solución del sistema no lineal mediante el método de Newton, es decir,

$$y^j = y^{j-1} - [Df(y^{j-1})]^{-1} f(y^{j-1}), \quad j = 1, 2, \dots,$$

donde la condición inicial $y^0 = (y_1^0, \dots, y_n^0) \in \mathbb{R}^n$ debe ser dada. Observemos que, en cada iteración, se ha de invertir la matriz Jacobiana $Df(y^{j-1})$. Esto es sencillo puesto que en nuestro caso dicha matriz Df es la matriz Hessiana de la función J_n y es, por lo tanto, Df es una matriz simétrica, es decir,

$$\frac{\partial f_i}{\partial y_j} = \frac{\partial f_j}{\partial y_i}.$$

Por otra parte, debido a (25.2), se tiene

$$\frac{\partial f_i}{\partial y_{i-1}} = \frac{\partial^2 J_n}{\partial y_i \partial y_{i-1}}$$

$$= -\frac{\partial^2 F}{\partial y \partial y'} \left(x_i, y_i, \frac{y_i - y_{i-1}}{h} \right) - \frac{\partial^2 F}{\partial y'^2} \left(x_i, y_i, \frac{y_i - y_{i-1}}{h} \right) \frac{1}{h} \quad (25.3)$$

$$\begin{aligned} \frac{\partial f_i}{\partial y_i} &= \frac{\partial^2 J_n}{\partial y_i^2} = \frac{\partial^2 F}{\partial y^2} \left(x_i, y_i, \frac{y_i - y_{i-1}}{h} \right) h + 2 \frac{\partial^2 F}{\partial y \partial y'} \left(x_i, y_i, \frac{y_i - y_{i-1}}{h} \right) \\ &\quad + \frac{\partial^2 F}{\partial y'^2} \left(x_i, y_i, \frac{y_i - y_{i-1}}{h} \right) \frac{1}{h} + \frac{\partial^2 F}{\partial y'^2} \left(x_{i+1}, y_{i+1}, \frac{y_{i+1} - y_i}{h} \right) \frac{1}{h} \end{aligned} \quad (25.4)$$

$$\begin{aligned} \frac{\partial f_i}{\partial y_{i+1}} &= \frac{\partial^2 J_n}{\partial y_i \partial y_{i+1}} = -\frac{\partial^2 F}{\partial y \partial y'} \left(x_{i+1}, y_{i+1}, \frac{y_{i+1} - y_i}{h} \right) \\ &\quad - \frac{\partial^2 F}{\partial y'^2} \left(x_{i+1}, y_{i+1}, \frac{y_{i+1} - y_i}{h} \right) \frac{1}{h}, \end{aligned} \quad (25.5)$$

siendo el resto de elementos de la matriz nulos, de modo que la matriz es además tridiagonal.

25.3. El problema de la braquistocrona

El problema de la *braquistocrona* o curva de descenso más rápida, es propuesto y resuelto por el eminente matemático de origen suizo Johann Bernoulli en 1696, aunque también dieron soluciones otros contemporáneos suyos como Jacob Bernoulli, Leibniz y Newton. Dicho problema consistía en determinar cuál, de entre todas las posibles trayectorias, era la que llevaba a una partícula sin rozamiento en el menor tiempo posible, desde un punto A hasta otro punto B en el mismo plano vertical, sujeta sólo a la acción de la gravedad. Para resolver este problema, se deben de considerar todas las infinitas curvas que pasan por A y B , de modo que, a cada una de ellas se le asigna un tiempo (el invertido por el punto material en descender desde A hasta B) y nos hemos de quedar con la curva con menor tiempo asociado.

Veamos cómo formularlo matemáticamente. Consideremos el plano vertical de coordenadas cartesianas x e y que contiene a los puntos A y B . Obviamente, la curva solución debe estar contenida en dicho plano. Tomemos el punto A como el origen de coordenadas, el eje x de abscisas con sentido positivo hacia donde apunta la gravedad y, sea $B = (x_1, y_1)$ con $x_1 > 0$ e $y_1 \geq 0$. Consideremos una curva en dicho plano arbitraria y descrita por la gráfica $y = y(x)$ con $0 \leq x \leq x_1$. Supondremos que la curva es suave, es decir, es de clase C^1 en su dominio. Además impondremos que los puntos inicial y final de dicha curva sean A y B respectivamente, es decir, $y(0) = 0$ e $y(x_1) = y_1$.

La masa m , en su descenso, conserva su energía mecánica puesto que no hay rozamiento. Supondremos que parte con velocidad inicial nula del punto A . Entonces, la energía mecánica en A vale $E_A = mgh_A$, siendo $E_A > 0$ y h_A la altura a la que se encuentra el punto A . Conservando la energía mecánica en cualquier otro punto de la trayectoria de la partícula se tiene

$$\frac{1}{2}mv^2 + mgh = mgh_A,$$

de modo que $v^2 = 2g(h_A - h)$. Con el sistema de referencia tomado está claro que $x = h_A - h$, de modo que, la ecuación diferencial que rige el movimiento de la masa m es

$$v := \frac{ds}{dt} = \sqrt{2gx} ,$$

siendo v el módulo de la velocidad que tiene la partícula cuando se encuentra en un punto con coordenada $(x, y(x))$ y ds el elemento de longitud de la trayectoria. Sabemos que, $ds = \sqrt{1 + y'(x)^2} dx$, de modo que se obtiene

$$dt = \frac{\sqrt{1 + y'(x)^2} dx}{\sqrt{2gx}} .$$

Integrado esta expresión se obtiene el tiempo total T que tarda la masa m en recorrer la curva $y = y(x)$ desde el punto A hasta el punto B . En concreto se tiene que

$$T = J(y) = \frac{1}{\sqrt{2g}} \int_0^{x_1} \sqrt{\frac{1 + y'(x)^2}{x}} dx .$$

Hallar la braquistocrona consiste en obtener, de entre todas las funciones $y(x)$ definidas para $0 \leq x \leq x_1$, con $y(0) = 0$ e $y(x_1) = y_1$, aquella que corresponda al menor valor del funcional $J(y)$.

Cabe destacar aquí que, la solución de este problema viene dado por la famosa curva llamada *cicloide*, cuya ecuación paramétrica es

$$x(t) = \frac{1}{k^2}(1 - \cos t) , \quad y(t) = \frac{1}{k^2}(t - \sin t) ,$$

siendo k una constante adecuada, ver por ejemplo [9].

Problema 25.1 Utilizar el algoritmo numérico de la sección anterior para aproximar la curva de descenso más rápida (braquistocrona) que pasa por los puntos de coordenadas $(x_0, y_0) = (0, 0)$ y $(x_1, y_1) = (3, 3)$.

- *Datos de entrada:* Los extremos a y b del intervalo de abscisas; el número de nodos n ; un criterio de parada en la iteración de Newton, por ejemplo, n_{\max} el número máximo de iteraciones; las condiciones de frontera $y(a) = y_a$ e $y(b) = y_b$.
- *Tomar,* para la iteración inicial del método de Newton, una interpolación lineal entre los extremos, es decir, la recta que pasa por (a, y_a) y (b, y_b) .
- *Salida:* La tabulación (x_i, y_i) con x_i abscisas equiespaciadas e $y(x_i) \approx y_i$.
- *Finalmente* comparar el resultado numérico aproximado con la solución exacta mediante superposición de gráficas.

Capítulo 26

Introducción a MATHEMATICA

26.1. Introducción

En este apéndice se pretende dar unas breves nociones sobre algunos de los comandos utilizados por el programa MATHEMATICA¹.

Además, la ayuda en MATHEMATICA para conocer información detallada sobre un cierto comando puede realizarse escribiendo ?? delante del nombre del comando o bien acudiendo al menú de ayuda.

26.2. Estructuras de datos: Listas

Una lista es para MATHEMATICA una estructura de datos cuyos elementos aparecen ordenados entre llaves y separados por comas. En realidad, los vectores y matrices son casos particulares de listas. Por ejemplo, el vector x y la matriz A siguientes

$$x = (1, 2, 3) \in \mathbb{R}^3, \quad A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}),$$

son representados por MATHEMATICA a través de las listas

$$x = \{1, 2, 3\}, \quad A = \{\{1, 2\}, \{3, 4\}\}.$$

Cuando un comando de MATHEMATICA es aplicado a una lista, este devuelve otra lista de las mismas dimensiones donde ha aplicado dicho comando a todos los elementos de la lista. Así, por ejemplo, si $y = \{x, \text{Sin}[x], \text{Exp}[x]\}$,

¹Por supuesto, una guía completa y mucho más detallada sobre las posibilidades de MATHEMATICA se encuentra en el manual del programa contenido en la página web <http://documents.wolfram.com/v5/TheMathematicaBook/index.html>

entonces puedo integrar respecto de la variable x a todos los elementos de la lista y de la forma $Integrate[y, x]$. El resultado que devuelve MATHEMATICA es $\{x^2/2, -Cos[x], Exp[x]\}$.

Si los elementos de una lista se generan a partir de una cierta función f , entonces, se puede utilizar el comando $Table[]$ de la forma $Table[f, \{i, i_{min}, i_{max}, d\}]$. Por ejemplo $Table[i \wedge 2, \{i, 0, 6, 2\}]$ produce la lista $\{0, 4, 16, 36\}$. De igual modo, por ejemplo, la matriz de Hilbert de orden 3 se genera de la forma $Table[1/(i + j - 1), \{i, 1, 3, 1\}, \{j, 1, 3, 1\}]$.

En la programación con MATHEMATICA resulta muy útil poder obtener de forma individualizada los elementos de una cierta lista. En concreto, $z[[i]]$ y $w[[i, j]]$ devuelven los elementos i -ésimo de la lista z y el elemento que ocupa la posición fila i columna j de la lista w . De esta forma, si $z = \{a, a^2, a^5\}$, entonces $z[[1]]$ vale a , $z[[2]]$ vale a^2 y $z[[3]]$ vale a^5 . De forma análoga, si $w = \{\{a, b\}, \{c, d\}\}$ entonces $w[[1]]$ equivale a $\{a, b\}$ y $w[[2]]$ es $\{c, d\}$. Sin embargo $w[[1, 1]]$ es a , $w[[1, 2]]$ es b , $w[[2, 1]]$ es c y $w[[2, 2]]$ es d .

Existen diversos comandos de MATHEMATICA que nos permiten modificar, añadir o eliminar elementos de listas. Por ejemplo, $Prepend[x, e]$ y $Append[x, e]$ añaden el elemento e a la lista x como el primer elemento y el último respectivamente. Por ejemplo $Prepend[\{a, b, c\}, e]$ devuelve $\{e, a, b, c\}$ y $Append[\{a, b, c\}, e]$ devuelve $\{a, b, c, e\}$.

Otra función interesante aplicable a listas es el comando $Length[x]$, que da como resultado el número de elementos de la lista x . Así, $Length[\{a, b, c\}]$ vale 3, mientras que $Length[\{\{a, b, c\}, \{d, e, f\}\}]$ es 2.

Una forma de visualizar los elementos de una lista x en forma de una tabla es mediante el comando $TableForm[x]$. Por ejemplo, la salida que produce MATHEMATICA tras el comando $TableForm[\{\{1, 2\}, \{3, 4\}\}]$ es

1	2
3	4

26.3. Funciones

Si se desea definir una función real de variable real $f : \mathbb{R} \rightarrow \mathbb{R}$, por ejemplo,

$$f(x) = \frac{x^3 \exp(2x)}{\sin x},$$

la forma de realizarlo es

$$f[x_] := (x \wedge 3 * Exp[2 * x]) / Sin[x].$$

Para evaluar la función f en un punto, por ejemplo x_0 , simplemente se escribe $f[x_0]$.

Por supuesto también es posible definir una función real de varias variables reales $f : \mathbb{R}^m \rightarrow \mathbb{R}$. Por ejemplo, para definir la función

$$f(x, y) = x^3 - 3xy ,$$

escribiríamos la sentencia

$$f[x_, y_] := x \wedge 3 - 3 * x * y$$

de modo que para evaluarla en el punto (x_0, y_0) simplemente se escribe $f[x_0, y_0]$. El argumento de la función de varias variables también puede ser único, es decir, en forma de una lista. Así, la anterior función se hubiera definido como

$$f[\{x_, y_-\}] := x \wedge 3 - 3 * x * y$$

y su evaluación de la forma $f[\{x_0, y_0\}]$.

Finalmente, para definir funciones vectoriales reales de varias variables reales $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ como puede ser

$$f(x, y) = (x^3 y, \cos x, e^{xy}) ,$$

escribiríamos la sentencia

$$f[x_, y_] := \{x \wedge 3 * y, \text{Cos}[x], E \wedge (x * y)\} .$$

26.4. Representación Gráfica

En esta sección veremos algunas de las capacidades gráficas del programa MATHEMATICA.

Para dibujar la gráfica de la función de una variable $y = f(x)$ en el intervalo $a \leq x \leq b$ el comando a utilizar es $\text{Plot}[f, \{x, a, b\}]$.

El comando $\text{Plot}[]$ tiene una gran variedad de opciones que nos permiten, por ejemplo, dar un título a la gráfica, etiquetar los ejes coordenados o bien modificarlos, elegir el origen de coordenadas, dar colores a las gráficas y un largo etcétera que puede consultarse en el manual del programa.

Si se tienen diversas gráficas, es posible realizar una superposición de ellas con el comando $\text{Show}[\text{grafica}_1, \text{grafica}_2, \dots, \text{grafica}_n]$.

Una instrucción muy útil que nos permitirá, dada una nube de puntos, presentarlos gráficamente es el comando $\text{ListPlot}[]$. La forma de utilizarlo es la siguiente. Consideremos una conjunto de puntos en el plano $(x_i, y_i) \in \mathbb{R}^2$ con $i = 1, \dots, n$. Entonces introduciendo

$$\text{ListPlot}[\{\{x_1, y_1\}, \{x_2, y_2\}, \dots, \{x_n, y_n\}\}, \text{PlotStyle} \rightarrow \text{PointSize}[0.02]]$$

se dibujan los puntos en el plano. *PlotStyle* \rightarrow tiene diversas opciones (ver el manual del programa), nosotros hemos utilizado una opción que nos permite controlar el tamaño de los puntos. Si además se utiliza

$$\text{ListPlot}[\{\{x_1, y_1\}, \{x_2, y_2\}, \dots, \{x_n, y_n\}\}, \text{PlotJoined} \rightarrow \text{True}]$$

entonces los puntos representados serán unidos por una recta. También es posible crear gráficas con colores mediante la opción *PlotStyle* \rightarrow *Hue*[*m*], con $0 \leq m \leq 1$, de manera que a medida que aumenta *m* se pasa por rojo, amarillo, verde, azul, etc...

En muchas ocasiones la nube de puntos proviene de un fichero de datos llamado, por ejemplo, *puntos.dat* con la estructura siguiente

$$\begin{array}{cc} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{array}$$

Entonces, con la instrucción

$$\text{ReadList}["\text{puntos.dat}", \text{Number}, \text{RecorLists} \rightarrow \text{True}]$$

se cargará la nube de puntos desde el fichero hacia una lista de la forma

$$\{\{x_1, y_1\}, \{x_2, y_2\}, \dots, \{x_n, y_n\}\}.$$

Para las gráficas en 3D se dispone de comandos análogos a los presentados anteriormente en 2D. Así, para dibujar la gráfica de la función de dos variable $z = f(x, y)$ en el dominio rectangular $a \leq x \leq b$, $c \leq y \leq d$ el comando a utilizar es *Plot3D*[*f*[*x*, *y*], {*x*, *a*, *b*}, {*y*, *c*, *d*}].

Por otra parte, consideremos una conjunto de puntos en el \mathbb{R}^3 (x_i, y_i, z_i) con $i = 1, \dots, n$. Entonces introduciendo

$$\text{ListPlot3D}[\{\{x_1, y_1, z_1\}, \dots, \{x_n, y_n, z_n\}\}, \text{PlotStyle} \rightarrow \text{PointSize}[0.02]]$$

se dibujan los puntos en el espacio.

Los gráficos obtenidos se pueden exportar a otras aplicaciones. En algunos casos simplemente copiando y pegando desde el menú Edit. También se puede utilizar el comando

$$\text{Display}["\text{nombre.ext}", \text{grafico}, "\text{tipo}"]$$

de modo que el gráfico llamado *grafico* se graba en un fichero llamado *nombre.ext* que es de un cierto *tipo*. Los posibles *tipo* son: EPS, PDF, Illustrator, PICT, Metafile, TIFF, GIF, PSImage, XBitmap, PCL, MGF y PBM. Por ejemplo:

$$\begin{aligned} \text{dibujo} &= \text{Plot}[\text{Sin}[x], \{x, 0, \text{Pi}\}] \\ \text{Display}["\text{grafica.eps}", \text{dibujo}, "\text{EPS}"] \end{aligned}$$

guarda el dibujo en el fichero *grafica.eps* que está en formato *EPS*. A partir de la versión 6 de Mathematica, la orden `Display[]` está obsoleta y se utiliza la instrucción

Export["nombre.ext", grafico]

26.5. Matrices y Sistemas Lineales

Para visualizar los elementos de una matriz A en MATHEMATICA se usa el comando *MatrixForm*[x]. Por ejemplo, el comando *MatrixForm*[$\{\{1, 2\}, \{3, 4\}\}$] produce

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

La matriz identidad de orden $n \in \mathbb{N}$ se puede genera con el comando

IdentityMatrix[n]

Por ejemplo *IdentityMatrix*[2] devuelve $\{\{1, 0\}, \{0, 1\}\}$.

Si A y B son dos matrices del mismo orden, $A + B$, $A - B$ y $k * A$ representan la suma, resta y producto por un escalar k . Si A y B se pueden multiplicar, entonces $A.B$ es su matriz producto.

Sea $A \in \mathcal{M}_n(\mathbb{R})$. Su determinante es *Det*[A]. Si además A es inversible, su matriz inversa es *Inverse*[A].

Por otra parte, la traspuesta de la matriz A es *Transpose*[A].

Sea $A \in \mathcal{M}_n(\mathbb{R})$ y $b \in \mathbb{R}^n$. La solución del sistema lineal $Ax = b$ viene dada por el comando *LinearSolve*[A, b]. Por ejemplo, en MATHEMATICA, el comando *LinearSolve*[$\{\{1, 1\}, \{1, -1\}\}, \{2, 0\}$] produce por salida el vector $\{1, 1\}$. Recordemos que si el sistema lineal es compatible indeterminado, entonces *LinearSolve*[A, b] sólo da una solución particular del sistema.

26.6. El Lenguaje de Programación

El programa MATHEMATICA, además de los múltiples comandos que lleva incorporados y nos permiten resolver de manera exacta o bien aproximada una gran variedad de problemas, dispone también de un lenguaje de programación de alto nivel.

26.6.1. Tipos de números

MATHEMATICA puede trabajar de forma simbólica, es decir, con precisión infinita. En realidad la precisión depende únicamente de la memoria que dispone

el ordenador con el que se trabaja. Las operaciones con números enteros se realizan de forma exacta. Por ejemplo, el número factorial $215!$ da

```
463908967738245731201181971316266541335354
606980986542278147922285826879934715156756
753035191305059570813964428091287791299849
47682780134232434682183959939102665927776
439857317019738245619747364381030569622906
549543129198965152594576297260803180404842
596116964782060639065966089166712338842834
343045860606296811709288710361986642298681
053856193817242616463360000000000000000000
000000000000000000000000000000000000
```

Los números racionales son también tratados de forma exacta. Por ejemplo $(2/5 + 7/4) \wedge 2 - 1/2$ devuelve $1649/400$. Nótese la diferencia con $(2/5 + 7/4) \wedge 2 - 0,5$ que devuelve la aproximación $4,1225$.

Respecto a los números irracionales, debido a sus infinitas cifras decimales, o bien se trabaja con ellos de forma simbólica o bien son aproximados con un número finito de decimales. Por ejemplo $Sqrt[2]$ devuelve $\sqrt{2}$. Sin embargo $N[Sqrt[2]]$ da la aproximación numérica $1,41421$ o bien la aproximación con 20 dígitos $1,4142135623730950488$. Otro ejemplo es la aproximación del número π con 30 dígitos dada por $N[\pi, 30]$ que da

```
3,14159265358979323846264338328
```

Finalmente, recordar que MATHEMATICA también trabaja con números complejos. La representación de $a + ib \in \mathbb{C}$ con $a, b \in \mathbb{R}$ y $i = \sqrt{-1}$ es $a + I * b$. Por ejemplo $(I + 3) \wedge 2 - 7/4 + 2 * I$ devuelve $25/4 + 8i$.

26.6.2. Entradas de datos

En esta sección describiremos cómo entrar datos desde MATHEMATICA así como desde un archivo.

Por ejemplo, si se requiere que el usuario introduzca un número real y este sea asignado a la variable z , la sintaxis será $z = \text{Input}["\text{introduzca un numero real}"]$.

En ocasiones los datos se introducen desde un fichero de forma automatizada. En primer lugar, el operador `!!` muestra el contenido de un fichero. Por ejemplo `!!valores.dat` muestra el contenido del fichero *valores.dat*. Para leer un fichero de datos y cargarlo en una lista, ver la sección Estructura de datos: Listas.

26.6.3. Asignación de variables

No es necesario declarar variables al principio de un programa. Los nombres de las variables no pueden coincidir con los nombres de ninguna de las instrucciones de MATHEMATICA. Además, recordar que MATHEMATICA asigna las letras *E*, *I* y *Pi* a las constantes $e = 2.71828\dots$, $\sqrt{-1} \in \mathbb{C}$ y $\pi = 3.141592\dots$

La forma de asignar un cierto valor a una variable es *variable* = *valor*. Se pueden utilizar estructuras típicas de la programación como son *variable* = *variable* + *numero* con lo cual, el actual valor de *variable* es el que tenía más *numero*.

Si se utiliza un punto y coma al final de una asignación, o de cualquier otra instrucción, MATHEMATICA realiza la asignación pero no devuelve ningún resultado por pantalla. Esto es muy útil en la programación cuando no se quieren ver por pantalla los cálculos intermedios de un programa.

Para borrar el valor de una variable se utiliza la instrucción *Clear*[*variable*].

26.6.4. Iteraciones y bucles

La sentencia *For*[*variable* = *inicial*, *condicion*, *incremento*, *proceso*] repite el *proceso* que se especifique para los valores de *variable* desde el valor *inicial*, mientras que la *condicion* sea cierta, incrementando en *incremento* los valores de *variable* después de cada iteración. Por ejemplo

$$\text{For}[i = 1, i \leq 5, i++, \text{Print}[\text{Integrate}[x \wedge i, x]]]$$

imprime por pantalla las primitivas de las funciones x^i para $i = 1, 2, \dots, 5$.

La sentencia *Do*[*proceso*, {*variable*, *inicial*, *final*, *incremento*}] repite lo descrito en *proceso* para los valores de *variable* desde el valor *inicial* hasta el *final*, aumentando en *incremento* los valores de *variable* después de cada iteración. Por ejemplo

$$\text{Do}[\text{Print}[\text{Integrate}[x \wedge i, x]], \{i, 1, 5, 1\}]$$

imprime otra vez por pantalla las primitivas de las funciones x^i para $i = 1, 2, \dots, 5$.

La sentencia *While*[*condicion*, *proceso*] realiza *proceso* mientras sea cierta *condicion*. Por ejemplo

$$\begin{aligned} i &= 1; \\ \text{While}[i \leq 5, \text{Print}[\text{Integrate}[x \wedge i, x]]; i &= i + 1] \end{aligned}$$

vuelve a imprimir por pantalla las primitivas de las funciones x^i para $i = 1, 2, \dots, 5$.

26.6.5. Condicionales y operadores lógicos

La sentencia $If[condicion, accion]$ realiza lo siguiente: si *condicion* es cierta entonces realiza *accion*. Incluso puede ser más general, de modo que la sentencia

$$If[condicion, accion - verdadera, accion - falsa]$$

realiza *accion - verdadera* si *condicion* es cierta y en caso contrario realiza *accion - falsa*.

Por ejemplo, se puede definir la función partida

$$f(x) = \begin{cases} \ln x & \text{si } x > 0, \\ \sin x & \text{si } x \leq 0, \end{cases}$$

de la forma

$$f[x_] = If[x > 0, Log[x], Sin[x]]$$

En muchas ocasiones, es necesario incluir en las sentencias condicionales anteriores algún operador lógico. Estos operadores son

Operador	Significado
!	NO
&&	Y
	O

26.6.6. Sentencias de control

Existen unas sentencias de control que permiten o bien detener o bien controlar el desarrollo de un cierto programa. Por ejemplo $Break[]$ abandona el bloque que se está realizando en ese momento. $Return[expresion]$ abandona el proceso que se realiza y devuelve el valor *expresion*.

26.7. Cálculo Numérico

26.7.1. Errores

Debido a errores de redondeo, se define el *epsilon de la máquina* como el mayor número positivo ϵ tal que $1 + x = 1$ para todo $x \in (0, \epsilon)$. Si intentamos averiguar el menor número $n \in \mathbb{N}$ tal que $\epsilon \geq 10^{-n}$ con un programa escrito en el lenguaje de MATHEMATICA podríamos intentar lo siguiente.

```
For[n = 1, n <= 1000, n ++,
  If[1 + 10^-n == 1, Print["El valor de n es", n]; Break[]]
]
```

Se comprueba que, con el anterior programa, no se produce ninguna salida. Pero esto no significa que $\epsilon < 10^{-1000}$. Lo que ha ocurrido es que MATHEMATICA ha

operado de modo simbólico, es decir exacto. Se puede forzar a un cálculo con decimales, por ejemplo, escribiendo un punto en la suma efectuada de la forma

```
For[n = 1, n <= 1000, n ++,
If[1. + 10 ^ (-n) == 1, Print["El valor de n es", n]; Break[]]
]
```

Ahora, la salida del programa es

El valor de n es 14

Un sencillo ejercicio donde se muestra el efecto debido al epsilon de la máquina es el siguiente. Se sabe que

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = 2,7182818.....$$

Vamos a ir almacenando en una lista *calculo* las aproximaciones al número *e* obtenidas a medida que se aumenta el número *n*.

```
calculo = {};
For[n = 1, n <= 20, n ++, AppendTo[calculo, (1. + 1/10 ^ n) ^ n];
]
calculo
```

Se obtiene la siguiente lista de salida.

```
{2.59374, 2.70481, 2.71692, 2.71815, 2.71827,
2.71828, 2.71828, 2.71828, 2.71828,
2.71828, 2.71828, 2.71852, 2.71611, 2.71611,
3.03504, 1., 1., 1., 1., 1.}
```

Se observa como a medida que *n* aumenta, las aproximaciones obtenidas del número *e* son mejores hasta que llega un momento, en que van empeorando, llegando a estabilizarse en 1.

Siempre que se restan magnitudes muy próximas, se produce un *error por cancelación* con la consiguiente pérdida de cifras significativas. Si además esta resta es multiplicada por un número grande, el error se verá aumentado. Por ejemplo, la expresión $x(\sqrt{x+1} - \sqrt{x})$ para $x \geq 0$ es equivalente, multiplicando y dividiendo por el conjugado, a la expresión

$$\frac{x}{\sqrt{x+1} + \sqrt{x}},$$

de donde se ve que, para valores grandes de *x* es aproximadamente igual a $\sqrt{x}/2$ puesto que $\sqrt{x+1} \approx \sqrt{x}$. Sin embargo, la primera expresión es prácticamente

cero debida a errores de redondeo en la resta entre paréntesis. Veamos este efecto de los errores por cancelación con el siguiente programa.

```

calculo = {};
For[n = 1, n <= 20, n++,
mal = 10 ^ n * (Sqrt[10 ^ n + 1.] - Sqrt[10 ^ n]);
bien = 10 ^ n / (Sqrt[10 ^ n + 1.] + Sqrt[10 ^ n]);
AppendTo[calculo, {mal, bien}];
]
TableForm[calculo, TableHeadings -> {None, {"MAL", "BIEN"}}]

```

La salida del programa es la siguiente.

<i>MAL</i>	<i>BIEN</i>
1.54347	1.54347
4.98756	4.98756
15.8074	15.8074
49.9988	49.9988
158.113	158.113
⋮	⋮
5.02914 10 ⁶	5. 10 ⁶
1.49012 10 ⁷	1.58114 10 ⁷
0.	5. 10 ⁷
0.	1.58114 10 ⁸
⋮	⋮
0.	5. 10 ⁹

26.7.2. Polinomio interpolador

Con MATHEMATICA es posible hallar el polinomio $P_n(x)$ de grado menor o igual que n que interpola a una nube de puntos (x_i, y_i) para $i = 0, 1, \dots, n$. La sintaxis a utilizar es

$$\text{InterpolatingPolynomial}[\{\{x_0, y_0\}, \{x_1, y_1\}, \dots, \{x_n, y_n\}\}, x] .$$

Observar que el primer argumento es una lista que contiene las coordenadas de la nube de puntos y el segundo argumento es el nombre de la variable del polinomio interpolador. Por ejemplo, la recta en el plano que pasa por los puntos $(0, 1)$ y $(3, 7)$ se puede obtener de la forma $\text{InterpolatingPolynomial}[\{\{0, 1\}, \{3, 7\}\}, x]$, siendo la solución $1 + 2x$.

26.7.3. Aproximación lineal por mínimos cuadrados

MATHEMATICA tiene la posibilidad de hallar la función $f^*(x) = \sum_{i=1}^m c_i \varphi_i(x)$ que mejor aproxima a la nube de puntos (x_i, y_i) para $i = 0, 1, \dots, n$. Aquí, $c_i \in \mathbb{R}$ y $\varphi_i(x)$ son funciones dadas linealmente independientes.

La sintaxis a utilizar es

$$Fit[\{\{x_1, y_1\}, \dots, \{x_n, y_n\}\}, \{\varphi_1[x], \dots, \varphi_m[x]\}, x] .$$

Observar que el primer argumento es una lista que contiene las coordenadas de la nube de puntos, el segundo argumento es una lista con las funciones base y el tercero el nombre de la variable. Por ejemplo, la recta de regresión en el plano que pasa por los puntos (x_i, y_i) para $i = 0, 1, \dots, n$ se puede obtener de la forma

$$Fit[\{\{x_1, y_1\}, \dots, \{x_n, y_n\}\}, \{1, x\}, x] .$$

26.7.4. Integración

Con el comando $NIntegrate[f, \{x, a, b\}]$ se obtiene una aproximación numérica de la integral

$$\int_a^b f(x) dx .$$

Por supuesto, si es posible, se puede obtener el valor exacto de la anterior integral mediante el comando $Integrate[f, \{x, a, b\}]$. Veamos algunos ejemplos:

- La integral $\int_0^e \sin x dx$ se calcula de forma exacta y numérica, es decir,

$$Integrate[Sin[x], \{x, 0, E\}]$$

produce $1 - Cos[e]$, mientras que con el comando numérico

$$NIntegrate[Sin[x], \{x, 0, E\}]$$

devuelve 1.91173 .

- Para la integral $\int_0^1 e^{x^2} dx$, el comando $Integrate[E \wedge (x \wedge 2), \{x, 0, 1\}]$ de MATHEMATICA devuelve $\frac{1}{2}\sqrt{\pi}Erfi[1]$, es decir, utiliza la función error imaginaria $Erfi[]$ que está definida como $Erfi[z] := \frac{2}{i\sqrt{\pi}} \int_0^{iz} \exp(-t^2) dt$.

Sin embargo $NIntegrate[E \wedge (x \wedge 2), \{x, 0, 1\}]$ da el valor aproximado 1.46265 .

26.7.5. Raíces de ecuaciones

Consideremos una ecuación polinomial de grado n , es decir, $p(x) = 0$ con $p \in \mathbb{R}_n[x]$, siendo $\mathbb{R}_n[x]$ el conjunto de los polinomios de grado n . Se sabe, gracias a la teoría de Galois, que en general no es posible hallar de manera exacta (mediante radicales) las raíces de dicha ecuación cuando $n \geq 5$.

El comando $Solve[p == 0, x]$ de MATHEMATICA intenta hallar la solución exacta de dicha ecuación. Cuando esto no sea posible, el comando de $NSolve[p == 0, x]$ obtiene la aproximación numérica de dichas raíces. En ambos casos, las raíces las ordena en una lista. Veamos algunos ejemplos:

- Las raíces de la ecuación polinomial $x^3 - 2x = 0$ se calculan de forma exacta y numérica, es decir, $Solve[x^3 - 2x == 0, x]$ produce $\{\{x - > 0\}, \{x - > -\sqrt{2}\}, \{x - > \sqrt{2}\}\}$, mientras que con el comando numérico $NSolve[x^3 - 2x == 0, x]$ devuelve $\{\{x - > 0\}, \{x - > -1.41421\}, \{x - > 1.41421\}\}$.
- A diferencia del apartado anterior, si se pretenden hallar de forma exacta las raíces de la ecuación polinomial $x^5 - x - 2 = 0$ utilizando $Solve[x^5 - x - 2 == 0, x]$, entonces MATHEMATICA fracasa. Sin embargo, de forma numérica, utilizando $NSolve[x^5 - x - 2 == 0, x]$ obtiene $\{\{x - > -0.894548 - 0.534149i\}, \{x - > -0.894548 + 0.534149i\}, \{x - > 0.260964 - 1.17723i\}, \{x - > 0.260964 + 1.17723i\}, \{x - > 1.26717\}\}$, donde $i = \sqrt{-1} \in \mathbb{C}$. Nótese cómo no sólo aproxima las raíces reales si no también las complejas.

Si se desea hallar de forma exacta o en su defecto una aproximación numérica de las soluciones de un sistema de ecuaciones polinomiales $p_i(x_1, \dots, x_m) = 0$ con $i = 1, \dots, m$, siendo $p_i \in \mathbb{R}_n[x_1, \dots, x_m]$, la forma de proceder es

$$Solve[\{p_1 == 0, \dots, p_m == 0\}, \{x_1, \dots, x_m\}]$$

o bien

$$NSolve[\{p_1 == 0, \dots, p_m == 0\}, \{x_1, \dots, x_m\}]$$

respectivamente. Así, los puntos de corte entre la circunferencia unidad $x^2 + y^2 = 1$ y la recta bisectriz del segundo cuadrante $y = -x$ se calculan de forma exacta mediante $Solve[\{x^2 + y^2 - 1 == 0, x + y == 0\}, \{x, y\}]$ obteniéndose las parejas de coordenadas en la lista $\{\{x - > -1/\sqrt{2}, y - > 1/\sqrt{2}\}, \{x - > 1/\sqrt{2}, y - > -1/\sqrt{2}\}\}$. De forma numérica se tiene $NSolve[\{x^2 + y^2 - 1 == 0, x + y == 0\}, \{x, y\}]$ que produce por resultado la aproximación $\{\{x - > -0.707107, y - > 0.707107\}, \{x - > 0.707107, y - > -0.707107\}\}$.

Consideremos ahora una ecuación trascendente $f(x) = 0$. El comando

$$NSolve[f == 0, x]$$

de MATHEMATICA intenta hallar una solución aproximada de la ecuación trascendente, aunque en este caso el éxito no está garantizado. Por ejemplo, si se utiliza $Solve[E \wedge x - \sin[x] == 0]$ no se obtendrá ninguna raíz como respuesta de MATHEMATICA. Sin embargo, existe otro comando cuya sintaxis es $FindRoot[f, \{x, x_0\}]$ que implementa el método de Newton-Raphson para hallar una raíz de la ecuación $f(x) = 0$ comenzando la iteración de Newton-Raphson en el punto inicial x_0 . Así, $FindRoot[E \wedge x - \sin[x], \{x, 1\}]$ converge hacia la raíz $\{x - > -6.28131\}$. Por supuesto no halla todas las raíces de la ecuación $f(x) = 0$, observar que en el ejemplo propuesto existen infinitas raíces negativas y ninguna positiva mediante un simple análisis gráfico de las funciones involucradas.

Se puede realizar un análisis gráfico previo a la resolución de la ecuación $f(x) = 0$ dibujando la gráfica de la función $y = f(x)$ y observando visualmente por qué zona del eje de abscisas se produce el corte. De este modo se tiene una estimación inicial de x_0 .

Si se desea hallar una aproximación numérica de alguna solución de un sistema de ecuaciones trascendente $f_i(x_1, \dots, x_m) = 0$ con $i = 1, \dots, m$, se puede utilizar el comando

$$\text{FindRoot}[\{f_1, \dots, f_m\}, \{y_1, \dots, y_m\}]$$

que utiliza el método de Newton-Raphson con un punto inicial $(y_1, \dots, y_m) \in \mathbb{R}^m$.

Por ejemplo, dado el sistema de ecuaciones $\sin x - y + 2 = 0$, $x^2 + \tan(xy) - 1 = 0$, el comando $\text{FindRoot}[\{\text{Sin}[x] - y + 2, x^2 + \text{Tan}[x * y] - 1\}, \{x, 0\}, \{y, 1\}]$ devuelve $\{x \rightarrow 1.05524, y \rightarrow 2.87002\}$.

26.7.6. Ecuaciones diferenciales ordinarias

Consideremos la ecuación diferencial ordinaria de primer orden

$$\frac{dy}{dx} = f(x, y) .$$

MATHEMATICA intenta hallar la solución general exacta $y(x)$ mediante el comando $\text{DSolve}[y'[x] == f[x, y[x]], y[x], x]$. Por ejemplo, para hallar la solución general exacta de la ecuación diferencial lineal de primer orden $dy/dx = y + 2x$, mediante la instrucción $\text{DSolve}[y'[x] == y[x] + 2 * x, y[x], x]$ se obtiene

$$\{\{y[x] \rightarrow -2(1 + x) + e^x C[1]\}\} .$$

La constante arbitraria de integración es denotada en la expresión anterior por $C[1]$. Si se pretende hallar una solución particular de manera exacta, es decir, resolver el problema de Cauchy

$$\frac{dy}{dx} = f(x, y) , \quad y(x_0) = y_0 , \quad (26.1)$$

entonces la condición inicial $y(x_0) = y_0$ se introduce de la forma $\text{DSolve}[\{y'[x] == f[x, y[x]], y[x_0] == y_0\}, y[x], x]$. De este modo, la solución exacta del problema de valor inicial $dy/dx = y + 2x$ con $y(0) = 1$ mediante el comando $\text{DSolve}[\{y'[x] == y[x] + 2 * x, y[0] == 1\}, y[x], x]$ es

$$\{\{y[x] \rightarrow -2(1 + x) + 3e^x\}\} .$$

Sin embargo, sabemos que existen muchas ecuaciones diferenciales de este tipo que no admiten expresar su solución general mediante funciones elementales. Por ejemplo, para la ecuación diferencial de Riccati $dy/dx = y^2 + 2xy$,

la sintaxis $DSolve[y'[x] == y[x] \wedge 2 + 2 * x * y[x], y[x], x]$ devuelve la solución general

$$\left\{ \left\{ y[x] -> -\frac{2e^{x^2}}{2C[1] + \sqrt{\pi} \text{Erfi}[x]} \right\} \right\},$$

expresada a través de la función error imaginaria Erfi que se ha visto anteriormente que es una expresión integral y no es por lo tanto de ayuda. Es en este tipo de casos donde los métodos numéricos tienen sentido de ser aplicados.

La solución numérica del problema de Cauchy (26.1) en el intervalo $x \in [a, b]$ se obtiene con MATHEMATICA mediante el comando $NDSolve[\{y'[x] == f[x, y[x]], y[x_0] == y_0\}, y, \{x, a, b\}]$. En este caso, MATHEMATICA devuelve lo que el llama una *InterpolatingFunction*. Esta solución puede ser dibujada e incluso evaluada en un punto.

Veamos un ejemplo de cómo proceder. Consideremos el problema de Cauchy definido por una ecuación diferencial de Abel $dy/dx = -y^3 + xy$ tal que $y(0) = 1$. Utilizando el comando

```
solucion = NDSolve[{y'[x] == -y[x]^3 + x*y[x], y[0] == 1}, y, {x, 0, 10}]
```

MATHEMATICA devuelve

```
{ {y -> InterpolatingFunction[0., 10., " <> "]} }
```

Observar que hemos llamado *solucion* a la anterior función. Podemos dibujar la solución $y(x)$ del problema de Cauchy anterior en el intervalo $0 \leq x \leq 10$ mediante el comando

```
Plot[y[x]/.solucion, {x, 0, 10}]
```

con lo cual MATHEMATICA produce la gráfica de la Figura 26.1.

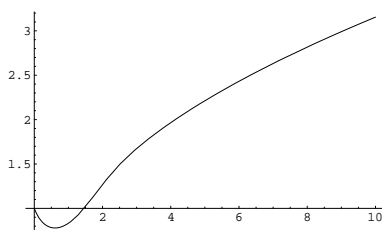


Figura 26.1: Gráfica de la solución del problema de cauchy asociado a la ecuación de Abel.

Además, mediante el comando $y[5]/.solucion$ se obtiene el valor de la solución $y(x)$ en $x = 5$. En concreto MATHEMATICA da $\{2.21258\}$.

Por supuesto MATHEMATICA permite también resolver de forma exacta si se puede o, en su defecto, de forma numérica una ecuación diferencial o problema de Cauchy de orden superior a 1. La sintaxis de los comandos es totalmente análoga a la descrita anteriormente para las ecuaciones diferenciales de primer orden. De este modo, la solución general de la ecuación diferencial de segundo orden $d^2y/dx^2 = 4y + x$ viene dada, mediante el comando $DSolve[y''[x] == 4 * y[x] + x, y[x], x]$, por

$$\left\{ \left\{ y[x] -> -\frac{x}{4} + C[1]e^{-2x} + C[2]e^{2x} \right\} \right\},$$

siendo $C[1]$ y $C[2]$ dos constantes arbitrarias de integración. Del mismo modo, el problema de Cauchy de segundo orden $d^2y/dx^2 = 4y + x$ con $y(0) = 1$, $dy/dx(0) = 2$ se resuelve de manera exacta mediante $DSolve[\{y''[x] == 4 * y[x] + x, y[0] == 1, y'[0] == 2\}, y[x], x]$ obteniéndose

$$\left\{ \left\{ y[x] -> \frac{1}{16}(-e^{-2x} + 17e^{2x} - 4x) \right\} \right\}.$$

Los problemas de Cauchy de orden superior se resuelven también de forma numérica. Por ejemplo, volviendo al caso $d^2y/dx^2 = 4y + x$ con $y(0) = 1$, $dy/dx(0) = 2$ se tendría

$solucion = NDSolve[\{y''[x] == 4 * y[x] + x, y[0] == 1, y'[0] == 2\}, y, \{x, 0, 10\}]$

MATHEMATICA devuelve la función $y(x)$ dada por

$$\{\{y -> InterpolatingFunction[0., 10., " <> "]\}\}$$

que obviamente puede graficarse y evaluarse como hemos comentado anteriormente.

Finalmente, comentemos que los sistemas de ecuaciones diferenciales también pueden resolverse de forma exacta o bien de forma numérica. Así, si se pretende hallar la solución general del sistema de ecuaciones diferenciales $dy_i/dx = f_i(x, y_1, \dots, y_m)$ con $i = 1, \dots, m$, el comando a utilizar es

$$DSolve[\{y1'[x] == f1[x, y1[x], \dots, ym[x]], \dots, ym'[x] == fm[x, y1[x], \dots, ym[x]], \{y1[x], \dots, ym[x]\}, x]$$

Si además se tienen las condiciones iniciales $y_1(x_0) = \alpha_1, \dots, y_m(x_0) = \alpha_m$ entonces se debe utilizar

$$DSolve[\{y1'[x] == f1[x, y1[x], \dots, ym[x]], \dots, ym'[x] == fm[x, y1[x], \dots, ym[x]], y1[x_0] = \alpha_1, \dots, ym[x_0] = \alpha_m\}, \{y1[x], \dots, ym[x]\}, x]$$

La versión numérica del anterior comando es

$$NDSolve[\{y1'[x] == f1[x, y1[x], \dots, ym[x]], \dots, ym'[x] == fm[x, y1[x], \dots, ym[x]], y1[x_0] = \alpha_1, \dots, ym[x_0] = \alpha_m\}, \{y1, \dots, ym\}, \{x, a, b\}]$$

Capítulo 27

Introducción a OCTAVE

27.1. Introducción a Matlab–Octave

Matlab–Octave tiene un lenguaje de programación muy intuitivo junto con una gran biblioteca de funciones de modo que se convierte en una herramienta extremadamente útil para ingenieros y científicos en general. En estas notas se pretende dar unas breves nociones sobre algunos de los comandos utilizados por el programa Matlab–Octave. Por supuesto, una guía completa y mucho más detallada sobre las posibilidades de Octave se encuentra en el manual del programa contenido, por ejemplo, en la página web:

http://www.sce.carleton.ca/faculty/adler/octave/octave_manual.pdf.

En la página web oficial de Octave <http://www.octave.org> se encuentra todo tipo de información sobre el programa.

27.1.1. Algunas convenciones

Existen algunas convenciones en Matlab–Octave. Así, una línea de código que comience por el símbolo `%` se interpreta que es un comentario y por lo tanto no se ejecuta. Cuando en el final de una línea de código se escriben tres puntos, es decir `...`, se toma la línea siguiente como continuación. Un punto y coma de la forma `;` sirve para unir varias sentencias en una misma línea y además no permite salir por pantalla el resultado.

Las instrucciones `exit` o bien `quit` cierra el intérprete mientras que `<CTRL>-c` para la ejecución de los comandos. El comando `help` seguido del nombre de una función permite acceder a la ayuda asociada a una función.

Es conveniente, en muchas situaciones, usar dos ventanas activas en el escritorio. En una ventana se tiene la sesión de trabajo con el intérprete Matlab–Octave y en la otra un editor de texto en el cual modificar el código fuente de un cierto programa.

27.1.2. Primeros cálculos: funciones elementales

El ejemplo más simple es utilizar Matlab–Octave a modo de calculadora, es decir, una entrada y una salida. Al lado del símbolo `>>` se introduce la entrada y la salida se almacena en una variable llamada por defecto **ans** (del inglés answer). Por ejemplo, introduciendo `3+2`, se devuelve **ans=5**.

Por otro lado, sabemos que $\sin(\pi/4) = \sqrt{2}/2 \in \mathbb{R} \setminus \mathbb{Q}$ es un número irracional y tiene por lo tanto infinitas cifras decimales. Sin embargo, introduciendo `sin(pi/4)`, el programa devuelve 0.70711 que no es más que una aproximación con 5 decimales del valor exacto. El mismo valor se obtiene, por supuesto, si se introduce `sqrt(2)/2`.

Recordar que Matlab–Octave, además de utilizar números reales, también trabaja con números complejos. En concreto, la representación del número $a + ib \in \mathbb{C}$ con $a, b \in \mathbb{R}$, siendo $i = \sqrt{-1}$ la unidad imaginaria es **a+bi**.

Matlab–Octave asigna por defecto las constantes **pi**, **e**, **i** a las constantes $\pi \approx 3,1416$, $e \approx 2,7183$ y $i = \sqrt{-1}$. Además, **eps** es la precisión en las operaciones en coma flotante (de hecho es la diferencia entre 1.0 y el número representable inmediatamente superior); **realmax** y **realmin** son el mayor y menor respectivamente número real positivo representable (típicamente del orden de 10^{308} y 10^{-308}); **±inf** números fuera del rango representable; **NaN** (not a number) se guarda para operaciones sin sentido como pueden ser, por ejemplo, indeterminaciones del tipo 0/0.

Todos los cálculos en Matlab–Octave se realizan en doble precisión. Sin embargo, el formato de salida se puede modificar. Por defecto, la salida contiene 4 decimales. En resumen, algunos formatos son:

Formato	Significado
format short	4 decimales
format long	14 decimales
format short e	notación científica con 4 decimales
format long e	notación científica con 14 decimales

La instrucción **who** o bien **whos** devuelve una lista con las variables almacenadas en la sesión de trabajo. Una variable *variable* se elimina de la memoria con el comando **clear variable**. Si se usa sólo **clear** se borran todas las variables de la memoria.

El comando **save** seguido de un nombre de archivo y un conjunto de variables sirve para guardar el archivo en el llamado *directorio de trabajo* (working directory) con el valor de las variables especificadas. Posteriormente, en otra sesión, utilizando el comando **load** se pueden recuperar todas las variables almacenadas. Por ejemplo, **save sesion v1 v2 v3** guarda en el archivo *sesion* el valor que tengan almacenadas las variables *v1*, *v2* y *v3*.

Algunas operaciones algebraicas elementales con números reales *x* e *y* son:

Operación	Sintaxis
$x \pm y$	x±y
xy	x*y
$\frac{x}{y}$	x/y
x^y	x^y

Algunas funciones elementales con números reales x e y son:

Función	Sintaxis	Descripción
$ x $	abs(x)	valor absoluto
$\text{sign}(x)$	sign(x)	-1 si $x < 0$; 0 si $x = 0$; 1 si $x > 0$
$\text{rem}(x/y)$	rem(x,y)	resto del cociente x/y
e^x	exp(x)	exponencial
$\ln x$	log(x)	logaritmo neperiano de $x > 0$
$\log_{10} x$	log10(x)	logaritmo decimal de $x > 0$
\sqrt{x}	sqrt(x)	raíz cuadrada de $x \geq 0$
$\sin(x)$	sin(x)	seno
$\cos(x)$	cos(x)	coseno
$\tan(x)$	tan(x)	tangente
$\arcsin(x)$	asin(x)	arco seno de $x \in [-1, 1]$ da ángulo en $[-\pi/2, \pi/2]$
$\arccos(x)$	acos(x)	arco coseno de $x \in [-1, 1]$ da ángulo en $[0, \pi]$
$\arctan(x)$	atan(x)	arco tangente da ángulo en $[-\pi/2, \pi/2]$

Recordar que los ángulos en los argumentos de las funciones trigonométricas han de estar siempre en radianes.

27.1.3. Matrices y vectores

Matlab significa laboratorio de matrices. Esto indica que las matrices y vectores son pieza clave en su funcionamiento. El programa en realidad no distingue entre escalares y matrices. Además, no es necesario dimensionar las matrices. Esto significa que realiza una asignación dinámica de memoria, de modo que, si se pretende controlar la memoria conviene dimensionar las matrices.

El vector fila $x = (1, 2, 3) \in \mathbb{R}^3$ se introduce de la forma **x=[1,2,3]**, el vector columna $y = (1, 2, 3)^T \in \mathbb{R}^3$ como **y=[1;2;3]** y la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathcal{M}_{2 \times 3}(\mathbb{R}) ,$$

de la forma **A=[1,2,3;4,5,6]**. Entonces, para denominar a sus elementos se utiliza la notación convencional, es decir, **x(2)** devuelve 2, **y(1)** devuelve 1, y finalmente **A(1,2)** es 2.

También se pueden extraer submatrices de una dada. Por ejemplo, si definimos la matriz

$$B = \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & \ell \\ m & n & o & p \end{pmatrix} \in \mathcal{M}_4(\mathbb{R}) ,$$

como $B=[a,b,c,d; e,f,g,h; i,j,k,l; m,n,o,p]$, entonces la submatriz que contiene las filas de la segunda a la cuarta y las columnas de la primera a la tercera

$$\begin{pmatrix} e & f & g \\ i & j & k \\ m & n & o \end{pmatrix}$$

se expresa como $B(2:4,1:3)$. Para tomar el vector de la columna segunda

$$\begin{pmatrix} b \\ f \\ j \\ n \end{pmatrix} \in \mathbb{R}^4,$$

se haría de la forma $B(:,2)$.

Una función interesante aplicable a vectores es el comando `length(x)`, que da como resultado el número de elementos del vector x . Así, `length([a,b,c])` vale 3.

Las matrices se pueden introducir de varias formas. (i) Mediante la introducción explícita de sus elementos (el método utilizado hasta el momento); (ii) Mediante instrucciones especiales: por ejemplo `rand(m,n)` crea una matriz de m filas y n columnas con elementos aleatorios uniformemente distribuidos en el intervalo $[0,1]$; (iii) Creando, en un editor de texto un fichero donde se introducen los elementos de la matriz distribuidos en filas-columnas y lo guardamos, por ejemplo, con el nombre *matriz.dat*, entonces, con el comando `load matriz.dat` se lee el fichero y se asigna la matriz en la variable *matriz*. Un posible fichero con una matriz puede ser simplemente

```
1 2 3 4
5 6 7 8
```

Se puede crear un vector fila x cuyas componentes comiencen en *inicio*, aumenten consecutivamente con *incremento* y acabe en *final* de la forma: $x=[inicio:incremento:final]$. Por ejemplo, $x=[0:0.2:1]$ es equivalente a $x=[0, 0.2, 0.4, 0.6, 0.8, 1]$. Si se omite *incremento*, entonces por defecto vale 1. Una elección particular de *inicio* e *incremento* puede dar lugar a no llegar a *final*. Por ejemplo $x=[0:0.3:1]$ es equivalente a $x=[0, 0.3, 0.6, 0.9]$.

Otra forma de generar un vector fila $x \in \mathbb{R}^n$ de componentes equiespaciadas cuya primera componente es *inicio* y su última componente es *fin* es mediante la instrucción `x=linspace(inicio, fin, n)`. Por ejemplo, `x=linspace(4,6,3)` es equivalente a $x=[4,5,6]$.

Sean $x, y \in \mathbb{R}^n$. Algunas funciones sobre vectores son las siguientes:

Función	Descripción
length(x)	Devuelve n si $x \in \mathbb{R}^n$
dot(x,y)	Producto escalar $x.y$
norm(x)	Norma euclídeana $\ x\ $
x'	Traspuesto conjugado
max(x)	Valor máximo de las componentes de x
min(x)	Valor mínimo de las componentes de x
sum(x)	Suma todas las componentes de x
prod(x)	Multiplica todas las componentes de x

Sea $A \in \mathcal{M}_{n \times m}(\mathbb{R})$. Por lo que se refiere al orden de las matrices, se dispone de la función **size(A)** que devuelve el vector fila $[n, m]$ con las filas y columnas de A .

Para definir rápidamente matrices que contienen ceros y unos se dispone, entre otras, de las siguientes funciones:

Función	Descripción
zeros(n)	Matriz nula cuadrada de orden n
zeros(n,m)	Matriz nula de orden $n \times m$
ones(n)	Matriz de unos cuadrada de orden n
ones(n,m)	Matriz de unos de orden $n \times m$
eye(n)	Matriz identidad de orden n
rand(n,m)	Matriz de orden $n \times m$ con elementos aleatorios uniformemente distribuidos en el intervalo $[0, 1]$

Por ejemplo, **zeros(size(A))** crea una matriz de elementos nulos del mismo orden que la matriz A .

Las operaciones básicas entre matrices A y B de órdenes adecuados para poder realizarse son:

Operación	Descripción
a*A	Producto del escalar a por la matriz A
A±B	Suma y resta
A*B	Multiplicación
A^n	Potencia
A'	Traspuesta conjugada

En Matlab–Octave existe lo que se conoce como operaciones entre matrices *elemento a elemento* y son de gran utilidad. Por supuesto, las matrices deben tener el orden adecuado para poder realizarse la operación. Así, si $A = (a_{ij})$, $B = (b_{ij}) \in \mathcal{M}_{n \times m}(\mathbb{R})$ y $k, l \in \mathbb{R}$ se tiene:

Operación	Descripción
C=k*A ± 1	$C = (c_{ij})$ tal que $c_{ij} = ka_{ij} \pm l$
C=A.*B	$C = (c_{ij})$ tal que $c_{ij} = a_{ij}b_{ij}$
C=A./B	$C = (c_{ij})$ tal que $c_{ij} = a_{ij}/b_{ij}$
C=A.^B	$C = (c_{ij})$ tal que $c_{ij} = a_{ij}^{b_{ij}}$
C=A.^k	$C = (c_{ij})$ tal que $c_{ij} = a_{ij}^k$

Algunas funciones para matrices son:

Función	Descripción
<code>size(A)</code>	Vector fila $[n, m]$ si $A \in \mathcal{M}_{n \times m}(\mathbb{R})$
<code>det(A)</code>	Determinante de A
<code>rank(A)</code>	Rango de A
<code>inv(A)</code>	Inversa de A
<code>x=A\b</code>	Solución del sistema lineal $Ax = b$
<code>eig(A)</code>	Vector columna con los valores propios de A
<code>[P,D]=eig(A)</code>	P es una matriz con los vectores propios de A en columnas
	D es una matriz diagonal con los valores propios de A en la diagonal
<code>poly(A)</code>	Polinomio característico de A
<code>expm(A)</code>	Matriz exponencial de A

Recordemos finalmente que, cualquiera de las funciones elementales que actúan sobre escalares también lo hacen sobre vectores y matrices mediante la actuación elemento a elemento. Por ejemplo, si $\mathbf{x}=[1, 4, 9]$, entonces `sqrt(x)` es justo el vector $[1,2,3]$ o bien `sin([pi/2, pi, 3*pi/2])` devuelve el vector $[1.00000, 0.00000, -1.00000]$.

27.1.4. Funciones

Una de las herramientas más potentes de Matlab–Octave es el concepto de función. A parte de las funciones elementales habituales en un curso de cálculo, como son las trigonométricas, logaritmos, etc..., Matlab–Octave dispone de una enorme biblioteca de funciones matemáticas menos conocidas. Por ejemplo la función gamma de Euler definida de la forma

$$\Gamma(z) = \int_0^{\infty} t^{z-1} \exp(-t) dt ,$$

está implementada en la función `gamma(z)`.

Matlab–Octave busca funciones no sólo en sus bibliotecas sino también en un directorio especial llamado *directorio de trabajo* (working directory). Para conocer el directorio de trabajo existe la instrucción `pwd`. Para cambiar el directorio de trabajo se usa la instrucción `cd`. Con la orden `dir` o bien `ls` se muestra el contenido del directorio de trabajo. Los comandos `delete` y `type` se pueden usar para borrar archivos o mostrarlos por pantalla respectivamente. La forma general de definir una función es

```
function [vs1,...,vsm]=nombre(ve1,...,ven)
...comandos...
endfunction
```

siendo $ve1, \dots, ven$ las n variables de entrada y $vs1, \dots, vsm$ las m variables de salida. Así, por ejemplo, una función que tome como argumento de entrada

un vector x y la salida sea un número real M que consiste en el valor medio de las componentes del vector x es la siguiente. Escribimos en un editor de texto la función

```
function [M]=media(x)
    M=sum(x)/length(x);
endfunction
```

y la grabamos con el mismo nombre, es decir `media.m`, en el directorio de trabajo. Desde el intérprete, se utilizaría, por ejemplo, de la manera siguiente. Si se teclea `media([1,2,3,4,5])` se obtiene 3.

Si se utilizan comentarios mediante `%` antes de la definición de la función, entonces esos comentarios aparecen cuando se escribe `help` seguido del nombre de la función. Si sólo hay una variable de salida, el corchete `[]` es opcional. Usualmente las funciones terminan cuando se alcanza un `endfunction`. Sin embargo se puede forzar un regreso más temprano utilizando el comando `return`.

Si se desea definir una función real de variable real $f : \mathbb{R} \rightarrow \mathbb{R}$, por ejemplo,

$$f(x) = \frac{x^3 \exp(2x)}{\sin x} ,$$

la forma de realizarlo es

```
function y=f(x)
    y=(x ^ 3*exp(2*x))/sin(x);
endfunction
```

Para evaluar la función f en un punto, por ejemplo 3, simplemente se escribe en el intérprete `f(3)` y Matlab-Octave devuelve `7.7187e+04` que significa $7,7187 \times 10^4$.

Por supuesto también es posible definir una función real de varias variables reales $g : \mathbb{R}^m \rightarrow \mathbb{R}$. Por ejemplo, para definir la función

$$g(x, y) = x^3 - 3xy ,$$

escribiríamos

```
function z=g(x,y)
    z=x^3-3*x*y;
endfunction
```

Para evaluar la función g en un punto, por ejemplo $(1, 2)$, se escribe `g(1,2)` y Matlab-Octave devuelve -5.

Finalmente, para definir funciones vectoriales reales de varias variables reales $h : \mathbb{R}^m \rightarrow \mathbb{R}^n$ como puede ser

$$h(x, y) = (x^3y, \cos x, e^{xy}) ,$$

escribiríamos, por ejemplo,

```
function y=h(x)
  y(1)=x(1)^3*x(2);
  y(2)=cos(x(1));
  y(3)=exp(x(1)*x(2));
endfunction
```

Notemos que la variable de entrada de la función h es vectorial de 2 componentes y la de salida vectorial de 3 componentes. La forma de utilizar esta función es del modo siguiente. Si queremos hallar $h(1, 2)$ escribimos `h([1,2])` y Matlab-Octave devuelve `[2.00000, 0.54030, 7.38906]`.

El comando **return**, al ser usado en una función, sale de ésta y devuelve el control al resto del programa que ha llamado a dicha función, llevándose el valor que tenga la variable de salida de la función.

Existe una forma de definir funciones sin necesidad de crear un archivo a parte. Esto se consigue con la sentencia **inline**. Por ejemplo, podemos definir la función $f(x) = \sin(x)\sqrt{x}$ y posteriormente evaluarla en $x = 1$ para obtener el valor de $f(1)$ de la forma:

```
f=inline('sin(x)*sqrt(x)');
f(1)
```

Se obtiene 0,84147 que es el valor de $f(1)$. También se puede utilizar el comando **feval(funcion, argumento)** para evaluar una función en un argumento. De este modo

```
f=inline('sin(x)*sqrt(x)');
feval(f,1)
```

muestra el valor de $f(1)$, es decir, 0,84147. El comando **feval** es necesario cuando se escribe una función que llama a otra función también escrita por nosotros en versiones antiguas de Octave.

27.1.5. Gráficos

Una de las mayores ventajas de Matlab–Octave es su capacidad gráfica para mostrar resultados. Se generan gráficas 2D y 3D. Los comandos básicos son:

`fplot` para dibujar funciones reales de variable real; `plot` para dibujar puntos en \mathbb{R}^2 y `plot3` para dibujar puntos en \mathbb{R}^3 .

Para dibujar una función $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ se utiliza `fplot(f, [a,b])`. Por ejemplo,

```
f=inline('sqrt(x)');
fplot(f, [0,1])
```

muestra la gráfica de la función $f(x) = \sqrt{x}$ en el intervalo $[0, 1]$

Sean $x, y \in \mathbb{R}^n$ vectores fila de componentes $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$. La orden `plot(x,y)` abre una ventana y dibuja los puntos $(x_i, y_i) \in \mathbb{R}^2$ para $i = 1, \dots, n$, donde dos puntos consecutivos son unidos mediante un segmento.

Así, una forma de dibujar una función $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ consiste en generar dos vectores x e y . El primero, x tiene por componentes los puntos de una partición fina y equiespaciada del intervalo $[a, b]$ mientras que el segundo, y , contiene los valores de la función f en las componentes de x . Por ejemplo, para dibujar la función $\cos(x)$ en el intervalo $[-\pi, \pi]$ podemos ejecutar el siguiente código:

```
x=[-pi:0.01:pi]; y=cos(x); plot(x,y)
```

Supongamos que se pretende superponer varias gráficas en una misma figura. Por ejemplo las gráficas de las funciones $\sin(x)$, $\cos(x)$, $\exp(x)$ y $x^2 + 5$ en el intervalo $[-2, 2]$. Entonces, se puede hacer de varias formas. Una de ellas es justo

```
x=[-2:0.01:2]; y=sin(x); z=cos(x); w=exp(x); r=x.^2+5;
plot(x,y,x,z,x,w,x,r)
```

Otra forma es creando una matriz Y cuyas columnas se desean dibujar, es decir,

```
x=[-2:0.01:2]; y=sin(x); z=cos(x); w=exp(x); r=x.^2+5;
Y=[y',z',w',r'];
plot(x,Y)
```

Finalmente, también existe la orden `hold on` que congela la figura actual, de modo que la siguiente gráfica se superpone a la actual. Con `hold off` se elimina esta opción. Entonces, la misma superposición gráfica que la obtenida anteriormente se puede realizar de la forma

```
x=[-2:0.01:2]; y=sin(x); z=cos(x); w=exp(x); r=x.^2+5;
plot(x,y); hold on; plot(x,z); plot(x,w); plot(x,r); hold off;
```

Sean $x, y, z \in \mathbb{R}^n$ vectores fila de componentes $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ y $z = (z_1, \dots, z_n)$. La orden `plot3(x,y,z)` abre una ventana y dibuja los puntos $(x_i, y_i, z_i) \in \mathbb{R}^3$ para $i = 1, \dots, n$, donde dos puntos consecutivos son unidos mediante un segmento.

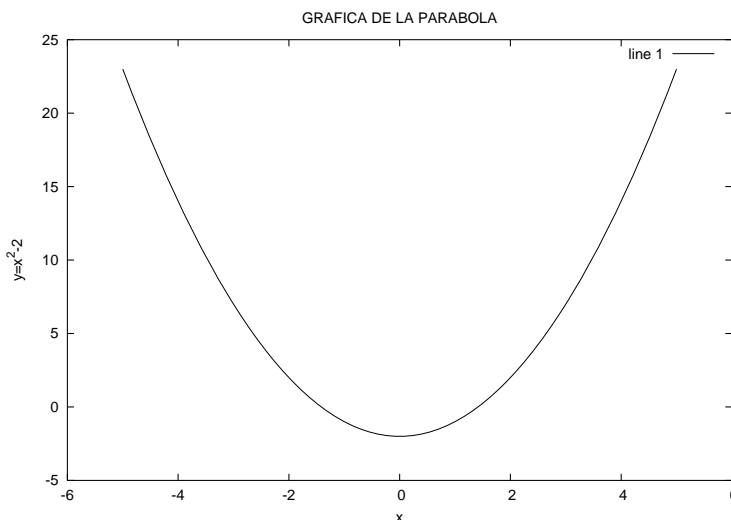
Sea $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^3$. Para dibujar la curva en \mathbb{R}^3 definida de forma paramétrica por las componentes de f en el intervalo de definición $[a, b]$, es decir, la curva $f(t) = (x(t), y(t), z(t))$ con $a \leq t \leq b$, se utiliza también la instrucción `plot3`. De esta forma, si se quiere dibujar la espiral definida de forma paramétrica como $(x(t), y(t), z(t)) = (\cos(t), \sin(t), t)$ para $0 \leq t \leq 10$, teclearemos

```
t=[0:0.01:10];x=cos(t); y=sin(t); z=t; plot3(x,y,z)
```

NOTA: Observar cómo, con el puntero del ratón colocado sobre la ventana gráfica, apretando el botón izquierdo y desplazando, se obtienen perspectivas diferentes de la gráfica 3D. Se puede conseguir el mismo efecto con las teclas de flechas del cursor.

En los gráficos se puede poner un título y etiquetar los ejes. Por ejemplo, el código siguiente genera la gráfica etiquetada de la parábola $y = x^2 - 2$ en el intervalo $[-5, 5]$.

```
x=[-5:0.01:5];y=x.^2-2; xlabel('x'), ylabel('y=x^2-2'),
title('GRAFICA DE LA PARABOLA')
plot(x,y)
```



NOTA: Una forma de grabar en un fichero una gráfica es mediante la instruc-

ción `print('c:nombre.eps','-deps')`. Ejecutándolo, se graba en el disco `c` el fichero `nombre.eps` en formato EPS.

27.1.6. Generación de números aleatorios

En muchos problemas en ingeniería es útil la generación de números aleatorios. Por ejemplo, cuando se quiere simular un sistema complejo o bien como una forma de simular que los datos han sido tomados en un laboratorio con los inherentes errores en las medidas. La función `rand` genera números aleatorios uniformemente distribuidos en el intervalo $[0, 1]$. Así, siendo $n, m \in \mathbb{N}$, `rand(n)` genera una matriz cuadrada de orden n con elementos aleatorios en el intervalo $[0, 1]$. `rand(n,m)` genera una matriz de orden $n \times m$ con elementos aleatorios en el intervalo $[0, 1]$.

27.1.7. Polinomios

Sea $p(x) \in \mathbb{R}_n[x]$ un polinomio a coeficientes reales de grado n , es decir, $f(x) = \sum_{i=0}^n a_i x^i$ con $a_i \in \mathbb{R}$. En Matlab–Octave, los polinomios se introducen como vectores fila de \mathbb{R}^{n+1} que contienen los coeficientes del polinomio p en orden descendente. De este modo, se debe introducir `p=[an, an-1, ..., a1, a0]`. Para evaluar el polinomio en un punto, es decir, calcular $p(c)$ con $c \in \mathbb{R}$, se utiliza el comando `polyval(p,c)`. La derivada de un polinomio $p'(x)$ se obtiene mediante la instrucción `polyder(p)` o `polyderiv(p)`. El polinomio $F(x)$ que es la primitiva del polinomio $p(x)$ tal que $F(0) = 0$, es decir, $F(x) = \int_0^x p(s) ds$, se calcula mediante la instrucción `polyinteg(p)`.

Dados dos polinomios $p(x) \in \mathbb{R}_n[x]$ y $q(x) \in \mathbb{R}_m[x]$, el polinomio producto $p(x)q(x) \in \mathbb{R}_{n+m}[x]$ se obtiene de la forma `conv(p,q)`. Veamos un ejemplo con los polinomios $p(x) = x^3 - x + 1$ y $q(x) = -x + 2$:

```
p=[1,0,-1,1]; q=[-1,2];
polyval(p,4)
polyder(p)
polyinteg(p)
conv(p,q)

devuelve los resultados

41
3  0  -1
0.25000  0.00000  -0.50000  1.00000  0.00000
-1  2  1  -3  2
```

Dado el vector $r = (r_1, \dots, r_n) \in \mathbb{R}^n$, la instrucción `poly(r)` devuelve un vector de \mathbb{R}^{n+1} cuyos coeficientes son los del polinomio mónico de grado n que tiene las raíces r_i para $i = 1, \dots, n$, es decir, del polinomio $\prod_{i=1}^n (x - r_i)$. Por ejemplo, si tomamos `r=[1,2,3]`, entonces `poly(r)` devuelve el vector `[1, -6,`

11, -6]. Notemos que el polinomio $(x - 1)(x - 2)(x - 3)$ tiene la expresión $x^3 - 6x^2 + 11x - 6$.

27.2. El Lenguaje de Programación

El programa Matlab-Octave, además de los múltiples comandos que lleva incorporados, dispone también de un lenguaje de programación de alto nivel.

Las entradas-salidas básicas por teclado-pantalla se realizan con las instrucciones `input` y `disp` respectivamente de la forma siguiente.

```
x=1;
disp('El valor de x es:')
disp(x)
```

devuelve por pantalla

```
El valor de x es:
1
```

Si, por ejemplo, ejecutamos `y=input('Introduce el valor de y:')`, entonces se imprime por pantalla `Introduce el valor de y:` y el ordenador espera a la escritura de un valor por el teclado, asignándolo a la variable `y`.

Recordemos que, los comandos `save` y `load` permiten escribir y leer del disco duro. Por ejemplo, podemos tabular la función $f(x) = \sin(x)$ en el intervalo $[-10, 10]$ con una longitud de paso $h = 1$ y guardar en el directorio de trabajo la tabulación obtenida en un fichero llamado *tabulacion.dat* de la forma siguiente:

```
h=1; x=[-10:h:10]; f=sin(x);
save tabulacion.dat x f
```

Posteriormente, en otra sesión, podemos leer los valores de las abscisas y ordenadas de la tabulación grabada de la forma:

```
load tabulacion.dat
```

de modo que, a partir de esta instrucción, el vector `x` y el vector `f` son reconocidos por el programa como las abscisas y ordenadas de la tabulación.

27.2.1. Iteraciones y bucles

La sentencia `for` funciona de la manera siguiente.

```
for variable = inicial:incremento:final
```

```
.....proceso.....  
endfor
```

Repite el *proceso* que se especifique para los valores de *variable* desde el valor *inicial* hasta el valor *final* saltando de *incremento* en *incremento*. Si se omite el *incremento* entonces este es por defecto 1. Por ejemplo, tras el código

```
for i = 2:2:8  
    disp(i), disp('es multiplo de 2')  
endfor
```

Matlab-Octave devuelve

```
2  
es multiplo de 2  
4  
es multiplo de 2  
6  
es multiplo de 2  
8  
es multiplo de 2
```

Si, en lugar de querer controlar un bucle con un contador se desea hacerlo mediante una condición lógica, podemos utilizar, por ejemplo, la sentencia **while**¹. Esta estructura se utiliza así.

```
while(condicion)  
.....proceso.....  
endwhile
```

El *proceso* se repite mientras *condicion* sea cierta. Por ejemplo

```
a=0;  
while(a <3)  
    disp(a)  
    a=a+1;  
endwhile
```

tiene por salida

```
0  
1
```

¹También es posible una estructura del tipo **do--until**.

2

Existen instrucciones que permiten irse de un bucle. Por ejemplo, la sentencia **break** dentro de un bucle hace que salga de este bucle y siga o bien ejecutando otro bucle externo o bien continúe el programa. Por el contrario, la sentencia **continue** no nos saca del bucle, simplemente se salta un paso.

Existe otra estructura controlada mediante una condición lógica. Es la sentencia **if**. Su estructura es

```
if(condicion)
.....proceso.....
endif
```

El *proceso* se ejecuta sólo si *condicion* sea cierta. Una generalización obvia es la siguiente:

```
if(condicion 1)
.....proceso 1.....
elseif(condicion 2)
.....proceso 2.....
else
.....proceso 3.....
endif
```

Por ejemplo, el siguiente código asigna a la variable *p* los valores 0, 1 y 2 dependiendo de si el número $n \in \mathbb{Z}$ es negativo o cero, positivo e impar o positivo y par respectivamente².

```
if(n <= 0)
  p=0;
elseif(rem(n,2)==0)
  p=2;
else
  p=1;
endif
```

27.2.2. Condicionales y operadores lógicos

En muchas ocasiones, es necesario incluir en las sentencias condicionales anteriores algún operador lógico. Estos operadores son

²Se utiliza la instrucción **rem(a,b)** que proporciona el resto de la división entre *a* y *b*.

Operador	Significado
!	NO
&	Y
	O

Los operadores de relación son:

Operador	Significado
<	menor que
>	mayor que
<=	menor o igual que
>=	mayor o igual que
==	igual que
!= o bien ~=	distinto que

27.2.3. Errores

Debido a errores de redondeo, se define el *epsilon de la máquina* como el mayor número positivo ϵ tal que $1 + x = 1$ para todo $x \in (0, \epsilon)$. Si intentamos averiguar ϵ con un programa escrito en el lenguaje de Matlab–Octave podríamos intentar lo siguiente.

```
x=1;
while(x+1>1)
    x=x/2;
endwhile
disp('El epsilon de la maquina es:')
disp(2*x)
```

La salida del programa es

```
El epsilon de la maquina es:
2.2204e-16
```

De hecho, el mismo valor está almacenado en la variable `eps`.

Un sencillo ejercicio donde se muestra el efecto debido al epsilon de la máquina es el siguiente. Se sabe que

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = 2,7182818.....$$

Sin embargo, si programamos

```
for i=1:10
    n=100\i; disp((1+1/n)^n)
```

```
endfor
```

la salida es la siguiente:

```
2.7048
2.7181
2.7183
2.7183
2.7183
2.7185
2.7161
1
1
1
```

Se observa que, a medida que n aumenta, las aproximaciones obtenidas del número e son mejores hasta que llega un momento en que van empeorando, llegando a estabilizarse en 1.

Siempre que se restan magnitudes muy próximas, se produce un *error por cancelación* con la consiguiente pérdida de cifras significativas. Si además esta resta es multiplicada por un número grande, el error se verá aumentado. Por ejemplo, la expresión $x(\sqrt{x+1} - \sqrt{x})$ para $x \geq 0$ es equivalente, multiplicando y dividiendo por el conjugado, a la expresión

$$\frac{x}{\sqrt{x+1} + \sqrt{x}},$$

de donde se ve que, para valores grandes de x es aproximadamente igual a $\sqrt{x}/2$ puesto que $\sqrt{x+1} \approx \sqrt{x}$. Sin embargo, la primera expresión es prácticamente cero debida a errores de redondeo en la resta entre paréntesis. Veamos este efecto de los errores por cancelación con el siguiente programa.

```
for n=1:10
    x=10^n;
    mal=x*(sqrt(x+1)-sqrt(x)); bien=x/(sqrt(x+1)+sqrt(x));
    A(n,:)=[mal,bien];
endfor
```

Observar que la primera columna de la matriz A tiende hacia cero a medida que aumenta la fila.

27.2.4. Comparando la eficiencia de algoritmos

La eficiencia de un algoritmo puede ser medido en términos de el número de operaciones en coma flotante realizados (flops) o bien en el tiempo total (en

segundos) de ejecución.

El comando `flops(0)` inicializa `flops` a cero. Así, poniendo `flops(0)` antes de comenzar un algoritmo y poniendo `flops` al finalizarse se obtiene el número de `flops` utilizados en el algoritmo.

El comando `tic` inicializa el reloj a cero y `toc` devuelve el tiempo pasado. Así, poniendo `tic` antes de comenzar un algoritmo y poniendo `toc` al finalizarse se obtiene los segundos de tiempo utilizados en realizarse el algoritmo.

Por ejemplo, dado un vector columna $b \in \mathbb{R}^n$ y una matriz inversible $A \in \mathcal{M}_n(\mathbb{R})$, se desea resolver el sistema lineal $Ax = b$. Con Matlab–Octave, existen dos formas diferentes de hacerlo. La primera es utilizando el algoritmo de eliminación Gaussiana de la forma $\mathbf{x} = \mathbf{A} \setminus \mathbf{b}$. La segunda es mediante el cálculo de la matriz inversa, es decir, $\mathbf{x} = \text{inv}(\mathbf{A}) * \mathbf{b}$. Se puede comparar la eficiencia de los dos algoritmos de la forma

```
flops(0), x = A \ b; flops
flops(0), x=inv(A)*b; flops
```

o bien de la forma

```
tic, x = A \ b; toc
tic, x=inv(A)*b; toc
```

¿Qué algoritmo es el más eficiente?

27.3. Métodos Numéricos

27.3.1. Polinomio interpolador

Con Matlab–Octave es posible hallar el polinomio $P_n(x) = \sum_{i=0}^n c_i x^i$ de grado menor o igual que n que interpola a una nube de puntos (x_i, y_i) para $i = 0, 1, \dots, n$. La sintaxis a utilizar es `polyfit(x,y,n)` siendo $\mathbf{x}=[x_0, \dots, x_n]$, $\mathbf{y}=[y_0, \dots, y_n]$ los vectores fila con las abscisas y ordenadas, respectivamente, de la nube de puntos. El resultado es el polinomio $P_n(x)$ en la notación usual de Matlab–Octave, es decir, el vector fila $(c_n, c_{n-1}, \dots, c_1, c_0)^T$ que contiene los coeficientes de $P_n(x)$ en orden descendente.

Existen otras formas de operar. Por ejemplo, `vander(x) \ y'`, siendo $\mathbf{x}=[x_0, \dots, x_n]$, $\mathbf{y}=[y_0, \dots, y_n]$ los vectores fila con las abscisas y ordenadas, respectivamente, de la nube de puntos. El resultado es el polinomio $P_n(x)$ dado como un vector columna que contiene los coeficientes de $P_n(x)$ en orden descendente.

Por ejemplo, la recta en el plano que pasa por los puntos $(0, 1)$ y $(3, 7)$ es $y = P_1(x) = 1 + 2x$. Se puede hallar de la forma `vander([0, 3]) \ [1, 7]'`, siendo el resultado el vector columna

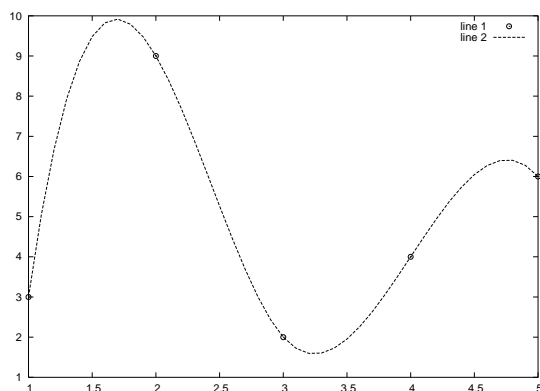
```
2
1
```

27.3.2. Splines Cúbicos

Se pretende calcular una función spline cúbica que interpole a la nube de puntos $(x_i, y_i) \in \mathbb{R}^2$, $i = 1, \dots, n$.

Matlab y Octave disponen de la función `spline(x,y,xe)` para el cálculo de una función spline cúbica. Sus dos primeros argumentos son $x = [x_1, \dots, x_n]$, $y = [y_1, \dots, y_n]$, es decir, dos vectores fila que contienen las abscisas y ordenadas respectivamente de los puntos a interpolar. El tercer argumento `xe` corresponde a un vector fila de \mathbb{R}^k , que contiene las k abscisas donde se desea evaluar la función spline. Como resultado, la función `spline(x,y,xe)` devuelve un vector fila de \mathbb{R}^k con los valores de la función spline en las abscisas `xe`. Si se toman las componentes de `xe` como los puntos de una partición fina del intervalo $[x_1, x_n] \subset \mathbb{R}$ de definición de la función spline, entonces luego se puede realizar un dibujo de la función spline. Veamos un ejemplo.

```
x=[1,2,3,4,5]; y=[3,9,2,4,6]; xe=[1:0.2:5];
ye=spline(x,y,xe);
plot(x,y, 'o', xe,ye)
```



Matlab–Octave devuelve como resultado la gráfica de la función spline en el intervalo $[1, 5]$ superpuesta a la nube de puntos de interpolación.

27.3.3. Integración Numérica

Para calcular una aproximación numérica de la integral $\int_a^b f(x) dx$, podemos utilizar la función `quad(f,a,b)`. La función $f(x)$ se debe introducir, o bien mediante un archivo de función o bien mediante la sentencia `inline`. Por ejemplo, para calcular $\int_0^1 \exp(-x^2) dx$, podemos introducir

```
quad(inline('exp(-x^2)'),0,1)
```

y se obtiene 0,74682.

Puede aproximar numéricamente integrales impropios. Por ejemplo, se sabe que $\int_0^\infty \exp(-x^2) dx = \sqrt{\pi}/2$. Entonces

```
quad(inline('exp(-x^2)'),0,inf)
```

da como resultado 0,88623.

27.3.4. Ecuaciones Diferenciales Ordinarias

En este caso, las diferencias entre Matlab y Octave son importantes, ganando Octave con diferencia. Esto es debido a que Octave incorpora uno de los paquetes de integración numérica de ecuaciones diferenciales ordinarias más robusto y efectivo llamado ODEPACK. Sólo explicaremos el uso en Octave en esta sección.

Consideremos la ecuación diferencial ordinaria de primer orden $\dot{x} = f(t, x)$. Sabemos que existen muchas ecuaciones diferenciales de este tipo que no admiten expresar su solución general mediante funciones elementales. Por ejemplo, la ecuación diferencial de Riccati $\dot{x} = x^2 + 2tx$. Es en este tipo de casos donde los métodos numéricos tienen sentido de ser aplicados.

Pretendemos pues hallar la solución numérica del problema de Cauchy

$$\dot{x} = f(t, x), \quad x(t_0) = x_0,$$

en el intervalo $t \in [t_0, t_f] \subset \mathbb{R}$. Aquí, $f : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$. La instrucción para conseguirlo es `lsode('f', x0, tau)`, donde $x_0 \in \mathbb{R}^n$ es un vector de condiciones iniciales y τ es un vector cuyas componentes contienen los tiempos donde se aproxima la solución $x(t)$ del problema de Cauchy. Por coherencia, la primera componente de τ es t_0 y la última t_f . La función $f(t, x)$ debe ser definida en un fichero de función a parte.

Por ejemplo, consideremos la ecuación de van der Pol $\ddot{x} + \mu(\dot{x}^2 - 1) + x = 0$. Pasamos la ecuación de segundo orden a un sistema de primer orden de la forma

$$\dot{x} = y, \quad \dot{y} = \mu(1 - x^2) - x. \quad (27.1)$$

Supongamos que se desea aproximar la solución de (27.1) utilizando la condición inicial $(x(0), y(0)) = (2, 0)$ en el intervalo de tiempos $t \in [0, 20]$. En primer

lugar, escribimos en un fichero la función $f(x, y)$ que define el sistema de van der Pol, por ejemplo con $\mu = 1$, de la forma:

```
function xp=f(x)
    mu=1;
    xp=zeros(2,1);
    xp(1)=x(2);
    xp(2)=mu*x(2)*(1-x(2)^2)-x(1);
endfunction
```

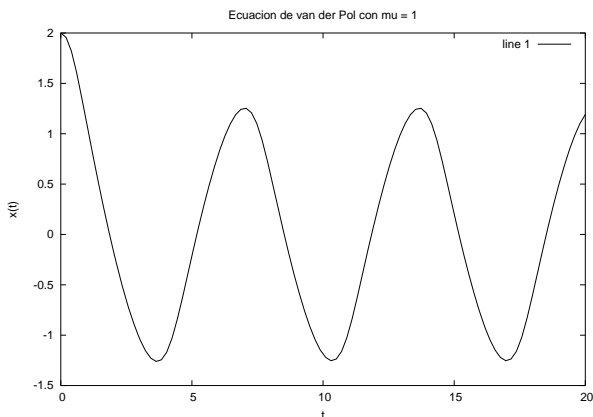
y lo guardamos en el directorio de trabajo con el nombre *f.m*. A continuación, tomamos 100 tiempos equiespaciados t_i con $i = 1, \dots, 100$, en el intervalo $t \in [0, 20]$ y resolvemos de la forma:

```
x0=[2,0];
tau=linspace(0,20,100);
sol=lsode('f',x0,tau);
```

Se obtiene una matriz 100×2 llamada *sol*, donde la fila i -ésima contiene una aproximación del vector $(x(t_i), y(t_i))$. Podemos dibujar la solución $x(t)$ en el intervalo $t \in [0, 20]$ de la forma

```
x=sol(:,1);
xlabel('t') , ylabel('x(t)') ,
title('Ecuacion de van der Pol con mu = 1') ,
plot(tau,x)
```

obteniendo la gráfica siguiente.



27.3.5. Raíces de Ecuaciones

Ecuaciones polinomiales

Consideremos una ecuación polinomial de grado n , es decir, $f(x) = 0$ con $f \in \mathbb{R}_n[x]$, siendo $\mathbb{R}_n[x]$ el conjunto de los polinomios de grado n , es decir, $f(x) = \sum_{i=0}^n a_i x^i$ con $a_i \in \mathbb{R}$. Se sabe, gracias a la teoría de Galois, que en general no es posible hallar de manera exacta (mediante radicales) las raíces de dicha ecuación cuando $n \geq 5$ con lo cual es necesario aproximar las raíces mediante métodos numéricos.

Pretendemos hallar todas las raíces x^* (reales o complejas) del polinomio f , es decir, todos los valores $x^* \in \mathbb{C}$ tales que $f(x^*) = 0$. Para ello, Matlab y Octave tienen la función `roots(p)`. El argumento de la función es un vector fila de \mathbb{R}^{n+1} que contiene los coeficientes del polinomio f en orden descendente. De este modo, se debe introducir $\mathbf{p}=[a_n, a_{n-1}, \dots, a_1, a_0]$. El resultado es un vector columna con todas las raíces del polinomio.

Por ejemplo, si $f(x) = 10 - 19x + 13x^2 - 5x^3 + x^4$, entonces se debe introducir el siguiente código.

```
p=[1,-5,13,-19,10]; x=roots(p)
```

Matlab-Octave devuelve como resultado:

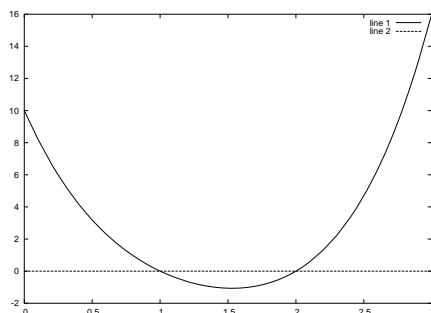
```
x=
1.0000 + 2.0000i
1.0000 - 2.0000i
2.0000
1.0000
```

El resultado obtenido no es otro que el que se obtendría mediante la Regla de Ruffini, es decir, la factorización $f(x) = (x-1)(x-2)(x^2-2x+5)$.

Se puede realizar finalmente una comprobación gráfica donde dibujemos el polinomio $f(x)$, por ejemplo en el intervalo $[0, 3] \subset \mathbb{R}$ y observemos las raíces reales como puntos de corte con el eje de abscisas. Una forma de hacerlo es la siguiente.

```
p=[1,-5,13,-19,10]; x=[0:0.1:3]; y=polyval(p,x);
plot(x,y) , hold
xx=[0,3]; yy=[0,0]; plot(xx,yy)
```

Matlab-Octave devuelve como resultado la superposición de la gráfica del polinomio $f(x)$ en el intervalo $[0, 3] \subset \mathbb{R}$ con el eje de abscisas (segmento que une el punto $(0, 0)$ con el punto $(0, 3)$).



Ecuaciones Trascendentes

Consideremos a partir de ahora una ecuación trascendente $f(x) = 0$. El comando `fzero(f, x0)` intenta hallar una solución aproximada de la ecuación trascendente $f(x) = 0$ comenzando la búsqueda en el punto inicial x_0 . Si no se converge hacia una solución, devuelve por respuesta `NaN`. También se puede utilizar de la forma `fzero(f, [x0, x1])` si sabemos que la raíz buscada se encuentra en el intervalo $[x_0, x_1] \subset \mathbb{R}$.

Así, supongamos que $f(x) = \exp(x) - \sin(x)$. Observar que, en el ejemplo propuesto, existen infinitas raíces negativas y ninguna positiva mediante un simple análisis gráfico de las funciones involucradas. Entonces el código

```
f=inline('exp(x)-sin(x)');
raiz=fzero(f,1)
```

da como resultado `raiz=NaN` de modo que no converge hacia ninguna solución. Se puede realizar un análisis gráfico previo a la resolución de la ecuación $f(x) = 0$ dibujando la gráfica de la función $y = f(x)$ y observando visualmente por qué zona del eje de abscisas se produce el corte. De este modo se tiene una estimación inicial de x_0 . Entonces, cambiando la condición inicial x_0 de la forma

```
f=inline('exp(x)-sin(x)');
raiz=fzero(f,-5)
```

da como resultado `raiz=-6.2813`. Por supuesto, no halla todas las raíces de la ecuación $f(x) = 0$, sólo ha hallado una de ellas.

Si se desea hallar una aproximación numérica de alguna solución de un sistema de ecuaciones trascendente $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ con $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ de componentes $f_i(x_1, \dots, x_m) = 0$ con $i = 1, \dots, m$, se puede utilizar el comando `fsolve('f', x0)`. Para su utilización, se debe crear a parte un fichero que defina la función $f(x)$. Por ejemplo, supongamos que se quiere hallar una solución del sistema de ecuaciones $\sin x - y + 2 = 0$, $x^2 + \tan(xy) - 1 = 0$ partiendo del

punto inicial $\mathbf{x}_0 = (0, 1) \in \mathbb{R}^2$. En primer lugar creamos el fichero

```
function y=f(x)
y(1)=sin(x(1))-x(2)+2;
y(2)=x(1)^2+tan(x(1)*x(2))-1;
endfunction
```

y lo guardamos en el directorio de trabajo con el nombre *f.m*. Posteriormente, ejecutamos en el intérprete

```
raiz=fsolve('f',[0;1])
```

y se obtiene

```
raiz=
1.05524
2.87002
```

27.3.6. Aproximación Polinomial por Mínimos Cuadrados

Matlab–Octave tiene la posibilidad de hallar el polinomio

$$f^*(x) = \sum_{i=0}^m c_i x^i \in \mathbb{R}_m[x]$$

que mejor aproxima (en el sentido de mínimos cuadrados) a la nube de puntos $(x_i, y_i) \in \mathbb{R}^2$ para $i = 0, 1, \dots, n$. La sintaxis a utilizar es `polyfit(x,y,m)` siendo $\mathbf{x}=[x_0, \dots, x_n]$, $\mathbf{y}=[y_0, \dots, y_n]$ los vectores fila con las abscisas y ordenadas, respectivamente, de la nube de puntos y $m \in \mathbb{N}$ el grado del polinomio de aproximación. El resultado es el polinomio f^* en la notación usual de Matlab–Octave, es decir, el vector columna que contiene los coeficientes de f^* en orden descendente, es decir, $(c_m, c_{m-1}, \dots, c_1, c_0)^T$. Como se ha visto en una sección anterior, en el caso particular de que $m = n$, `polyfit(x,y,n)` devuelve el polinomio interpolador de grado menor o igual que n .

Por ejemplo, la recta de regresión en el plano que pasa por los puntos $(0, 1)$, $(2, 3)$ y $(3, 5)$ se puede obtener de la forma:

```
x=[0,2,3]; y=[1,3,5];
p=polyfit(x,y,1)
```

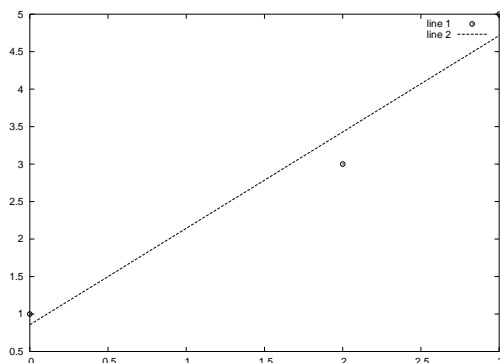
Matlab–Octave devuelve como resultado

```
1.28571
0.85714
```

que se interpreta diciendo que la recta de regresión es $y = f^*(x) = 1,28571x + 0,85714$.

Podemos visualizar el resultado, superponiendo la gráfica de la nube de puntos y la gráfica de la recta de regresión:

```
xx=[0:0.1:3];
pp=polyval(p,xx);
plot(x,y,'*')
hold on
plot(xx,pp)
```



Sea $Ax = b$ un sistema lineal sobredeterminado, es decir, con más ecuaciones que incógnitas. De este modo, se tiene que $A \in \mathcal{M}_{m \times n}$, $x \in \mathbb{R}^n$ y $b \in \mathbb{R}^m$ con $m > n$. Entonces, el comando `A \ b` calcula la “solución” del sistema dentro del contexto de mínimos cuadrados, es decir, calcula el valor de $x \in \mathbb{R}^n$ tal que la norma $\|b - Ax\|_2$ sea mínima.

Bibliografía

- [1] M. ABRAMOWITZ, I.A. STEGUN. *Handbook of mathematical functions with formulas, graphs and mathematical tables*. Dover, 1964.
- [2] A. AUBANELL, A. BENSENY, A. DELSHAMS. *Eines Bàsiques de Càlcul Numèric*. Manuals de la UAB, 1991. Versión en castellano: *Útiles básicos de cálculo numérico*. Labor, 1993.
- [3] G. J. BORSE. *Programación en FORTRAN 77 con aplicaciones de cálculo numérico en ciencias e ingeniería*. Anaya, 1989.
- [4] J. CHAVARRIGA, I.A. GARCÍA, J. GINÉ. *Manual de Métodos Numéricos*. Eines **35**, Edicions de la Universitat de Lleida, 1999.
- [5] G. DALQUIST, A. BJÖRCK. *Numerical methods*. Prentice Hall, 1974.
- [6] B.P. DEMIDOVICH, I.A. MARON. *Cálculo Numérico Fundamental*. Paraninfo, 1988.
- [7] L. ELDÉN, L. WITTMAYER-KOCH. *Numerical Analysis*. Academic Press, 1990.
- [8] C. FROBERG. *Introducción al análisis numérico*. Vicens-Vives, 1974.
- [9] I.A. GARCÍA. *Teoría de Estabilidad y Control*. Eines **52**, Edicions de la Universitat de Lleida, 2005.
- [10] M. GRAU, M. NOGUERA. *Càlcul Numèric*. Edicions UPC, 1993.
- [11] P. HENRICI. *Elements of numerical analysis*. John Wiley, New York, 1964.
- [12] E. ISAACSON, B. KELLER. *Analysis of numerical methods*. John Wiley & sons, 1966.
- [13] D. KINCAID, W. CHENEY. *Análisis numérico*. Addison-Wesley, 1994.
- [14] R. KRESS. *Numerical Analysis*. Graduate texts in Mathematics, Springer-Verlag, 1998.
- [15] S.S. KUO. *Numerical methods and computers*. Addison-Wesley, 1965.

- [16] J. H. MATHEWS. *Numerical Methods for Mathematics, Science and Engineering*. Prentice Hall, 1992.
- [17] A. RALSTON. *Introducción al análisis numérico*. Limusa-Wiley, 1970.
- [18] H.R. SCHWARZ. *Numerical Analysis*. John Wiley & sons, 1989.
- [19] J. STOER, R. BULIRSCH. *Introduction to Numerical Analysis*. Texts in Applied Mathematics, 12. Springer-Verlag, 1972.
- [20] F. SCHEID, R.E. DI COSTANZO. *Métodos numéricos*. MacGraw-Hill, 1991.

$$f(x) \; = \; 0 \; .$$

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

$$P_n(x) \; = \; \sum_{k=0}^n f(x_k) \prod_{\substack{j=0 \\ j \neq k}}^n \frac{x-x_j}{x_k-x_j} \; .$$

$$\begin{array}{l} \frac{dy}{dx} \; = \; f(x,y) \; , \; \; y(x_0) = y_0 \; . \\ y(x_{i+1}) \; \approx \; y_{i+1} = y_i + hf(x_i,y_i) \; . \end{array}$$

$$f(x) \; = \; 0 \; .$$

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



Universitat de Lleida

$$P_n(x) \; = \; \sum_{k=0}^n f(x_k) \prod_{\substack{j=0 \\ j \neq k}}^n \frac{x-x_j}{x_k-x_j} \; .$$

$$\begin{array}{l} \frac{dy}{dx} \; = \; f(x,y) \; , \; \; y(x_0) = y_0 \; . \\ y(x_{i+1}) \; \approx \; y_{i+1} = y_i + hf(x_i,y_i) \; . \end{array}$$

$$\begin{array}{l} \frac{dy}{dx} \; = \; f(x,y) \; , \; \; y(x_0) = y_0 \; . \\ y(x_{i+1}) \; \approx \; y_{i+1} = y_i + hf(x_i,y_i) \; . \end{array}$$

$$f(x) \; = \; 0 \; .$$

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