

Herramientas Computacionales — Notas de clase

Andrés González Mancera

23 de abril de 2010

Andrés González Mancera
2010



Este documento está bajo una licencia Reconocimiento-Compartir bajo la misma licencia 2.5 Colombia de Creative Commons. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-sa/2.5/co/> o envíe una carta a Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

Índice general

1. Introducción	4
1.1. Conceptos básicos	4
1.1.1. Límites, continuidad y derivadas	4
1.1.2. Teorema de Taylor	5
1.2. Análisis de error	7
1.2.1. Pérdida de significancia	8
1.2.2. Orden de Aproximación $\mathcal{O}(h^n)$	10
2. Introducción a la programación en MATLAB	13
2.1. Introducción	13
2.2. Matemática simple	13
2.3. Scripts en archivos-M (M-Files)	15
2.4. Arreglos y operaciones sobre arreglos	17
2.5. Algebra lineal	20
2.6. Control de flujo	22
2.6.1. Ciclos FOR	22
2.6.2. Estructura IF-ELSE-END	22
3. Algebra lineal	24
3.1. Distribución de temperatura en una placa	24
3.2. Métodos Iterativos	30
3.2.1. Iteraciones de punto fijo para una ecuación	30
3.2.2. Iteraciones de punto fijo para un sistema de ecuaciones	31
3.2.3. Métodos iterativos de Jacobi, Gauss-Seidel y SOR	32
3.3. Solución de flujo potencial el dos dimensiones usando diferencias finitas	35
4. Solución de ecuaciones no lineales	39
4.1. Ecuación de estado de Redlich-Kwong en termodinámica	39
4.2. Raíces de ecuaciones	40
4.2.1. Método de la bisección	40
4.2.2. Método de Newton	41
4.2.3. Método de la secante	45

5. Ajuste de curvas	46
5.1. Regresión lineal por mínimos cuadrados	47
5.1.1. Ajuste lineal por mínimos cuadrados	47
5.1.2. Ajuste de ecuación de potencia $y = Ax^M$	48
5.1.3. Cuantificación del error en la regresión lineal	49
5.1.4. Método de linealización de datos para $y = Ce^{Ax}$	50
5.2. Regresión polinomial por mínimos cuadrados	52
5.3. Regresión lineal múltiple	53
6. Interpolación	55
6.1. Interpolación polinomial	55
6.1.1. Sistema de Vandermode	56
6.1.2. Interpolación de Lagrange	56
6.1.3. Interpolación de Neville	58
6.1.4. Interpolación de Newton y diferencias divididas	59
6.1.5. Errores en la interpolación polinomial	61
6.2. Interpolación por partes y Trazas	62
6.2.1. Interpolación lineal	62
6.2.2. Interpolación cuadrática	63
6.2.3. Interpolación por trazadores cúbicos.	64
7. Aproximando derivadas	66
7.1. Diferencias finitas	66
8. Integración numérica	68
8.1. Cálculo de integrales en una dimensión	68
8.1.1. ¿Cuándo necesitamos integración numérica?	68
8.1.2. Estrategia general para la integración numérica	69
8.2. Integración mediante interpolación polinomial local: Reglas de Newton-Cotes	69
8.2.1. Regla del trapecoide	69
8.2.2. Regla de Simpson 1/3	70
8.2.3. Problemas	70
9. Solución de ecuaciones diferenciales	73

Capítulo 1

Introducción

En este curso se desarrollarán los conocimientos y habilidades para el estudio, desarrollo y análisis de algoritmos para obtener soluciones numéricas de varios problemas matemáticos y de ingeniería. Estudiaremos lo que se conoce como las *matemáticas de la computación científica*.

Este capítulo comienza con un repaso de algunos temas importantes para el curso.

1.1. Conceptos básicos

Comenzamos repasando algunos conceptos básicos del cálculo.

1.1.1. Límites, continuidad y derivadas

Si f es una función real de una variable real, entonces el *límite* de la función f en c (si este existe) se define de la siguiente forma.

Definición 1.1 *La ecuación*

$$\lim_{x \rightarrow c} f(x) = L$$

significa que para cada ϵ positivo corresponde un δ positivo tal que la distancia entre $f(x)$ y L es menor que ϵ cuando la distancia entre x y c es menor que δ .

$$|f(x) - L| < \epsilon \text{ cuando } 0 < |x - c| < \delta$$

Definición 1.2 *Asuma que $f(x)$ se define en un conjunto S de números reales y que $x_0 \in S$. Se dice que la función f es continua en c si*

$$\lim_{x \rightarrow x_0} f(x) = f(x_0)$$

Se dice que la función f es continua en S si es continua en cada punto $x \in S$. La notación $C^n(S)$ se usa para referirse al conjunto de todas las funciones f tal que f y sus primeras n derivadas son continuas en S .

Ahora un teorema sencillo.

Teorema 1.1 Teorema del valor intermedio para funciones continuas. Asuma que $f \in C[a, b]$ y L es cualquier número entre $f(a)$ y $f(b)$. Entonces existe un número c , con $c \in (a, b)$, tal que $f(c) = L$.

Ejemplo 1.1 La función $f(x) = \cos(x - 1)$ es continua en el intervalo $[0, 1]$, y la constante $L = 0,8 \in (\cos(0), \cos(1))$. La solución para $f(x) = 0,8$ en el intervalo $[0, 1]$ es $c_1 = 0,356499$. De igual manera $f(x)$ es continua en el intervalo $[1, 2,5]$, y la $L = 0,8 \in (\cos(1), \cos(2,5))$. La solución para $f(x) = 0,8$ en el intervalo $[1, 2,5]$ es $c_2 = 1,643502$.

Ahora definamos la derivada.

Definición 1.3 La derivada de f en c (si esta existe) se define por la ecuación

$$f'(c) = \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c} \quad (1.1)$$

Como este límite no tiene que existir para una función y un c particular, es posible que la derivada no exista para dicha función. Si f es una función para la que $f'(c)$ existe, entonces se dice que f es *diferenciable* en c . La derivada en un punto x es la la pendiente de la línea tangente a la curva en $f(x)$.

Ahora un poco de notación. El conjunto de todas las funciones que son continuas en toda la línea de números reales \mathbb{R} se denota como $C(\mathbb{R})$. El conjunto de todas las funciones para las que f' es continuo en todas partes se denota como $C^1(\mathbb{R})$. De igual forma definimos $C^n(\mathbb{R})$ para cada número natural n como el conjunto de todas las funciones para las que $f^{(n)}$ es continuo. Por último, $C^\infty(\mathbb{R})$ es el conjunto de funciones para el que todas sus derivadas son continuas. Por ejemplo, e^x .

De la misma forma, definimos $C^n[a, b]$ como el conjunto de funciones f para las cuales $f^{(n)}$ existe y es continua en el intervalo cerrado $[a, b]$.

Teorema 1.2 Teorema del Valor Medio. Asuma que $f \in C[a, b]$ y que $f'(x)$ existe para todo $x \in (a, b)$. Entonces existe un número $c \in (a, b)$, tal que

$$f'(c) = \frac{f(b) - f(a)}{b - a} \quad (1.2)$$

Geométricamente, el Teorema del Valor Medio dice que existe por lo menos un número $c \in (a, b)$ tal que la pendiente de la línea tangente a la gráfica de $y = f(x)$ en el punto $(c, f(c))$ es igual a la pendiente de la línea secante que pasa por los puntos $(a, f(a))$ y $(b, f(b))$.

1.1.2. Teorema de Taylor

Un importante teorema concerniente a funciones pertenecientes al conjunto $C^n[a, b]$ es el Teorema de Taylor, el cual aparecerá a todo lo largo del estudio y desarrollo de algoritmos numéricos en la computación científica.

Teorema 1.3 Teorema de Taylor. Si $f \in C^n[a, b]$ y si $f^{(n+1)}$ existe en el intervalo abierto (a, b) , entonces para cualquier punto x y cualquier punto c en el intervalo cerrado $[a, b]$ se tiene

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(c)(x-c)^k}{k!} + E_n(x), \quad (1.3)$$

donde para algún punto ξ entre c y x , el término del error es

$$E_n(x) = \frac{f^{(n+1)}(\xi)(x-c)^{(n+1)}}{(n+1)!}$$

Para el caso particular cuando $c = 0$, la ecuación 1.3 se denomina una serie de MacLaurin para $f(x)$,

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(0)(x)^k}{k!} + E_n(x), \quad (1.4)$$

donde

$$E_n(x) = \frac{f^{(n+1)}(\xi)x^{n+1}}{(n+1)!}.$$

Por ejemplo, tenemos las siguientes series de Taylor (con $c = 0$) para algunas funciones importantes,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \quad (1.5)$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \quad (1.6)$$

$$\cos x = x - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots \quad (1.7)$$

Usaremos este teorema una y otra vez durante el curso. Su principal uso es para aproximar una función usando los primeros términos de la serie de Taylor. El teorema nos dice que la aproximación es tan buena como su último término, conocido como el término del error.

Hagamos la siguiente manipulación:

$$\begin{aligned} f(x) &= \sum_{k=0}^n \frac{f^{(k)}(c)(x-c)^k}{k!} + E_n(x) \\ f(x) - \sum_{k=0}^n \frac{f^{(k)}(c)(x-c)^k}{k!} &= E_n(x) \\ \left| f(x) - \sum_{k=0}^n \frac{f^{(k)}(c)(x-c)^k}{k!} \right| &= |E_n(x)| = \frac{|f^{(n+1)}(\xi)| |x^{(n+1)}|}{n+1!} \end{aligned} \quad (1.8)$$

El lado izquierdo de la ecuación 5.2 es la diferencia entre $f(x)$ y su aproximación con la serie de Taylor. Usaremos nuestro conocimiento sobre $f^{(n+1)}(\xi)$ en el intervalo $[a, b]$ para encontrar alguna constante M de forma que

$$\left| f(x) - \sum_{k=0}^n \frac{f^{(k)}(c)(x-c)^k}{k!} \right| = |E_n(x)| \leq M |(x-c)|^{(n+1)}. \quad (1.9)$$

Ejemplo 1.2 Encuentre una aproximación para $f(x) = \sin x$, expandido alrededor de $c = 0$, usando $n = 3$.

Solución: Encontrar las derivadas $f^{(k)}$ para esta función es relativamente sencillo. Se encuentra entonces que

$$\begin{aligned} f(x) = \sin x &= \sin(0) + \frac{\cos(0)}{1!}x - \frac{\sin(0)}{2!}x^2 - \frac{\cos(0)}{3!}x^3 + \frac{\sin(\xi)}{4!}x^4 \\ &= x - \frac{x^3}{6} + \frac{\sin(\xi)}{24}x^4. \end{aligned}$$

En la ecuación anterior para $\sin(x)$ se tiene que los primeros dos términos del lado derecho son un polinomio de orden tres que aproxima la función cuando x esta cerca de 0. El último término es el término del error, nos dice como este polinomio difiere de $\sin(x)$.

Con el conocimiento que tenemos de la función $\sin(x)$, en particular que $|\sin(\xi)| \leq 1$, se puede entonces estimar el valor del error.

$$\left| \sin(x) - \left(x - \frac{x^3}{6} \right) \right| = \left| \frac{\sin(\xi)x^4}{24} \right| \leq \frac{x^4}{24}.$$

El caso especial cuando $n = 0$ del Teorema de Taylor es frecuentemente usado en el desarrollo de métodos numéricos. Se conoce como el *Teorema del valor medio*.

1.2. Análisis de error

En la práctica del análisis numérico es importante tener en cuenta que las soluciones computadas no son soluciones matemáticas exactas. La precisión de las soluciones numéricas se puede ver afectada de maneras sutiles por varios factores. El entendimiento de estas dificultades puede a menudo guiarnos en la implementación adecuada y/o en el desarrollo de mejores algoritmos numéricos.

Definición 1.4 Suponga que \hat{p} es una aproximación de p . El **error absoluto** es $E_p = |p - \hat{p}|$, y el **error relativo** es $R_p = |p - \hat{p}|/|p|$, dado que $p \neq 0$.

Definición 1.5 Se dice que el número \hat{p} aproxima a p con d cifras significativas (dígitos significativos) si d es el entero positivo más grande para el cual

$$\frac{|p - \hat{p}|}{|p|} < \frac{10^{-d}}{2}$$

Ejemplo 1.3 Determine el número de cifras significativas para la siguientes aproximaciones.

- Si $x = 3,141592$ y $\hat{x} = 3,14$, entonces $|x - \hat{x}|/|x| = 0,000507 < 10^{-2}/2$. Por lo tanto, \hat{x} aproxima a x con dos cifras significativas.
- Si $y = 1000000$ y $\hat{y} = 999996$, entonces $|y - \hat{y}|/|y| = 0,000004 < 10^{-5}/2$. Por lo tanto, \hat{y} aproxima a y con cinco cifras significativas.
- Si $z = 0,000012$ y $\hat{z} = 0,000009$, entonces $|z - \hat{z}|/|z| = 0,25 < 10^{-0}/2$. Por lo tanto, \hat{z} aproxima a z sin ninguna cifras significativa.

1.2.1. Pérdida de significancia

En términos generales, un computador guarda un número x en memoria como una mantisa y un exponente, por ejemplo $x = \pm r \times 10^k$, donde r es un número racional $[0,1, 1)$ de un número dado de dígitos (dependiente de la arquitectura del hardware), y k es un entero dentro de un cierto rango.

El número de dígitos significativos en r usualmente es determinado por la entrada del usuario. Las operaciones ejecutadas sobre los números guardados en la memoria del computador siguen la regla que la precisión no se puede ganar pero si se puede perder. Por ejemplo, si se suman dos números, 0,171717 y 0,51, entonces el resultado tendrá solamente dos dígitos significativos; la precisión del primer número se pierde debido a la incertidumbre en el segundo.

El anterior ejemplo no muestra nada nuevo y muestra la pérdida de presión a la que estamos acostumbrados en cualquier tipo de análisis. Sin embargo, se puede también incurrir en la pérdida de significancia si se restan dos números casi iguales. Por ejemplo, si se resta 0,177241 de 0,177589, el resultado $0,348 \times 10^{-3}$ tiene tan solo 3 dígitos significativos y los otros tres se han perdido. Este tipo de pérdida de precisión se conoce con el nombre de ‘cancelación por resta’ y puede ser evitada re-escribiendo la expresión. Esto será demostrado en los ejemplos.

También se puede incurrir en errores cuando dos cantidades de magnitudes radicalmente diferentes son sumadas. Por ejemplo, $0,1234 + 5,6789 \times 10^{-20}$ puede dar como resultado 0,1234 por un sistema con 16 dígitos significativos. Esto puede arrojar resultados inesperados.

Las estrategias para re-escribir expresiones que puedan presentar problemas son completar el cuadrado, factorización, o usar expansiones de Taylor como se ilustra en los siguientes ejemplos.

Ejemplo 1.4 Considere la precisión de $\sqrt{x+1} - 1$ cuando x se encuentra cerca de 0. Re-escriba la expresión para eliminar la cancelación por resta.

Solución: Suponga que $x = 1,2345678 \times 10^{-5}$. Entonces $\sqrt{x+1} \approx 1,000006173$. Si su computador o calculadora tiene tan solo 8 dígitos significativos entonces esta cantidad va a ser aproximada a 1,0000062. Cuando a esta cantidad se le resta 1, el resultado es $6,2 \times 10^{-6}$. De tal forma que se han perdido 6 dígitos significativos del número original.

Para solucionar este problema podemos *racionalizar* la expresión.

$$\sqrt{x+1} - 1 = \sqrt{x+1} - 1 \frac{\sqrt{x+1} + 1}{\sqrt{x+1} + 1} = \frac{x+1-1}{\sqrt{x+1} + 1} = \frac{x}{\sqrt{x+1} + 1}.$$

Esta expresión no contiene restas y por lo tanto no es susceptible a la cancelación por resta. Ahora, de nuevo para $x = 1,2345678 \times 10^{-5}$, esta expresión resulta en

$$\frac{1,2345678}{2,0000062} \approx 6,17281995 \times 10^{-6}$$

en una máquina con 8 dígitos de precisión no hay pérdida de precisión.

La mayoría de computadores y calculadoras modernas guardan resultados intermedios de los cálculos en formatos con mayor precisión. Esto minimiza pero no elimina los problemas expuestos en el ejemplo anterior. Es por esto que es importante tener esto en mente cuando se realicen soluciones a problemas de ingeniería.

Ejemplo 1.5 *Escriba un algoritmo que encuentre las raíces de la ecuación $x^2 + bx + c = 0$.*

Solución: La fórmula tradicional para solucionar la ecuación cuadrática es

$$x_{\pm} = \frac{-b \pm \sqrt{b^2 - 4c}}{2}$$

Si $b \gg c > 0$ entonces la expresión dentro de la raíz cuadrada puede ser aproximado a b^2 lo que daría como resultado dos raíces $x_+ = 0$ y $x_- = -b$. De estas dos x_- es casi la solución correcta, mientras x_+ no tiene ningún dígito significativo correcto.

Para corregir este problema, multiplicamos el numerador y denominador de x_+ por $-b - \sqrt{b^2 - 4c}$ obteniendo

$$x_+ = \frac{2c}{-b - \sqrt{b^2 - 4c}}$$

Ahora si $b \gg c > 0$, esta expresión da la raíz $x_+ = c/b$, lo cual es muy cercano al valor correcto. Esto lleva al par de raíces:

$$x_- = \frac{-b - \sqrt{b^2 - 4c}}{2}, \quad x_+ = \frac{2c}{-b - \sqrt{b^2 - 4c}}$$

Note que estas dos raíces son casi recíprocas, de tal forma que si se calcula x_- entonces x_+ puede ser calculada con tan solo dos cálculos adicionales. Este tipo de análisis va a ser muy útil en el desarrollo de algoritmos computacionales.

Ejemplo 1.6 *Re-escriba la expresión $e^x - \cos x$ para que no pierda precisión cuando x se encuentra cerca de 0.*

Solución: Cuando $x \rightarrow 0$ entonces $e^x \approx 1$, $\cos x \approx 1$ y al restar las dos cantidades entraríamos a tener cancelación por resta. Calculemos la expansión de Taylor de estas dos funciones:

$$\begin{aligned} e^x - \cos x &\approx \left[1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \dots \right] - \left[1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \right] \\ &= x + x^2 + \frac{x^3}{3!} + \mathcal{O}(x^5) \end{aligned}$$

Esta expresión no tiene restas y por lo tanto no está sujeta a cancelación por resta. Note que proponemos calcular $x + x^2 + x^3/6$ como una *aproximación* de $e^x - \cos x$. Como suponemos que x esta muy cerca de cero la aproximación es bastante buena. Si x esta *muy* cerca de cero es posible que con tan solo el primero o dos primeros términos sea suficiente. Si x no está tan cerca de cero es posible que necesitemos todos los términos propuestos (o incluso más). Si x esta lejos de cero debemos usar otra técnica.

1.2.2. Orden de Aproximación $\mathcal{O}(h^n)$

Claramente las secuencias $\{\frac{1}{n^2}\}_{n=1}^{\infty}$ y $(\frac{1}{n})_{n=1}^{\infty}$ convergen las dos a cero. Adicionalmente, se puede observar que la primera secuencia está convergiendo a cero mucho más rápido que la segunda secuencia. Nos interesa poder describir la velocidad a la que una secuencia converge (si converge) a un número dado.

Definición 1.6 Se dice que la función $f(h)$ es **O grande** de $g(h)$, denotados como $f(h) = \mathcal{O}(g(h))$, si existen las constantes C y c tal que

$$|f(h)| \leq C|g(h)| \quad \text{cuando } h \leq c. \quad (1.10)$$

Ejemplo 1.7 Considere las funciones $f(x) = x^2 + 1$ and $g(x) = x^3$. Como $x^2 \leq x^3$ y $1 \leq x^3$ para $x \geq 1$, entonces se tiene que $x^2 + 1 \leq 2x^3$ para $x \geq 1$. Entonces, $f(x) = \mathcal{O}(g(x))$.

La notación **O grande** es útil para describir la velocidad de crecimiento de una función en términos de funciones elementales bien conocidas (eg. x^n , $x^{1/n}$, a^x , $\log_a x$, etc.)

La velocidad de convergencia de una secuencia se puede describir de manera similar.

Definición 1.7 Suponga que $\{x_n\}_{n=1}^{\infty}$ y $\{y_n\}_{n=1}^{\infty}$ son dos secuencias. Se dice que la secuencia $\{x_n\}$ es de orden **O grande** de $\{y_n\}$, denotado $x_n = \mathcal{O}(y_n)$, si existen las constantes C y N tal que

$$|x_n| \leq C|y_n| \quad (1.11)$$

Ejemplo 1.8 $\frac{n^2-1}{n^3} = \mathcal{O}(\frac{1}{n})$, ya que $\frac{n^2-1}{n^3} \leq \frac{n^2}{n^3} = \frac{1}{n}$ siempre que $n \geq 1$.

A menudo una función $f(h)$ se aproxima con otra función $p(h)$ y se sabe que el error en esta aproximación está acotado por $M|h^n|$. Esto lleva a la siguiente definición.

Definición 1.8 Asuma que $f(h)$ se aproxima usando la función $p(h)$ y que existe una constante real $M > 0$ y un entero positivo n tal que

$$\frac{|f(h) - p(h)|}{|h^n|} < M \quad \text{para un } h \text{ suficientemente pequeño.} \quad (1.12)$$

Se dice entonces que $p(h)$ aproxima $f(h)$ con un orden de aproximación $\mathcal{O}(h^n)$ y escribimos

$$f(h) = p(h) + \mathcal{O}(h^n). \quad (1.13)$$

Note que cuando la relación 1.12 se escribe en la forma $|f(h) - p(h)| \leq M|h^n|$, se ve que la notación $\mathcal{O}(h^n)$ se usa en lugar del término del error $M|h^n|$. El siguiente teorema nos muestra como aplicar la definición a combinaciones simples de dos funciones.

Teorema 1.4 Asuma que $f(h) = p(h) + \mathcal{O}(h^n)$, $g(h) = q(h) + \mathcal{O}(h^m)$, y que $r = \min\{m, n\}$. Entonces,

$$f(h) + g(h) = p(h) + q(h) + \mathcal{O}(h^r), \quad (1.14)$$

$$f(h)g(h) = p(h)q(h) + \mathcal{O}(h^r) \quad (1.15)$$

y

$$\frac{f(h)}{g(h)} = \frac{p(h)}{q(h)} + \mathcal{O}(h^r) \quad \text{siempre y cuando } g(h) \neq 0 \text{ y } q(h) \neq 0 \quad (1.16)$$

En este momento es importante que empecemos a pensar en $p(x)$ como la aproximación polinomial de Taylor de orden n de la función $f(x)$; entonces el término del error se puede asignar simplemente a $\mathcal{O}(h^{n+1})$, lo que representa los términos omitidos comenzando con el término correspondiente a h^{n+1} . El término del error converge a cero a la misma velocidad que h^{n+1} converge a cero cuando h se aproxima a cero, lo cual se puede expresar por la relación

$$\mathcal{O}(h^{n+1}) \approx Mh^{n+1} \approx \frac{f^{(n+1)}(c)}{(n+1)!} h^{n+1} \quad (1.17)$$

para h suficientemente pequeñas. Entonces la notación $\mathcal{O}(h^{n+1})$ reemplaza la cantidad Mh^{n+1} .

Ejemplo 1.9 Considere las siguientes expansiones polinomiales de Taylor,

$$e^h = 1 + h + \frac{h^2}{2!} + \frac{h^3}{3!} + \mathcal{O}(h^4) \quad y \quad \cos(h) = 1 - \frac{h^2}{2!} + \frac{h^4}{4!} + \mathcal{O}(h^6)$$

Determine el orden de la aproximación de su suma y producto.

Para la suma tenemos

$$\begin{aligned} e^h + \cos(h) &= 1 + h + \frac{h^2}{2!} + \frac{h^3}{3!} + \mathcal{O}(h^4) + 1 - \frac{h^2}{2!} + \frac{h^4}{4!} + \mathcal{O}(h^6) \\ &= 2 + h + \frac{h^3}{3!} + \mathcal{O}(h^4) + \frac{h^4}{4!} + \mathcal{O}(h^6) \end{aligned}$$

Como $\mathcal{O}(h^4) + \frac{h^4}{4!} = \mathcal{O}(h^4)$ y $\mathcal{O}(h^4) + \mathcal{O}(h^6) = \mathcal{O}(h^4)$, esto se reduce a

$$e^h + \cos(h) = 2 + h + \frac{h^3}{3!} + \mathcal{O}(h^4),$$

y el orden de la aproximación es $\mathcal{O}(h^4)$. El producto se puede tratar de forma similar:

$$\begin{aligned} e^h \cos h &= \left(1 + h + \frac{h^2}{2!} + \frac{h^3}{3!} + \mathcal{O}(h^4)\right) \left(1 - \frac{h^2}{2!} + \frac{h^4}{4!} + \mathcal{O}(h^6)\right) \\ &= \left(1 + h + \frac{h^2}{2!} + \frac{h^3}{3!}\right) \left(1 - \frac{h^2}{2!} + \frac{h^4}{4!}\right) \\ &\quad + \left(1 + h + \frac{h^2}{2!} + \frac{h^3}{3!}\right) \mathcal{O}(h^6) + \left(1 - \frac{h^2}{2!} + \frac{h^4}{4!}\right) \mathcal{O}(h^4) \\ &\quad + \mathcal{O}(h^4)\mathcal{O}(h^6) \\ &= 1 + h - \frac{h^3}{3} - \frac{5h^4}{24} - \frac{h^5}{24} + \frac{h^6}{48} + \frac{h^7}{144} \\ &\quad + \mathcal{O}(h^4) + \mathcal{O}(h^6) + \mathcal{O}(h^4)\mathcal{O}(h^6) \end{aligned}$$

Como $\mathcal{O}(h^4)\mathcal{O}(h^6) = \mathcal{O}(h^{10})$ y

$$-\frac{5h^4}{24} - \frac{h^5}{24} + \frac{h^6}{48} + \frac{h^7}{144} + \mathcal{O}(h^4) + \mathcal{O}(h^6) + \mathcal{O}(h^{10}) = \mathcal{O}(h^4),$$

Entonces la ecuación previa se puede simplificar de siguiente manera

$$e^h \cos h = 1 + h - \frac{h^3}{3} + \mathcal{O}(h^4)$$

y el orden de la aproximación es $\mathcal{O}(h^4)$.

Ejercicio 1.1 Encuentre los números c , a los que se refiere el teorema del valor intermedio, para cada una de las funciones mostradas, para el intervalo y valor dado de L .

a. $f(x) = -x^2 + 2x + 3$ en el intervalo $[-1, 0]$ usando $L = 2$

b. $f(x) = \sqrt{x^2 - 5x - 2}$ en el intervalo $[6, 8]$ usando $L = 3$

Ejercicio 1.2 Encuentre el polinomio de Taylor de grado $n = 4$ expandido alrededor del valor dado x_0 , para cada una de las funciones mostradas a continuación.

a. $f(x) = \sqrt{x}$, $x_0 = 1$

b. $f(x) = x^5 + 4x^2 + 3x + 1$, $x_0 = 0$

c. $f(x) = \cos(x)$, $x_0 = 0$

Ejercicio 1.3 Encuentre el error E_x y el error relativo R_x . También determine el número de cifras significativas de la aproximación.

a. $x = 2,71828182$, $\hat{x} = 2,7182$

b. $x = 98350$, $\hat{x} = 98000$

c. $x = 0,000068$, $\hat{x} = 0,00006$

Ejercicio 1.4 Dadas las siguientes expansiones polinomiales de Taylor

$$e^h = 1 + h + \frac{h^2}{2!} + \frac{h^3}{3!} + \frac{h^4}{4!} + \mathcal{O}(h^5)$$

y

$$\sin(h) = h - \frac{h^3}{3!} + \mathcal{O}(h^5)$$

Determine el orden de la aproximación y la expresión de su suma y producto.

Ejercicio 1.5 Utilice el resultado del ejemplo 1.5 para desarrollar un algoritmo y un programa en MATLAB que le permita calcular con exactitud las raíces de una ecuación cuadrática en todas las situaciones, incluyendo los casos problemáticos cuando $|b| \approx \sqrt{b^2 - 4ac}$.

Capítulo 2

Introducción a la programación en MATLAB

2.1. Introducción

Esta es una guía introductoria al lenguaje de programación científica MATLAB. No pretende ser una revisión profunda y extensa sino tan solo una introducción muy básica. Se recomienda al estudiante que continúe sus estudios con la extensa y excelente ayuda incluida con MATLAB.

La presente guía recopila y está fuertemente inspirada en el libro “Mastering MATLAB 7” de Duane Hanselman y Bruce Littlefield. Para los estudiantes interesados, con este libro pueden profundizar y extender los temas cubiertos en esta guía. También puede servir como referencia para programadores avanzados.

2.2. Matemática simple

En MATLAB se pueden hacer simples cálculos como en cualquier calculadora. Supongamos que quiere calcular su nota definitiva de una materia y tiene las siguientes notas: quices (15 %) 4,0, proyectos (35 %) 4,5, parciales (40 %) 3,0 y tareas (10 %) 4,8. Para solucionar esto en MATLAB se podrían hacer los cálculos igual que en una calculadora:

```
>> 4.0*0.15 + 4.5*0.35 + 3.0*0.40 + 4.8*0.10
ans =
    3.8550
```

Note que MATLAB no tiene en cuenta los espacios y que las multiplicaciones preceden a las sumas. También note que MATLAB guarda la solución en una variable llamada **ans** por defecto. Ésta podrá ser usada en operaciones subsiguientes. Alternativamente, el problema podría haberse solucionado guardando la información en variables:

```
>> quices = 4.0
quices =
    4
>> proyectos = 4.5
proyectos =
    4.5
>> parciales = 3.0
parciales =
    3
>> tareas = 4.8;
>> notadefinitiva = quices*.15 + proyectos*.35 + parciales*.40 + tareas*.10
notadefinitiva =
    3.8550
```

En este ejemplo se crearon cuatro variables en MATLAB: ‘quices’, ‘proyectos’, ‘parciales’ y ‘tareas’ para guardar la nota correspondiente a cada rubro. Observe que MATLAB muestra el resultado después de cada operación excepto en el caso de ‘tareas’. El punto y coma al final de la línea le dice a MATLAB que evalúe pero no muestre el resultado. Observe que ahora el resultado final quedó guardado en la variable ‘notadefinitiva’ y no ‘ans’ como en el ejemplo inicial. Las variables en MATLAB solo pueden ser una palabra, por eso se usa la línea para unir las dos palabras.

Las variables usadas quedan guardadas en el *workspace* de MATLAB y podrán ser consultadas en cualquier momento. Por ejemplo escribiendo el nombre de la variable:

```
>> parciales
parciales =
    3
```

Si no recuerda los nombres de las variables que ha usado puede consultarlas usando el comando de MATLAB ‘who’.

Si se reutiliza una variable (por ejemplo ‘parciales’) o se asigna un nuevo valor, el valor previo es reemplazado y perdido. Sin embargo, cualquier otra expresión calculada con el valor previo no cambia. Considere el siguiente ejemplo:

```
>> ejes = 4;
>> rodamientos = 6;
>> engranajes = 2;
>> objetos = ejes + rodamientos + engranajes
objetos =
    12
>> ejes = 6
ejes =
    6
>> objetos
objetos =
    12
```

Aquí se calcula el número total de objetos entre ejes, rodamientos y engranajes. Después se cambia el número de ejes a seis, reemplazando el valor anterior de 4. El número total de objetos no ha cambiado. A diferencia de las hojas de cálculo, MATLAB no vuelve a calcular el número total de objetos basado en el nuevo número de ejes. *Cuando MATLAB hace un cálculo, lo hace usando los valores que conoce en el momento en el que el cálculo se hace.*

2.3. Scripts en archivos-M (M-Files)

Para problemas simples, escribir las instrucciones en la línea de comandos de MATLAB es simple y eficiente. Sin embargo, a medida que el número de comandos aumenta, o cuando se desea cambiar el valor de una o varias variables y volver a calcular el resultado, usar la línea de comando se vuelve tedioso. MATLAB provee una solución a este problema, permite escribir los comandos en un archivo de texto para después correrlos tal como si se hubieran escrito en la línea de comandos. Estos archivos de texto se llaman *scripts* o *archivos M*. El término ‘script’ significa que MATLAB simplemente lee las instrucciones línea por línea y las ejecuta. El término ‘archivo M’ se debe a que los archivos de texto que contienen el script deben tener la extensión ‘.m’, por ejemplo, `ejemplo1.m`.

Para crear un archivo M haga click en el icono con una página en blanco de la barra de herramientas de MATLAB o seleccione **New** del menú **File** y seleccione **M-file**. Este procedimiento abre un editor de texto donde usted puede escribir comandos de MATLAB. El siguiente script muestra los comandos del ejemplo anterior:

```
% script M-file ejemplo1.m
quices = 4.0;
proyectos = 4.5; %promedio de los proyectos.
parciales = 3.0
tareas = 4.8;
notadefinitiva = quices*0.15 + proyectos*0.35 + ...
                parciales*0.40 + tareas*0.10
```

Guarde este archivo como ‘ejemplo1.m’. Este archivo puede ser guardado en el disco y ejecutado inmediatamente de varias formas. (1) Escoja **Save and Run** del menú **Debug**, (2) presione el botón **Save and Run** en la barra de herramientas del *editor*, o (3) simplemente presione la tecla **F5**. Alternativamente, puede guardar este archivo seleccionando **Save** del menú **File** y después, en la línea de comandos de MATLAB, ejecutando el archivo escribiendo el nombre del archivo sin la extensión ‘.m’:

```
>> ejemplo1
parciales =
     3
notadefinitiva =
    3.8550
```

Note que el texto después de un % es un comentario y es ignorada por MATLAB. También observe que la regla de los punto y coma también se cumple en la ejecución de scripts.

Para calcular la nota de muchos estudiantes se podrían cambiar las notas repetidas veces en el script y ejecutarlo para calcular la nueva nota. Sin embargo sería más sencillo solicitar las notas al ejecutar el script sin tener que modificarlo. Supongamos que la única nota que cambia es la de parciales para cada estudiante y estudie el script a continuación:

```
% script M-file ejemplo2.m
quices = 4.0;
proyectos = 4.5; %promedio de los proyectos.
parciales = input('Introduzca la nota de parciales: ');
tareas = 4.8;
notadefinitiva = quices*0.15 + proyectos*0.35 + ...
                parciales*0.40 + tareas*0.10
```

Guarde este archivo como ‘ejemplo2.m’ y ejecútelo:

En respuesta a la solicitud la nota 3.0 fue entrada y la tecla **Return** o **Enter** fue presionada. El resto de los comandos se calcula como en el ejemplo anterior. Note que la entrada puede ser cualquier expresión, por ejemplo:

```
>> ejemplo2
Introduzca la nota de parciales: 3.0
notadefinitiva =
    3.8550

>> ejemplo2
Introduzca la nota de parciales: (4.0+3.0+2.0)/3
notadefinitiva =
    3.8550
```

2.4. Arreglos y operaciones sobre arreglos

Todos los cálculos considerados hasta ahora han involucrado números escalares. Cuando se desea efectuar la misma operación sobre más de un número a la vez, ejecutando la misma operación repetidamente sobre escalares es lento y engorroso. Para esto, MATLAB tiene definidas operaciones sobre arreglos.

Considere el problema de calcular la función seno sobre la mitad de su periodo, es decir entre $0 \leq x \leq \pi$. Como es imposible calcular el seno en todos los puntos en este rango, hay infinitos! Escogemos un número finito de puntos en los cuales evaluar la expresión. Para este ejemplo calcularemos la función seno cada $0,1\pi$.

Primero podemos construir un vector que contenga los puntos donde se va a calcular la expresión y después se calculará el seno para cada punto:

```
>> x = [0 .1*pi .2*pi .3*pi .4*pi .5*pi .6*pi .7*pi .8*pi .9*pi pi]
x =
    Columns 1 through 7
         0    0.3142    0.6283    0.9425    1.2566    1.5708    1.8850
    Columns 8 through 11
    2.1991    2.5133    2.8274    3.1416
>> y = sin(x)
y =
    Columns 1 through 7
         0    0.3090    0.5878    0.8090    0.9511    1.0000    0.9511
    Columns 8 through 11
    0.8090    0.5878    0.3090    0.0000
```

Para crear un arreglo en MATLAB basta con abrir un paréntesis cuadrado escribir los valor separados por un espacio, o una coma, y al terminar cerrar el paréntesis. Calcular el seno de cada uno de los términos es trivial. el comando `y = sin(x)` calcula el seno de cada uno de los elementos de 'x' y pone el resultado en la variable 'y'.

En el ejemplo anterior se creó un arreglo con una fila y 11 columnas; en lenguaje matemático, es un vector fila de longitud 11 o un arreglo de 1 por 11.

Los elementos del arreglo se pueden acceder de manera individual mediante subíndices. En

particular, $x(1)$ corresponde al primer elemento (0), $x(2)$ al segundo elemento (0.3142) y así sucesivamente. Por ejemplo,

```
>> x(3) % el tercer elemento de x
ans =
    0.6283
>> y(5) % el quinto elemento de y
ans =
    0.9511
```

También se pueden acceder varios elementos a la vez mediante la notación de los dos puntos:

```
>> x(1:5)
ans =
     0     0.3142     0.6283     0.9425     1.2566
```

Aquí se muestran los primeros cinco elementos del vector x . La notación 1:5 indica, comience con 1 y cuente hasta 5. También se puede usar:

```
>> x(7:end)
ans =
    1.885    2.1991    2.5133    2.8274    3.1416
```

La notación puede ser extendida para definir el paso que se debe dar al contar. En los ejemplos anteriores se usó el paso por defecto de 1, se habría podido escribir `7:1:end` y los resultados serían idénticos.

El índice de la mitad indica el paso a dar en al contar. Este índice puede tomar cualquier valor. En el ejemplo a continuación se listan el segundo, cuarto y quinto elemento del arreglo `x`. La notación `2:2:7` indica que se comience en el segundo elemento, aumente de 2 y se detenga cuando llegue a 7 (note que el elemento 7 no se ha incluido).

```
>> x(2:2:7)
ans =
    0.3142    0.9425    1.5708
```

El valor del paso también puede ser negativo, por ejemplo `11:-1:1` mostrará el vector completo en orden inverso. Pruébenlo!

Al principio de la sección creamos el arreglo ‘`x`’ escribiendo sus elementos entre paréntesis cuadrados separados por espacios. Esto creó un vector fila con los elementos del arreglo. Este método resultó eficiente para llenar un vector de 11 elementos, ¿pero que tal que quisiéramos construir un vector de 100 o 1000 elementos? Otro método para construir el vector ‘`x`’ sería:

```
>> x = (0:0.1:1)*pi
x =
Columns 1 through 7
    0    0.3142    0.6283    0.9425    1.2566    1.5708    1.8850
Columns 8 through 11
    2.1991    2.5133    2.8274    3.1416
```

El comando `(0:0.1:1)` crea un vector con el primer elemento 0 y el último 1 avanzando en pasos de 0.1. En este caso se multiplicó cada uno de los elementos del nuevo vector por π . Con este método se controla el elemento inicial, el final y el paso entre cada elemento, sin embargo no hay forma de controlar el número de elementos.

Si se desea tener un número determinado de elementos se puede usar la función de MATLAB **`linspace`**. Los argumentos de la función son:

```
linspace(primer_valor , ultimo_valor , numero_de_elementos)
```

Con esta función se puede generar el vector ‘`x`’ de la siguiente manera:

```
>> x = linspace(0,pi,11)
x =
Columns 1 through 7
    0    0.3142    0.6283    0.9425    1.2566    1.5708    1.8850
Columns 8 through 11
    2.1991    2.5133    2.8274    3.1416
```

En este momento podemos aprovechar para demostrar algunas de las capacidades gráficas de MATLAB. Ejecute los siguientes comandos:

```
>> plot(x,y)
>> title('Funcion_Seno')
>> xlabel('x')
>> ylabel('Seno_de_x')
```

2.5. Algebra lineal

En sus orígenes MATLAB era un sistema para solucionar sistemas lineales de ecuaciones. De hecho su nombre (MATrix LABoratory) refleja esto. Con el pasar de los tiempos MATLAB ha evolucionado para solucionar muchos problemas en todas las ramas del análisis. En esta sección exploramos esta funcionalidad de MATLAB.

Uno de los problemas más comunes en el álgebra lineal es encontrar la solución de un sistema lineal de ecuaciones. Por ejemplo consideren el conjunto de ecuaciones

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 366 \\ 804 \\ 351 \end{bmatrix}$$

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{y}$$

donde el producto punto (\cdot) es una operación entre una matriz y un vector. En MATLAB este producto se representa con un asterisco (*). Existen numerosos métodos para encontrar la solución del sistema lineal tales como eliminación Gaussiana, factorización LU o utilizando directamente la inversa de la matriz \mathbf{A} , \mathbf{A}^{-1} . La solución formal del sistema de ecuaciones está dado por:

$$\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$$

Para resolver el sistema de ecuaciones en MATLAB se deben ingresar primero la matriz A y el vector y.

```
>> A = [1 2 3; 4 5 6; 7 8 0]
A =
     1     2     3
     4     5     6
     7     8     0
>> y = [366;804;351]
y =
    366
    804
    351
```

Note que en este caso se usan punto y comas para ingresar varias filas. Con estos datos ingresados se pueden hacer cálculos sobre la matriz. Pruebe, por ejemplo, calcular el determinante de la matriz.

```
>> det(A)
ans = 27.000
```

¿Qué es el determinante de una matriz? ¿Por qué es importante al resolver un sistema lineal de ecuaciones? ahora que la matriz y el vector han sido ingresados se puede resolver fácilmente el sistema de ecuaciones.

```
>> x = inv(A)*y
x =
    25
    22
    99
```

Esta es la manera directa de solucionar el sistema de ecuaciones, sin embargo no es el más eficiente. MATLAB tiene definida la siguiente notación para resolver el sistema de ecuaciones que es mucho más eficiente que calcular explícitamente el inverso de la matriz. ¿Qué es el inverso de una matriz?

```
>> x = A\y
x =
    25
    22
    99
```

2.6. Control de flujo

2.6.1. Ciclos FOR

Los ciclos ‘*for*’ permiten repetir grupos de comandos por un numero de veces fijo y predeterminado. La forma general de los ciclos ‘*for*’ es

```
for x = arreglo
    (comandos)
end
```

Los (comandos) entre el **for** y el **end** se ejecutan una ver por cada columna en el arreglo.

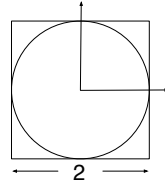
2.6.2. Estructura IF-ELSE-END

Muchas veces, una secuencia de comandos debe ser evaluada condicionalmente en la base de una prueba relacional. En los lenguajes computacionales, esta lógica está dada por alguna variación de la estructura IF-Else-End. La forma más simple de un If-Else-End es

```
if expresion
    (comandos)
elseif expresion2
    (otros comandos)
end
```

Ejercicio 2.1 *Método estadístico para calcular π*

Considere un círculo de radio unitario circunscrito por un cuadrado de lado 2 como se muestra en la figura. Suponga que se distribuyen al azar N puntos dentro del cuadrado. A medida que N tiende a infinito la probabilidad $P(N)$ de que un punto esté dentro del círculo es la relación entre el área del círculo y el área del cuadrado que puede ser calculada como $\pi/4$.



Escriba un programa que calcule el valor de π usando el método descrito en el párrafo anterior. Discuta la convergencia de los resultados con respecto al valor de N . Específicamente prepare una gráfica de $P(N) - \pi/4$ como función de N en escala log-log. Examine y discuta su gráfica. Elabore y discuta otras gráficas que considere necesarias.

Para resolver este ejercicio le puede ser útil la función **rand** de MATLAB. Esta función genera números pseudo-aleatorios a partir de una distribución aleatoria en el intervalo $[0, 1]$. Investigue esta función utilizando la ayuda de MATLAB (puede usar **help rand**). Por ejemplo, si necesita generar 100 valores en el intervalo $[a, b]$ puede usar la función $r = a + (b - a) \cdot \text{rand}(100, 1)$, donde r será un vector que contiene 100 números generados de forma pseudo-aleatoria.

¿Como determinar si el punto está dentro del círculo? Como el círculo puede ser descrito de forma analítica de la forma $f(x, y) = 0$ entonces el signo de $f(x, y)$ cambiará cuando se cruza la frontera del círculo de tal forma que se puede decidir si un punto está dentro o fuera del círculo simplemente inspeccionando el signo de la función. En concreto, para un círculo con radio a , definimos $f(x, y) = x^2 + y^2 - a^2$. Si $f(x_0, y_0) > 0$ entonces el punto (x_0, y_0) está por fuera del círculo, de lo contrario estará dentro del círculo. ¿Qué pasa si $f(x_0, y_0) = 0$? ¿Qué probabilidad hay de que esto ocurra?

Capítulo 3

Algebra lineal

Las operaciones matriciales son fundamentales en la mayoría de los métodos numéricos usados en ingeniería y su dominio es requisito fundamental en la computación numérica. No es exagerado decir que resolver la gran mayoría de problemas en ciencia e ingeniería se reduce, de alguna forma u otra, a encontrar la solución de sistemas lineales de ecuaciones usando métodos numéricos. En las etapas finales de la solución buscamos un vector \mathbf{x} que al multiplicarse por la matriz \mathbf{A} de como resultado el vector conocido \mathbf{b} ,

$$\mathbf{Ax} = \mathbf{b} \quad (3.1)$$

Las cantidades \mathbf{A} , \mathbf{b} y \mathbf{x} pueden tener una diversidad de significados físicos dependiendo del contexto matemático o físico de la ecuación 3.1.

En la siguiente sección se ilustra la manera en que un sistema lineal de ecuaciones resulta a partir de la solución de un problema práctico de ingeniería.

3.1. Distribución de temperatura en una placa

En esta sección se ilustra como un sistema de ecuaciones lineales resulta de la discretización de una ecuación diferencial parcial usando el método de las *diferencias finitas*. Este método será estudiado en detalle en el Capítulo 7, pero por ahora lo usaremos para demostrar la aparición de sistemas de ecuaciones lineales en la solución de problemas típicos de ingeniería.

Considere la distribución de temperatura en estado estable sobre una placa plana en dos dimensiones. La conservación de energía requiere que la distribución de temperatura, denotada por f , satisfaga la ecuación de Poisson en dos dimensiones

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = g(x, y) \quad (3.2)$$

donde $g(x, y)$ es una función que expresa la pérdida o producción local de energía a través de calor, el signo positivo representa pérdida (por ejemplo por radiación) y el signo negativo significa producción (por ejemplo por una resistencia eléctrica, reacción química, etc.).

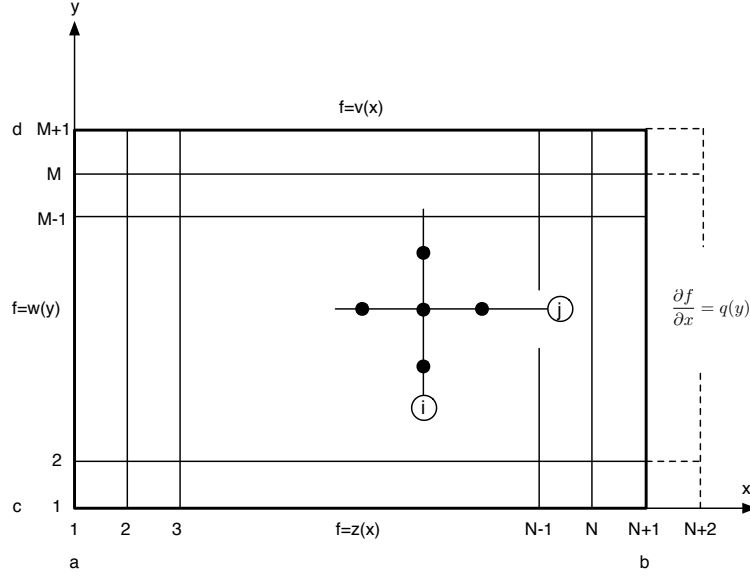


Figura 3.1: Una malla de diferencias finitas para resolver la ecuación de Poisson en un dominio rectangular.

Para completar la definición del problema, debemos especificar la geometría y condiciones de frontera. Para simplificar la implementación numérica, consideremos una placa con forma rectangular confinada entre $a \leq x \leq b$ y $c \leq y \leq d$ como se muestra en la figura 3.1. Como condiciones de frontera especificamos distribución de temperaturas a lo largo de la fronteras izquierda y horizontales de la placa; a lo largo de la frontera derecha especificamos un flujo de energía en forma de calor. Matemáticamente se expresa como *condición de frontera de Dirichlet*

$$f = w(y) \text{ en } x = a, \quad f = z(x) \text{ en } y = c, \quad f = v(x) \text{ en } y = d \quad (3.3)$$

y la *condición de frontera de Neumann*

$$\frac{\partial f}{\partial x} = q(y) \text{ en } x = b. \quad (3.4)$$

Por definición, la condición de frontera de Dirichlet da el valor de la función en la frontera, mientras que la condición de frontera de Neumann da el valor de la derivada normal.

Procedemos entonces a desarrollar un método de diferencias finitas sencillo. Comenzamos dividiendo el intervalo (a, b) en N subintervalos uniformes de ancho $\Delta x = (b - a)/N$ y dibujamos las líneas de malla $x = x_i$, donde

$$x_i = a + (i - 1)\Delta x \quad (3.5)$$

$i = 1, 2, \dots, N + 1$ como se muestra en la figura 3.1. De manera similar dividimos en intervalo (c, d) en M subintervalos uniformes de altura $\Delta y = (d - c)/M$ y dibujamos las líneas de malla $y = y_i$, donde

$$y_j = c + (j - 1)\Delta y \quad (3.6)$$

$j = 1, 2, \dots, M + 1$. Las intersecciones entre las líneas de malla verticales y horizontales definen los *puntos de la malla* o *nodos*. Es en estos nodos donde vamos a calcular de manera discreta los valores de la función. Por simplicidad, denotamos los valores de la función f en el nodo (i, j) como

$$f_{i,j} \equiv f(x_i, y_j) \quad (3.7)$$

La condición de frontera de Dirichlet ecuación 3.3 quedaría entonces en términos de valores discretos como

$$f_{1,j} = w(y_j), \quad f_{i,1} = z(x_i), \quad f_{i,M+1} = v(x_i) \quad (3.8)$$

Como estos valores son conocidos, nuestro objetivo es calcular los valores desconocidos de $f_{i,j}$ en los nodos $i = 1, 2, \dots, N + 1$, $j = 2, \dots, M$ para un total $K = N(M - 1)$ valores desconocidos.

Para construir el sistema de ecuaciones lineales para las incógnitas mencionadas aplicamos la ecuación diferencial 3.2 en cada nodo (i, j) y aproximamos la segunda derivada parcial usando diferencias finitas centrales,

$$\begin{aligned} (\partial^2 f / \partial x^2) &\approx (f_{i+1,j} - 2f_{i,j} + f_{i-1,j}) / \Delta x^2 \\ (\partial^2 f / \partial y^2) &\approx (f_{i,j+1} - 2f_{i,j} + f_{i,j-1}) / \Delta y^2 \end{aligned} \quad (3.9)$$

Como se discutirá a profundidad en el Capítulo 7, estas aproximaciones introducen un error numérico que es comparable al tamaño de Δx^2 o Δy^2 (el error es de orden Δx^2 o Δy^2). Recuerde además que este es tan solo uno de los errores que introducimos en nuestra solución numérica.

Las aproximaciones 3.9 nos permiten convertir la ecuación parcial diferencial 3.2 aplicada al nodo (i, j) a la ecuación de diferencias finitas

$$\frac{f_{i+1,j} - 2f_{i,j} + f_{i-1,j}}{\Delta x^2} + \frac{f_{i,j+1} - 2f_{i,j} + f_{i,j-1}}{\Delta y^2} = g_{i,j} \quad (3.10)$$

la cual puede ser organizada de la siguiente forma

$$f_{i+1,j} - 2(1 + \beta)f_{i,j} + f_{i-1,j} + \beta f_{i,j+1} + \beta f_{i,j-1} = \Delta x^2 g_{i,j} \quad (3.11)$$

para $i = 2, \dots, N$ y $j = 2, \dots, M$, donde

$$\beta \equiv \left(\frac{\Delta x}{\Delta y} \right)^2 \quad (3.12)$$

La ecuación 3.11 puede ser aplicada a los $L = (N - 1)(M - 1)$ nodos internos correspondientes a $i = 2, \dots, N$ y $j = 2, \dots, M$ pero no en los nodos de la frontera derecha ya que algunos de los nodos usados estarían por fuera de dominio de la solución. Entonces necesitamos producir $K - L = N(M - 1) - (N - 1)(M - 1) = M - 1$ ecuaciones adicionales.

Las ecuaciones faltantes deben involucrar la condición de frontera de Neumann sobre el borde derecho de la placa en $x = b$ que no ha sido utilizada aún. Una forma de implementar esta condición con un error comparable al tamaño de Δx^2 , el mismo debido a la discretización por diferencias finitas, es extender el dominio de la solución mas allá de la frontera física $x = b$ e introducir una *línea de malla fantasma* $x_{N+2} = b + \Delta x$ con valores asociados $f_{N+2,j}$.

A partir de la expansión de Taylor se puede mostrar que la primera derivada $(\partial f / \partial x)_{N+1,j}$ se puede aproximar con la fórmula de diferencias centrales $(f_{N+2,j} - f_{N,j}) / (2\Delta x)$ introduciendo un error numérico comparable al tamaño de Δx^2 . Adoptando esta aproximación podemos reemplazar la ecuación 3.4 con la ecuación de diferencias finitas $(f_{N+2,j} - f_{N,j}) / (2\Delta x) = q_j$ la cual podemos escribir como

$$f_{N+2,j} - f_{N,j} = 2\Delta x q_j \quad (3.13)$$

Habiendo introducido los nodos fantasma, podemos aplicar la ecuación 3.11 a los nodos de la frontera con $i = N + 1$ y $j = 2, \dots, M$. Usando la ecuación 3.13 para expresar $f_{N+2,j}$ en términos de $f_{N,j}$ encontramos

$$-2(1 + \beta)f_{N+1,j} + 2f_{N,j} + \beta f_{N+1,j+1} + \beta f_{N+1,j-1} = \Delta x^2 g_{N+1,j} - 2\Delta x q_j \quad (3.14)$$

Las ecuaciones 3.11 y 3.14 conforman el sistema de $N(M - 1)$ ecuaciones algebraicas para la misma cantidad de incógnitas. Para obtener la forma estándar dada en la ecuación 3.1 ubicamos los valores $f_{i,j}$, $i = 2, \dots, N + 1$, $j = 2, \dots, M$ en un vector de longitud $N(M - 1)$, \mathbf{s} , fila por fila comenzando desde abajo. Organizando las ecuaciones también fila por fila comenzando de abajo formamos el sistema lineal $\mathbf{A}\mathbf{s} = \mathbf{b}$. La matriz de coeficientes \mathbf{A} y el vector constante \mathbf{b} se muestran en las ecuaciones 3.15 a 3.21.

$$\begin{bmatrix} T & D & 0 & \dots & \dots & \dots & 0 \\ D & T & D & 0 & \dots & \dots & \dots \\ 0 & D & T & D & 0 & \dots & \dots \\ \dots & 0 & & \ddots & & 0 & \dots \\ \dots & \dots & \dots & 0 & D & T & D \\ 0 & \dots & \dots & \dots & 0 & D & T \end{bmatrix} \begin{bmatrix} f^{(2)} \\ f^{(3)} \\ \dots \\ \dots \\ \dots \\ f^{(M-1)} \\ f^{(M)} \end{bmatrix} = \begin{bmatrix} b^{(2)} \\ b^{(3)} \\ \dots \\ \dots \\ \dots \\ b^{(M-1)} \\ b^{(M)} \end{bmatrix} \quad (3.15)$$

donde,

$$T = \begin{bmatrix} -2(1 + \beta) & 1 & 0 & \dots & \dots & 0 \\ 1 & -2(1 + \beta) & 1 & 0 & \dots & 0 \\ 0 & 1 & -2(1 + \beta) & 1 & 0 & \dots \\ \dots & 0 & & \ddots & & 0 \\ \dots & \dots & 0 & 1 & -2(1 + \beta) & 1 \\ 0 & \dots & \dots & 0 & 2 & -2(1 + \beta) \end{bmatrix}, \quad (3.16)$$

$$D = \begin{bmatrix} \beta & 0 & \dots & \dots & 0 \\ 0 & \beta & 0 & \dots & \dots \\ \dots & 0 & \ddots & 0 & \dots \\ \dots & \dots & 0 & \beta & 0 \\ 0 & \dots & \dots & 0 & \beta \end{bmatrix}, \quad (3.17)$$

$$f^{(j)} = (f_{2,j}, f_{3,j}, \dots, f_{N+1,j})^T, \quad (3.18)$$

$$b^{(2)} = \begin{bmatrix} -w(y_2) - \beta z(x_2) + \Delta x^2 g_{2,2} \\ -\beta z(x_3) + \Delta x^2 g_{3,2} \\ \dots \\ -\beta z(x_N) + \Delta x^2 g_{N,2} \\ -2\Delta x q(y_2) - \beta z(x_{N+1}) + \Delta x^2 g_{N+1,2} \end{bmatrix}, \quad (3.19)$$

$$b^{(j)} = \begin{bmatrix} -w(y_j) + \Delta x^2 g_{2,j} \\ \Delta x^2 g_{3,j} \\ \dots \\ \Delta x^2 g_{N,j} \\ -2\Delta x q(y_j) + \Delta x^2 g_{N+1,j} \end{bmatrix} \text{ para } j = 3, \dots, M-1, \quad (3.20)$$

y

$$b^{(M)} = \begin{bmatrix} -w(y_M) - \beta z(x_2) + \Delta x^2 g_{2,M} \\ -\beta z(x_3) + \Delta x^2 g_{3,M} \\ \dots \\ -\beta z(x_N) + \Delta x^2 g_{N,M} \\ -2\Delta x q(y_M) - \beta z(x_{N+1}) + \Delta x^2 g_{N+1,M} \end{bmatrix} \quad (3.21)$$

Hemos estudiado como sistemas de ecuaciones lineales resultan de la discretización de ecuaciones parciales diferenciales usando el método de las diferencias finitas. En particular, se estudió que la distribución de temperatura en estado estable en una placa rectangular en dos dimensiones satisface la ecuación de Poisson en dos dimensiones,

$$\nabla^2 f \equiv \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = g(x, y) \quad (3.22)$$

donde ∇^2 es el operador Laplaciano en dos dimensiones y $g(x, y)$ es una función conocida de la posición que expresa producción o pérdida de energía en forma de calor. Un valor positivo de $g(x, y)$ significa pérdidas de energía, por ejemplo debido a la radiación, y un valor negativo significa producción, por ejemplo debido a una reacción química.

Estudie la derivación del sistema de ecuaciones lineales y realice los siguientes ejercicios.

Ejercicio 3.1 *Forma explícita del sistema de ecuaciones lineales*

Escriba de forma explícita el sistema lineal $\mathbf{Ax} = \mathbf{b}$ para el problema de la distribución de temperatura en la placa para $N=4$ y $M=4$.

Ejercicio 3.2 *Condición de frontera de Neumann en todas las fronteras: un sistema singular.*

Derive un sistema de $(N + 1)(M + 1)$ ecuaciones lineales resultantes de la discretización por diferencias finitas de la ecuación de Poisson con condiciones de frontera de Neumann que especifica la derivada normal a la frontera de la función desconocida en todas las fronteras del dominio rectangular. Verifique que la matriz de coeficientes es *singular*, y discuta las razones físicas. Recuerde que en clase consideramos un sistema con condiciones de frontera de Dirichlet (temperatura definida) en tres de los cuatro lados del rectángulo y condición de Neumann en el cuarto lado del rectángulo.

Ejercicio 3.3 Considere el flujo a través de un conducto con sección transversal rectangular como se muestra en la figura 3.2. La distribución de velocidad en una sección transversal del conducto se puede determinar usando la ecuación de Poisson en dos dimensiones,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{1}{\mu} \frac{\partial P}{\partial z} \quad (3.23)$$

donde μ es la viscosidad del fluido y $\partial P / \partial z$ es el gradiente de presión en la dirección del flujo y es una constante conocida.

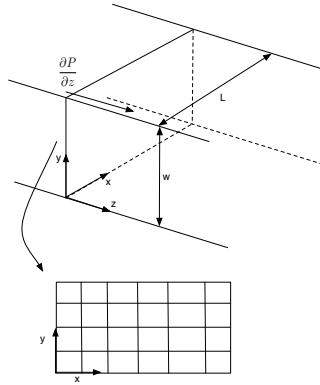


Figura 3.2: Conducto con sección transversal rectangular.

Haciendo mínimas modificaciones a su función de diferencias finitas para calcular la distribución de temperatura sobre una placa plana, usted puede implementar la solución a la ecuación 3.23.

Para un flujo viscoso se debe satisfacer la condición de no deslizamiento, esta dice que si un fluido se encuentra en contacto con una superficie este se mueve a la misma velocidad que la superficie. ¿A qué velocidad se mueve el fluido en las fronteras? ¿Qué tipo de condiciones de frontera tiene su problema?

A medida que la relación L/W aumenta los efectos de borde se vuelven pequeños en el centro del conducto y el flujo allí se aproxima al flujo entre dos placas planas. El flujo entre dos placas planas infinitamente largas se denomina flujo de Couette. La velocidad entre las placas satisface por la ecuación diferencial:

$$\frac{\partial^2 u}{\partial y^2} = \frac{1}{\mu} \frac{\partial P}{\partial z} \quad (3.24)$$

Resuelva esta ecuación analíticamente exprese la velocidad entre las placas como función de la posición y .

Establezca para que valor de L/W la solución analítica al flujo de Couette es una buena aproximación al perfil de velocidades en el centro del conducto. ¿Cómo afecta la viscosidad este resultado?

3.2. Métodos Iterativos

Hasta el momento hemos implementado modelos computacionales en los cuales se encuentra un sistema de ecuaciones lineales que aproximadamente describen el problema bajo consideración y hemos resuelto el sistema de ecuaciones lineales de forma directa (invirtiendo la matriz y multiplicándola por el vector constante). Después de un número finito de operaciones computacionales la solución del sistema se puede encontrar y si los cálculos se hicieran con una precisión infinita, la solución sería exacta.

Nos encontraremos con dificultades prácticas cuando el tamaño del sistema lineal es grande, es decir que contenga más de unos cientos de ecuaciones. Para un método directo el número de operaciones computacionales para resolver el sistema lineal es del orden de $\mathcal{O}(n^3)$ imagínese que pasa cuando usted tiene un sistema con cientos de miles de ecuaciones! Sistemas de estos tamaños (y mucho más grandes) resultan con frecuencia en la solución de problemas prácticos de ingeniería. Debido a esto ha sido necesario el desarrollo de *métodos iterativos* para la solución de estos problemas. En los métodos iterativos se generan soluciones aproximadas sucesivas basadas en algoritmos que emplean multiplicaciones vectoriales.

En un método iterativo típico se procede seleccionando un vector solución inicial $\mathbf{x}^{(0)}$ y después calculando la secuencia de vectores

$$\mathbf{x}^{(k+1)} = \mathbf{P}\mathbf{x}^{(k)} + \mathbf{c} \quad (3.25)$$

donde \mathbf{P} es una matriz de iteración adecuada y \mathbf{c} es un vector constante. \mathbf{P} y \mathbf{c} se diseñan en términos de la matriz de coeficientes \mathbf{A} y el vector constante \mathbf{b} de tal forma que la secuencia $\mathbf{x}^{(k)}$ converja, eso es, que su límite coincida con la solución del sistema lineal $\mathbf{A}\mathbf{x} = \mathbf{b}$.

3.2.1. Iteraciones de punto fijo para una ecuación

Comencemos estudiando una ecuación escalas que contenga una sola incógnita x ,

$$ax = b \quad (3.26)$$

Aplicando un algoritmo bastante conocido podemos encontrar la solución esperada en un solo paso, $x = b/a$.

Por razones que se harán aparentes muy pronto, escribimos la ecuación 3.26 en la forma $(a + n)x = nx + b$ donde n es un número real diferente de $-a$. Reorganizando la última ecuación obtenemos

$$x = px + c \quad (3.27)$$

donde hemos definido

$$p = n/(a + n), \quad c = b/(a + n) \quad (3.28)$$

Proponemos entonces calcular la solución de la siguiente manera iterativa:

1. Seleccionar un valor inicial para $x^{(0)}$.
2. Calcular una secuencia de números usando la fórmula de iteración:

$$x^{(k+1)} = px^{(k)} + c \quad (3.29)$$

Note que esta formula resulta directamente de la ecuación 3.27. Para encontrar $x^{(k+1)}$, multiplicamos $x^{(k)}$ por el coeficiente p y lo desplazamos c .

Si ocurriese que la secuencia 3.29 converge a un límite dado X , entonces por definición este límite satisface la ecuación

$$X = pX + c \quad (3.30)$$

y por lo tanto la ecuación original 3.26, de tal forma que X es la solución deseada.

El número X se denomina el punto fijo de la secuencia expresada por la ecuación 3.29 correspondiente a la función de iteración lineal

$$f(x) = px + c \quad (3.31)$$

en otras palabras, X es el punto donde $X = f(x)$.

3.2.2. Iteraciones de punto fijo para un sistema de ecuaciones

Tomando las ideas de la sección anterior un paso más allá, consideramos ahora el sistema de ecuaciones lineales

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (3.32)$$

y escribámosla en la forma equivalente $(\mathbf{A} + \mathbf{N})\mathbf{x} = \mathbf{N}\mathbf{x} + \mathbf{b}$ donde \mathbf{N} es alguna matriz. Entonces se tiene

$$\mathbf{M}\mathbf{x} = \mathbf{N}\mathbf{x} + \mathbf{b} \quad (3.33)$$

donde hemos definido

$$\mathbf{M} = \mathbf{A} + \mathbf{N}. \quad (3.34)$$

Se dice entonces que $\mathbf{A} = \mathbf{M} - \mathbf{N}$ es la separación de la matriz \mathbf{A} . Pre-multiplicando ambos lados de la ecuación 3.33 por \mathbf{M}^{-1} obtenemos

$$\mathbf{x} = \mathbf{P}\mathbf{x} + \mathbf{c} \quad (3.35)$$

donde

$$\mathbf{P} = \mathbf{M}^{-1}\mathbf{N}, \quad \mathbf{c} = \mathbf{M}^{-1}\mathbf{b} \quad (3.36)$$

El método iterativo procede de la siguiente manera:

1. Seleccionar un valor inicial para $\mathbf{x}^{(0)}$
2. Calcular una secuencia de vectores usando la formula de iteración:

$$\mathbf{x}^{(k+1)} = \mathbf{P}\mathbf{x}^{(k)} + \mathbf{c} \quad (3.37)$$

3.2.3. Métodos iterativos de Jacobi, Gauss-Seidel y SOR

Habiendo establecido nuestros cimientos teóricos procedemos a desarrollar funciones de iteración específicas y a discutir su rendimiento.

Método de Jacobi

En este método la matriz de iteración \mathbf{P} y el vector de corrimiento \mathbf{c} definidos en las ecuaciones 3.36 se construyen simplemente resolviendo las ecuaciones escalares individuales del sistema original $\mathbf{Ax} = \mathbf{b}$ para las incógnitas de la diagonal, obteniendo

$$x_i^{(k+1)} = \frac{1}{A_{i,i}} \left(b_i - \sum_{j=1, j \neq i}^N A_{i,j} x_j^{(k)} \right). \quad (3.38)$$

Haciendo referencia a la ecuación 3.34, la matriz \mathbf{M} es el componente diagonal de \mathbf{A} y la matriz \mathbf{N} es el negativo de \mathbf{A} con los elementos diagonales igualados a cero. La matriz de proyección está dada por

$$P_{i,j} = \begin{cases} -\frac{A_{i,j}}{A_{i,i}} & \text{para } i \neq j \\ 0 & \text{para } i = j \end{cases} \quad (3.39)$$

y el vector de desplazamiento esta dado por

$$c_i = -b_i/A_{i,i} \quad (3.40)$$

donde los índices repetidos no implican suma.

La programación del método de Jacobi se hace de forma directa con base a la ecuación 3.38. *Vale la pena anotar que la aproximación actual $x^{(k)}$ y la siguiente $x^{(k+1)}$ deben ser guardadas en dos vectores diferentes.*

Por ejemplo, para el sistema de 2×2 ,

$$\begin{aligned} 5x_1 + 3x_2 &= 8 \\ -x_1 + 4x_2 &= 3 \end{aligned} \quad (3.41)$$

obtenemos

$$\begin{aligned} x_1^{(k+1)} &= -\frac{3}{5}x_2^{(k)} + \frac{8}{5} \\ x_2^{(k+1)} &= \frac{1}{4}x_1^{(k)} + \frac{3}{4} \end{aligned} \quad (3.42)$$

Seleccionando $x_1^{(0)} = 0$ y $x_2^{(0)} = 0$ obtenemos $x_1^{(1)} = \frac{8}{5}$ y $x_2^{(1)} = \frac{3}{4}$ en la primera iteración. Continuando con las substituciones llegamos a los valores fijos $X_1 = 1$ y $X_2 = 1$, los cuales son la solución exacta al sistema 3.41.

La convergencia se puede verificar usando el criterio:

$$|\varepsilon_{a,i}| = \left| \frac{x_i^{(k)} - x_i^{(k-1)}}{x_i^{(k)}} \right| < \varepsilon_s \quad (3.43)$$

para todas las i , donde ϵ_s es algún valor definido por el usuario.

Las condiciones bajo las cuales la secuencia $\mathbf{x}^{(k)}$ convergerá a su punto fijo pueden ser complicadas y están relacionadas con el radio espectral de la matriz \mathbf{P} . Éste es difícil de calcular en la práctica. Alternativamente una condición suficiente más no necesaria de convergencia es que la matriz \mathbf{A} o su transpuesta sea diagonalmente dominante; esto es

$$|A_{i,i}| > \sum_{j=1, j \neq i}^N |A_{i,j}| \quad \text{or} \quad |A_{i,i}| > \sum_{j=1, j \neq i}^N |A_{j,i}| \quad (3.44)$$

para todos los valores de i . Es decir, el coeficiente diagonal en cada una de las ecuaciones debe ser mayor que la suma del valor absoluto de los otros coeficientes de la ecuación. Por fortuna, en la ingeniería, muchos problemas de importancia práctica satisfacen este requerimiento. Estudie la matriz de coeficientes resultante de la discretización de la ecuación de Poisson usando diferencias finitas. ¿Es diagonalmente dominante la matriz de coeficientes? Esta condición se satisface en el sistema 3.41. Si \mathbf{A} no es diagonalmente dominante, puede ser que reorganizando las ecuaciones o cambiando el nombre de las incógnitas se logre que lo sea, aunque no existe una forma sencilla y eficiente de programarlo.

Método de Gauss-Seidel

El método de Gauss-Seidel surge de una pequeña modificación del método de Jacobi. La característica distintiva es que los valores de los componentes individuales de $\mathbf{x}^{(k+1)}$ reemplazan los valores correspondientes de $\mathbf{x}^{(k)}$ tan pronto como estos son calculados. Por ejemplo, con referencia a las ecuaciones 3.42, seleccionamos $x_1^{(0)} = 0$ y $x_2^{(0)} = 0$ y calculamos $x_1^{(1)} = \frac{8}{5}$ y $x_2^{(1)} = (\frac{1}{4})(\frac{8}{5}) + \frac{3}{4} = \frac{23}{20}$.

La programación del método es un poco más simple que la del método de Jacobi ya que no es necesario guardar $\mathbf{x}^{(k)}$ y $\mathbf{x}^{(k+1)}$ en dos vectores diferentes.

El algoritmo de Gauss-Seidel se puede escribir explícitamente como

$$x_i^{(k+1)} = \frac{1}{A_{i,i}} \left(b_i - \sum_{j=1}^{i-1} A_{i,j} x_j^{(k+1)} - \sum_{j=i+1}^N A_{i,j} x_j^{(k)} \right). \quad (3.45)$$

En términos generales, cuando el método de Jacobi converge, el método de Gauss-Seidel también converge de manera aún más rápida.

Claramente existe un sesgo en la manera como se actualizan las variables desconocidas: la última incógnita utiliza la información más reciente, mientras la primera incógnita usa los valores viejos. Para prevenir este tratamiento injusto, podemos invertir el orden de actualización después de cada ciclo completo. Esta modificación resulta en el método de Gauss-Seidel simétrico.

Ejercicio 3.4 *Utilización de los métodos de Jacobi y Gauss-Seidel para la solución de sistemas lineales.*

Encuentre la solución del siguiente sistema de ecuaciones con una precisión de tres cifras significativas usando el método de Jacobi:

$$\begin{aligned}x + 9y &= 10 \\8x - y &= 7\end{aligned}$$

Repita la solución usando el método de Gauss-Seidel y compare la velocidad de convergencia. Para este ejercicio no es necesario programar nada, hágalo a mano.

Ejercicio 3.5 Programación de los métodos de Jacobi y Gauss-Seidel.

- Implemente un programa en algún lenguaje de alto nivel que resuelva un sistema de ecuaciones lineales usando el método de Jacobi con la opción de utilizar el método de Gauss-Seidel. Los cálculos deben terminar cuando la magnitud del cambio de cada valor escalar de la solución sea más pequeño que un valor específico.
- Use el programa desarrollado en la parte (a) para resolver el siguiente sistema de ecuaciones lineales con una precisión de cinco cifras significativas usando los dos métodos:

$$\begin{aligned}0,2x_1 + 0,2x_2 + 12,0x_3 + x_4 &= 1,0 \\0,3x_1 + 15,0x_2 + x_3 + 0,5x_4 &= 9,0 \\0,1x_1 + 0,4x_2 + 0,2x_3 + 19,0x_4 &= 5,0 \\10,0x_1 + 0,5x_2 + 0,2x_3 + x_4 &= 0,0\end{aligned}\tag{3.46}$$

Discuta la diferencia en la velocidad de convergencia.

Sobre relajación sucesiva (Successive Over-relaxation SOR)

Volvemos al algoritmo de Jacobi, ecuación 3.38, pero ahora la escribimos en la forma residual-corrección

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{r}^{(k)}\tag{3.47}$$

donde el residual o corrección $\mathbf{r}^{(k)}$ es

$$\mathbf{r}^{(k)} = (\mathbf{P} - \mathbf{I})\mathbf{x}^{(k)} + \mathbf{c}.\tag{3.48}$$

Si todos los componentes de $\mathbf{r}^{(k)}$ tienen el mismo signo en las iteraciones, la convergencia se acelerará mediante una sobre-corrección; por otro lado si los signos cambian la convergencia se acelerará mediante una sub-corrección. Estas consideraciones han motivado la modificación de las correcciones mediante la introducción de parámetro de relajación ω y estableciendo

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \omega\mathbf{r}^{(k)}\tag{3.49}$$

El algoritmo asociado es

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \frac{\omega}{A_{i,i}} \left(b_i - \sum_{j=1, j \neq i}^N A_{i,j}x_j^{(k)} \right).\tag{3.50}$$

Si las componentes de $\mathbf{x}^{(k)}$ se actualizan tan pronto como los nuevos valores son calculados, como se hizo en el método de Gauss-Seidel, el algoritmo es efectivamente el método de la sobre relajación sucesiva o SOR por el nombre del algoritmo en inglés (Successive Over-relaxation). En la práctica el método se implementa con base a la ecuación 3.50 pero los valores nuevos reemplazan los anteriores tan pronto son calculados. El algoritmo efectivo del método SOR es

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \frac{\omega}{A_{i,i}} \left(b_i - \sum_{j=1}^{i-1} A_{i,j}x_j^{(k+1)} - \sum_{j=i+1}^N A_{i,j}x_j^{(k)} \right). \quad (3.51)$$

La selección de ω se hace para optimizar la velocidad de convergencia. Se puede demostrar que una condición necesaria para que las sustituciones sucesivas converjan es que $0 < \omega < 2$.

Si $\omega = 1$ es igual a $1 - \omega = 0$ y el resultado es igual al método de Gauss-Seidel. Si a ω se le asigna un valor entre 0 y 1, el resultado es un promedio ponderado de los resultados actuales y anteriores. Este tipo de modificación se conoce como subrelajación. Se emplea comúnmente para hacer que un sistema no convergente, converja o apresure la convergencia al amortiguar sus oscilaciones.

Para valores de ω de 1 a 2, se le da una ponderación extra al valor actual. En este caso, hay una suposición implícita de que el nuevo valor se mueve en la dirección correcta hacia la solución verdadera, pero con una velocidad demasiado lenta. Por lo tanto, se pretende que la ponderación adicional de ω mejore la aproximación al llevarla más cerca de la verdadera. De aquí que este tipo de modificación, a la cual se le llama sobrerrelajación, permite acelerar la convergencia de un sistema que ya es convergente.

La elección de un valor adecuado de ω es específico al problema y se determina en forma empírica. Para la solución de un solo sistema de ecuaciones, con frecuencia es innecesaria. No obstante si el sistema bajo estudio se va a resolver de manera repetitiva, la eficiencia que se introduce por una prudente elección de ω puede ser de fundamental importancia. Buenos ejemplos de lo anterior son los sistemas muy grandes de ecuaciones diferenciales parciales, que frecuentemente se presentan cuando se modelan variaciones continuas de variables.

Ejercicio 3.6 Programación del método SOR.

Implemente un programa en algún lenguaje de alto nivel que resuelva un sistema de ecuaciones lineales usando el método de SOR. Los cálculos deben terminar cuando la magnitud del cambio de cada valor escalar de la solución sea más pequeño que un valor específico. Utilice este programa para resolver el sistema de ecuaciones del ejercicio 3.5, pruebe con diferentes valores de ω y encuentre para qué valor de ω la convergencia es más rápida. Discuta.

3.3. Solución de flujo potencial en dos dimensiones usando diferencias finitas

Para un flujo incompresible en dos dimensiones se tiene que las componentes de la velocidad en las coordenadas x y y , u y v respectivamente, deben satisfacer la ecuación de continuidad en dos

dimensiones

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \quad (3.52)$$

para satisfacer la conservación de masa. Esta ecuación se satisface idénticamente si definimos una función $\psi(x, y)$ tal que la ecuación 3.52 se convierta en

$$\frac{\partial}{\partial x} \left(\frac{\partial \psi}{\partial y} \right) + \frac{\partial}{\partial y} \left(-\frac{\partial \psi}{\partial x} \right) \equiv 0 \quad (3.53)$$

La comparación de las ecuaciones 3.52 y 3.53 muestra que esta nueva función $\psi(x, y)$ debe ser definida de tal forma que

$$u = \frac{\partial \psi}{\partial y} \quad y \quad v = -\frac{\partial \psi}{\partial x} \quad (3.54)$$

ψ se denomina función de corriente y tiene sentido matemático y físico.

¿Si tiene sentido lo que acabamos de hacer? Por supuesto, esto es tan solo un truco matemático para reemplazar dos variables (u y v) por una sola variable ψ . Sin embargo, más que un simple truco matemático, ψ , tiene una interpretación geométrica: Las líneas con ψ constante son *líneas de corriente* del flujo. Una línea de corriente es una línea que es tangencial a la velocidad en todos los puntos. Por definición no hay flujo a través de una línea de corriente. Otra interpretación relaciona ψ al caudal. En particular se puede demostrar que el caudal entre dos líneas de corriente es igual al cambio en la función de corriente ψ entre las dos líneas. Es decir

$$Q_{1 \rightarrow 2} = \psi_2 - \psi_1 \quad (3.55)$$

Más aún, la dirección del flujo se puede determinar mirando si ψ aumenta o disminuye. Si la función ψ aumenta hacia arriba el flujo es hacia la derecha, de lo contrario es hacia la izquierda.

Una aplicación importante de la función de corriente es en flujo no viscoso, incompresible e *irrotacional* en dos dimensiones. La condición de irrotacionalidad se reduce a la ecuación de Laplace para ψ

$$2\omega_z = 0 = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \frac{\partial}{\partial x} \left(-\frac{\partial \psi}{\partial x} \right) + \frac{\partial}{\partial y} \left(\frac{\partial \psi}{\partial y} \right) \quad (3.56)$$

donde ω_z es la vorticidad. Esto se reduce a

$$\nabla^2 \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = 0 \quad (3.57)$$

Las condiciones de frontera son velocidades conocidas en los flujos libres y cero flujo normal a cualquier superficie sólida:

$$\text{Flujo libre: } \frac{\partial \psi}{\partial x}, \frac{\partial \psi}{\partial y} \text{ conocidas} \quad (3.58)$$

$$\text{Superficie sólida: } \psi_{\text{superficie}} = \text{const} \quad (3.59)$$

La condición 3.59 es particularmente interesante ya que *cualquier* línea de ψ constante en un flujo puede ser interpretada como un cuerpo sólido!

Este análisis hace parte del estudio de flujos *potenciales*. Por ahora solo nos interesa saber que si resolvemos la ecuación 3.57 para una geometría dada y condiciones de frontera adecuadas podremos calcular la velocidad a partir de la ecuación 3.54.

Ejercicio 3.7 Resuelva el problema de flujo potencial en la expansión del conducto mostrado en la figura 3.3 usando el método de las diferencias finitas con un error de truncamiento de $\mathcal{O}(\Delta x^2)$ para discretizar las ecuaciones y el método SOR para resolver el sistema de ecuaciones lineales resultante. El flujo entra a una velocidad uniforme de 10 m/s, donde el conducto tiene 1 m de ancho y se asume que sale del conducto a velocidad uniforme de 5 m/s donde el conducto tiene 2 m de ancho. El conducto tiene una sección recta de 1 m de largo, una expansión de 45° y una sección recta final de 1 m de largo.

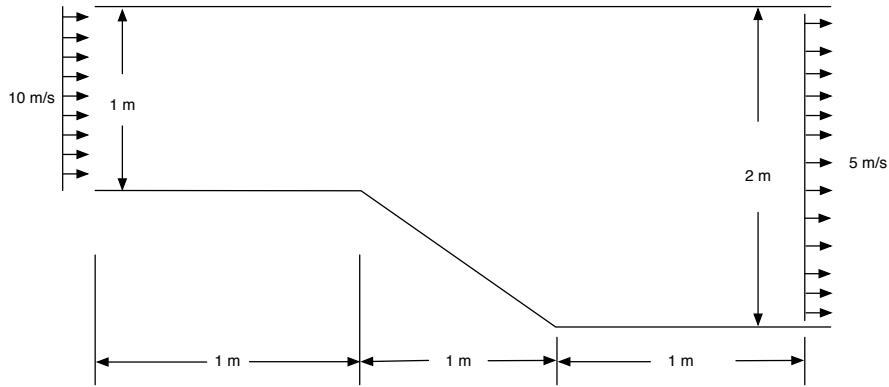


Figura 3.3: Conducto bidimensional con expansión a 45° .

Por conveniencia puede definir la función de corriente como cero a lo largo de la pared inferior. De esta forma como conocemos el caudal en el conducto, conocemos también el valor de la función de corriente en la pared superior. En las superficies de entrada y salida del conducto la función de corriente debe variar linealmente para dar velocidades uniformes. Todas estas condiciones de frontera deben ser entradas en su programa de solución al igual que un valor inicial para la velocidad en todos los nodos internos.

- Realice un estudio de la velocidad de convergencia como función del valor del parámetro de relajación ω y establezca un valor óptimo.
- Muestre algunas líneas de corriente a lo largo del conducto. Analice.
- Calcule los vectores de velocidad y muéstrellos en gráficas de vectores. ¿Qué se puede decir acerca de la aceleración del fluido?
- Podemos calcular la distribución de velocidad a lo largo de la pared. Si p_1 y V_1 son la presión y la velocidad a la entrada del conducto, las condiciones en cualquier otro punto pueden ser calculadas a partir de la ecuación de Bernoulli despreciando la gravedad.

$$p + \frac{1}{2}\rho V^2 = p_1 + \frac{1}{2}\rho V_1^2 \quad (3.60)$$

La presión se puede expresar en términos de un coeficiente de presión adimensional

$$C_p = \frac{p - p_1}{\frac{1}{2}\rho V_1^2} = 1 - \left(\frac{V}{V_1}\right)^2 \quad (3.61)$$

De esta forma C_p se puede determinar después de que V se ha determinado a partir de la función de corriente. Compare el valor del coeficiente de presión a lo largo de las paredes con la aproximación a partir de la ecuación de continuidad en una dimensión $V_1 A_1 \approx V(x) A(x)$

$$C_p(\text{una dimensión}) \approx 1 - \left(\frac{A_1}{A} \right)^2 \quad (3.62)$$

Capítulo 4

Solución de ecuaciones no lineales

Sistemas de ecuaciones no lineales resultan en el estudio teórico y práctico de una amplia gama de problemas de ingeniería. Su ocurrencia frecuente es un reflejo de la naturaleza predominantemente no-lineal de las leyes físicas. Comencemos con un ejemplo práctico donde el sistema está caracterizado por una ecuación no lineal.

4.1. Ecuación de estado de Redlich-Kwong en termodinámica

Una ecuación de estado relaciona el volumen v ocupado por 1 kmol de un gas con la presión p y la temperatura absoluta en Kelvin T del gas. Existen muchas ecuaciones de estado con varios niveles de precisión y sofisticación; la más simple es la ley de los gases ideales $pv = RT$. La constante R es la constante universal de los gases con un valor de 8314 kJ/(kmol · K).

Predicciones más precisas se logran con la utilización de la ecuación de estado de Redlich-Kwong [Van Wylen, Sonntag, Borgnakke, Fundamentals of Classical Thermodynamics, Fourth Ed., Wiley, pg. 464]:

$$p = \frac{RT}{v - b} - \frac{a}{v(v + b)\sqrt{T}} \quad (4.1)$$

donde a y b son dos constantes físicas relacionadas con la *temperatura crítica* T_c y con la *presión crítica* P_c a través de las ecuaciones

$$a = 0,42748 \frac{R^2 T_c^{2,5}}{P_c}, \quad b = 0,0867 \frac{RT_c}{P_c} \quad (4.2)$$

Cuando $a = b = 0$, la ecuación de Redlich-Kwong se reduce a la ecuación de los gases ideales. La temperatura crítica se define como la temperatura mínima arriba de la cual una sustancia o mezcla no puede existir en fase líquida sin importar que tan alta es la presión. La presión necesaria para licuar un fluido a la temperatura crítica es la presión crítica.

Un ingeniero se pone a la tarea de almacenar una determinada cantidad de hidrogeno en un tanque. Para diseñar dicho tanque, el ingeniero debe calcular el volumen a partir de la ecuación no-lineal 4.1 para valores específicos de la presión.

4.2. Raíces de ecuaciones

Desde hace años usted aprendió a usar la fórmula cuadrática:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (4.3)$$

para resolver

$$f(x) = ax^2 + bx + c = 0 \quad (4.4)$$

A los valores calculados con la ecuación 4.3 se les llama las “raíces” de la ecuación 4.4 que representan los valores de x que hacen a la ecuación 4.4 igual a cero. Por lo tanto, se define la raíz de una ecuación como el valor de x que hace $f(x) = 0$.

Aunque la fórmula cuadrática es útil para resolver la ecuación 4.4 existen muchas funciones donde las raíces no se pueden determinar tan fácilmente.

4.2.1. Método de la bisección

El método más simple es el de la bisección. El siguiente teorema garantiza el éxito del método:

Teorema 4.1 Teorema del valor medio. Si $f(x)$ es continua en $[a, b]$ entonces para cualquier y tal que y esté entre $f(a)$ y $f(b)$ existe un $c \in [a, b]$ tal que $f(c) = y$.

En particular, el teorema del valor medio nos dice que si $f(x)$ es continua y conoce a, b tal que $f(a), f(b)$ tienen signo diferente, entonces existe una raíz en $[a, b]$. Un estimativo decente para la raíz es $c = \frac{a+b}{2}$. Podemos verificar si $f(c) = 0$. Si esto no es cierto entonces una y solo una de las siguientes afirmaciones es cierta:

- (1) $f(a), f(c)$ tienen signos diferentes.
- (2) $f(c), f(b)$ tienen signos diferentes.

Ahora escogemos aplicar la bisección ya sea a $[a, c]$ o $[c, b]$, respectivamente, dependiendo de cual de estas afirmaciones es cierta. El método de la bisección procede de acuerdo a los siguientes pasos:

- (1) Encuentre dos valores x_1 y x_2 para los cuales $f(x_1)$ y $f(x_2)$ tienen diferentes signos.
- (2) Calcule $x_M = \frac{1}{2}(x_1 + x_2)$ y evalúe $f(x_M)$. El valor x_M es nuestro estimado actual de la raíz. (nota: Para minimizar el error de redondeo es mejor calcular $d = x_2 - x_1$ y definir $x_M = x_1 + \frac{1}{2}d$.)
- (3) Si $f(x_1)$ y $f(x_M)$ tienen signos diferentes, reemplace x_2 por x_M , de lo contrario, si $f(x_2)$ y $f(x_M)$ tienen signos diferentes, reemplace x_1 por x_M y regrese al paso 2. Si $f(x_M) = 0$ dentro de una tolerancia determinada, detenga los cálculos: x_M es la raíz buscada.

Convergencia

Para cada iteración del método de la bisección el intervalo $|b - a|$ es la mitad del tamaño del intervalo anterior. Formalmente llamaremos el primer intervalo $[a_0, b_0]$ y el primer punto medio c_0 . El segundo intervalo será $[a_1, b_1]$, etc. Para la iteración n tendremos

$$b_n - a_n = \frac{b_{n-1} - a_{n-1}}{2} \quad (4.5)$$

$$= \frac{b_0 - a_0}{2^n} \quad (4.6)$$

La convergencia del método es lineal, esto es, el tamaño del intervalo actual es igual al intervalo anterior por una constante.

Teorema 4.2 Teorema del método de la bisección. Si $f(x)$ es una función continua en $[a, b]$ tal que $f(a)f(b) < 0$, entonces después de n iteraciones el algoritmo de la bisección retornará c tal que $|c - c'| \leq \frac{b-a}{2^n}$ donde c' es alguna raíz de f .

Ejercicio 4.1 Considere el método de la bisección aplicado a encontrar el cero de la función $f(x) = x^2 - 5x + 3$ con $a_0 = 0$, $b_0 = 1$. ¿Cuáles son a_1 , b_1 ? ¿Cuáles son a_2 , b_2 ?

4.2.2. Método de Newton

El método de Newton es un método iterativo para encontrar raíces. Esto es, comenzando a partir de algún valor inicial, x_0 , una iteración del algoritmo produce un número x_1 que se supone estará más cerca de la raíz; valores x_2, x_3, \dots, x_n siguen de forma idéntica.

El método de Newton utiliza *linealización* para encontrar una raíz aproximada. Recordamos que el Teorema de Taylor no dice que

$$f(x+h) \approx f(x) + f'(x)h \quad (4.7)$$

Esta aproximación es mejor cuando $f''(x)$ se comporta bien entre x y $x+h$. El método de Newton trata de encontrar un h tal que

$$0 = f(x+h) \approx f(x) + f'(x)h$$

Esta ecuación se resuelve fácilmente así

$$h = \frac{-f(x)}{f'(x)}.$$

Una iteración del método de Newton toma un valor inicial x_k y devuelve x_{k+1} definido como

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad (4.8)$$

La utilización del método de Newton requiere que la función $f(x)$ sea diferenciable. Mas aún, la derivada de la función debe ser conocida. Como en el caso del método de la bisección, nuestro algoritmo no puede explícitamente verificar la continuidad de $f(x)$. Adicionalmente, el éxito del método depende del valor inicial x_0 . Este también era el caso para el método de la bisección, pero para este método la prueba de la validez del intervalo inicial era sencilla, por ejemplo $f(a_0)f(b_0) < 0$.

Nuestro algoritmo verificará que tan bueno es el estimado actual mirando el valor de $|f(x)|$. El algoritmo también revisará si la derivada es muy cercana a cero. Note que en el caso cuando $f'(x_k) = 0$ entonces h estará indefinido.

Algoritmo 1 Algoritmo para encontrar la raíz con el método de Newton

Entrada: f, f', x_0, N, tol

```
1:  $x \leftarrow x_0, n \leftarrow 0$ 
2: while  $n \leq N$ 
3:    $fx \leftarrow f(x)$ 
4:   if  $|fx| < tol$  then
5:     Salida:  $x$ 
6:   end if
7:    $fpx \leftarrow f'(x)$ 
8:   if  $|fpx| < tol$  then
9:     Warning " $f'(x)$  is muy pequeño, saliendo"
10:    Salida:  $x$ 
11:  end if
12:   $x \leftarrow x - fx/fpx$ 
13:   $n \leftarrow n + 1$ 
14: end while
```

Convergencia

La ecuación 4.8 se puede escribir en términos de una ecuación de iteración que producirá la secuencia de aproximaciones

$$g(x) = x - \frac{f(x)}{f'(x)} \quad (4.9)$$

donde $g(x)$ es la ecuación de iteración. Note que la ecuación $x = g(x)$ debe implicar que la ecuación $f(x) = 0$.

La convergencia del método puede ser investigado de una forma más rigurosa restando a ambos lados de la ecuación de iteración, $x^{k+1} = g(x^{(k)})$, el punto fijo X . (Note que este es una problema de punto fijo!) Introduciendo el error $e^{(k)} = x^{(k)} - X$ obtenemos,

$$e^{(k+1)} = g(x^{(k)}) - X \quad (4.10)$$

Procedemos asumiendo que el punto $x^{(k)}$ está lo suficientemente cerca de X , expandimos $g(x^k)$ usando una serie de Taylor alrededor del punto X , y notando que $g(X) = X$, para encontrar

$$e^{(k+1)} = g'(X)e^{(k)} + \frac{1}{2}g''(X)e^{(k)2} + \dots \quad (4.11)$$

Cuando la magnitud de $e^{(k)}$ es suficientemente pequeña, el segundo término y los términos subsiguientes (representados por los puntos) pueden ser despreciados, aproximando el comportamiento lineal

$$e^{(k+1)} \approx g'(X)e^{(k)} \quad (4.12)$$

Tomando la norma en ambos lados, encontramos que cada vez que hacemos una iteración, efectivamente multiplicamos la norma del error por el factor $|g'(X)|$. Dado esto las iteraciones convergerán dado que

$$|g'(X)| < 1. \quad (4.13)$$

Esta desigualdad es la única condición de convergencia del método numérico, dado que el estimado inicial esté lo suficientemente cerca de la raíz.

Cuando el criterio 4.12 se cumple, la velocidad de convergencia será al menos lineal, lo que significa que cuando la aproximación actual está lo suficientemente cerca de la raíz, hacer una nueva iteración causará que el error disminuya aproximadamente por un factor constante. Cuando la aproximación no está lo suficientemente cerca de la raíz, el análisis anterior no es válido y la condición de convergencia no puede ser establecida.

La ecuación 4.12 parece indicar que, cuando $g'(X = 0)$, el error $e^{(1)}$ será cero independientemente de la magnitud de $e^{(0)}$, esto es, independientemente de la aproximación inicial. En este caso, pareciera que podemos encontrar la solución en una sola iteración. El problema es que la ecuación 4.12 resultó de despreciar los términos de segundo orden y mayores en el lado derecho de 4.11. Manteniendo el término de mayor orden, obtenemos

$$e^{(k+1)} \approx \frac{1}{2}g''(X)e^{(k)2} \quad (4.14)$$

lo que demuestra que ahora la velocidad de convergencia es cuadrática: Cuando la aproximación actual está lo suficientemente cerca de la raíz, hacer una nueva iteración causa que el error sea elevado a la segunda potencia y después sea multiplicado por un factor constante $\frac{1}{2}g''(X)$.

La aplicación sucesiva de la fórmula 4.14 lleva a la expresión

$$e^{(k+1)} \approx \frac{1}{\alpha} \left(\alpha e^{(0)} \right)^{2^{k+1}} \quad (4.15)$$

donde

$$\alpha = \frac{1}{2}g''(X) \quad (4.16)$$

lo que muestra que la magnitud de $e^{(k+1)}$ tenderá a decrecer mientras que $|e^{(0)}| < 1/|\alpha|$. Con esto concluimos que la convergencia del método está garantizada mientras que la norma de $e^{(0)}$ sea lo suficientemente pequeña, esto es, mientras que la aproximación inicial esté lo suficientemente cerca de la solución.

De esta forma, si $e^{(0)}$ es de orden 10^{-1} y si $\frac{1}{2}g''(X)$ es de orden uno, entonces $e^{(1)}$ será de orden 10^{-2} , $e^{(2)}$ será de orden 10^{-4} y $e^{(3)}$ será de orden 10^{-8} . Precisión de ocho cifras significativas se alcanza en tan solo tres iteraciones.

Ahora queda por demostrar que el método de Newton tiene convergencia cuadrática. Para hacer esto debemos mostrar que $g'(X) = 0$. Sacando la derivada con respecto a x a ambos lados de la ecuación 4.9, encontramos

$$g'(x) = f(x) \frac{f''(x)}{f'(x)^2} \quad (4.17)$$

Por definición $f(X) = 0$, asumimos que $f'(X) \neq 0$ y encontramos que $g'(X) = 0$, lo que confirma el comportamiento cuadrático.

Procedemos entonces a calcular el coeficiente $g''(X)$ del lado derecho de la ecuación 4.14 en términos de la función f . Tomando la derivada con respecto a x a ambos lados de la ecuación 4.17, encontramos

$$g''(x) = \frac{f''(x)}{f'(x)} + f(x) \frac{f'''(x)}{f'(x)^2} - 2f(x) \frac{f''(x)^2}{f'(x)^3} \quad (4.18)$$

De nuevo, por definición $f'(X \neq 0)$, asumiendo que $f'(X) \neq 0$ encontramos $g''(X) = f''(X)/f'(X)$. Sustituyendo esta expresión en el lado derecho de la ecuación 4.14, encontramos que, cuando $f'(X) \neq 0$ el error evoluciona de acuerdo a la siguiente regla

$$e^{(k+1)} \approx \frac{1}{2} \frac{f''(X)}{f'(X)} e^{(k)^2} \quad (4.19)$$

Por lo tanto la convergencia es cuadrática.

Posibles problemas

Como se mencionó arriba, la convergencia del método depende de $f(x)$ y del estimado inicial x_0 . Una serie de problemas pueden resultar, ahora estudiaremos unos de estos problemas.

Ejemplo 4.1 Considere el método de Newton aplicado a la función $f(x) = x^j$ con $j > 1$ y con el estimado inicial $x_0 \neq 0$.

Solución: Note que $f(x) = 0$ tiene una sola raíz en $x = 0$. Ahora note que

$$x_{k+1} = x_k - \frac{k_k^j}{j x_k^{j-1}} = \left(1 - \frac{1}{j}\right) x_k$$

Como la ecuación tiene una sola raíz en $x = 0$ vemos que x_k está convergiendo a la raíz. Sin embargo, la convergencia no es cuadrática.

Ejemplo 4.2 Considere el método de Newton aplicado a la función $f(x) = \frac{\ln x}{x}$ con una aproximación inicial $x_0 = 3$.

Solución: Note que $f(x)$ es continua en \mathbb{R}^+ . Tiene una sola raíz en $x = 1$. Nuestro estimado inicial no está muy lejos de esta raíz. Sin embargo, considere la derivada

$$f'(x) = \frac{x^{\frac{1}{x}} - \ln x}{x^2} = \frac{1 - \ln x}{x^2}$$

Si $x > e^1$, entonces $1 - \ln x < 0$, y entonces $f'(x) < 0$. Sin embargo, para $x > 1$, sabemos que $f(x) > 0$. Entonces tomando

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)} > x^k$$

Los estimados sucesivos se ‘escaparán’ de la raíz $x = 1$.

Ejercicio 4.2 Aproxime $\sqrt{10}$ usando el método de Newton con un estimado inicial de $x_0 = 3$. Su respuesta debe tener 5 cifras decimales significativas.

Ejercicio 4.3 Use el método de Newton para encontrar una secuencia x_0, x_1, \dots tal que $x_k \rightarrow \ln 10$.

Ejercicio 4.4 Diseñe una subrutina que tan solo usando resta y multiplicación pueda encontrar la inversa multiplicativa de algún número entrado z , esto es $(1/z)$.

4.2.3. Método de la secante

No siempre es posible encontrar la derivada de la función $f(x)$, $f'(x)$. En una variación del método de Newton, aproximamos la derivada de la función $f(x)$, de la cual buscamos los ceros, con una diferencia finita que resulta de reemplazar la gráfica de $f(x)$ con una línea recta que pasa por dos puntos consecutivos (x^k, f^k) y (x^{k-1}, f^{k-1}) . Esto significa reemplazar la derivada $f'(x^{(k)}) = (f^{(k)} - f^{(k-1)})/(x^{(k)} - x^{(k-1)})$. La iteración se hará entonces con la fórmula de dos puntos

$$x^{(k+1)} = x^{(k)} - f^{(k)} \frac{x^{(k)} - x^{(k-1)}}{f^{(k)} - f^{(k-1)}} \quad (4.20)$$

Se deben hacer dos adivinanzas para empezar la iteraciones, pero solo *una* función se debe evaluar en cada paso.

Para investigar la convergencia del método, trabajamos como en la sección 4.2.2 y encontramos que

$$e^{(k+1)} = \frac{1}{2} \frac{f''(X)}{f'(X)} e^{(k)} e^{(k-1)} + \dots \quad (4.21)$$

Este resultado garantiza que el método convergerá dado que las dos aproximaciones iniciales estén lo suficientemente cerca de la raíz.

Para cuantificar la velocidad de convergencia, asumimos que, cuando estamos suficientemente cerca de la raíz, dos errores sucesivos están relacionados de acuerdo a la ley de potencia

$$\left| e^{(k+1)} \right| = \alpha \left| e^{(k)} \right|^p \quad (4.22)$$

donde α es una constante positiva y p es un exponente positivo que expresa el orden del método. Por ejemplo $p = 2$ significa que el método es cuadrático. Escribiendo

$$\left| e^{(k+1)} \right| = |\alpha|^{-1/p} \left| e^{(k)} \right|^{1/p} \quad (4.23)$$

y sustituyendo el lado derecho de las dos ecuaciones en la ecuación 4.21, encontramos que p satisface la ecuación cuadrática $p = 1 + 1/p$. La raíz con el valor absoluto más grande es $p = 1,62$.

Encontramos entonces que cada vez que llevamos a cabo una iteración, la magnitud del error se eleva aproximadamente a la 1,62 potencia y luego es multiplicado por un factor casi constante. El método de la secante convergerá para cualquier grupo de aproximaciones iniciales mientras estos sean razonables.

Ejercicio 4.5 *Derive la ecuación 4.21.*

Capítulo 5

Ajuste de curvas

Muchas aplicaciones de técnicas numéricas en ciencia y tecnología involucran el ajuste de datos experimentales. Por ejemplo, en 1601 el astrónomo alemán Johannes Kepler formuló la tercera ley del movimiento planetario, $T = Cx^{3/2}$, donde x es la distancia al sol medida en millones de kilómetros, T es el periodo medido en días, y C es una constante. Los datos medidos para los primeros cuatro planetas: Mercurio, Venus, Tierra y Marte, son (58, 88), (108,225), (150,365), y (228,687), y se muestran en la figura 5.1. Se puede encontrar la constante C mediante el ajuste por mínimos cuadrados a los datos. Con este método la constante se puede determinar como $C = 0,199769$.

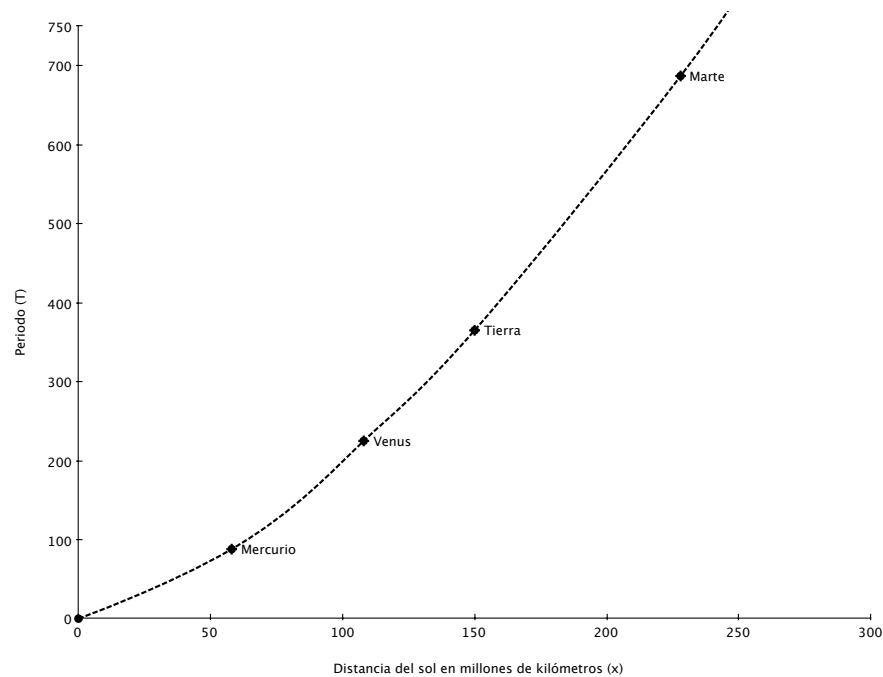


Figura 5.1: Datos de periodo como función de la distancia al sol para los primeros 4 planetas del sistema solar.

5.1. Regresión lineal por mínimos cuadrados

Procedimientos experimentales que producen una serie de puntos $(x_1, y_1), \dots, (x_N, y_N)$ son comunes en ciencia e ingeniería. Un objetivo de los métodos numéricos es el de determinar una fórmula $y = f(x)$ que relacione las variables x y y . Usualmente, se elige una clase de fórmulas posibles y los coeficientes son determinados. Hay muchos posibles tipos de funciones que se pueden usar. Algunas veces existe un modelo matemático, proveniente de una situación física, que determina la forma de la función. En esta sección se usará la clase de funciones lineales de la forma

$$y = f(x) = Ax + B. \quad (5.1)$$

En interpolación, se desea ajustar una función de tal forma que pase por todos los puntos, esto se estudiará en un capítulo posterior. Experimentalmente los datos contienen errores de diferente naturaleza. Seguir los puntos uno a uno puede dar una idea errónea del comportamiento de los datos. ¿Cómo ajustamos funciones se ajusten de la mejor manera posible a los datos, pero no necesariamente que pasen por todos ellos? Para contestar esta pregunta, debemos primero discutir el concepto de **error** (también conocido como **desviación** o **residual**):

$$e_k = f(x_k) - y_k \quad \text{para } 1 \leq k \leq N. \quad (5.2)$$

Existen varias formas para usar los residuales en la ecuación 5.2 para medir que tan lejos la curva $y = f(x)$ está de los datos. Por ejemplo, se podría usar el error máximo, el promedio del error o el error RMS ('root mean square'). Éste último es el más utilizado y se expresa como

$$E_2(f) = \left(\frac{1}{N} \sum_{k=1}^N |f(x_k) - y_k|^2 \right)^{1/2} \quad (5.3)$$

Minimizar la ecuación 5.3 se puede hacer minimizando la función

$$S^2 = \sum_{i=1}^n [y_i - y(x_i)]^2 \quad (5.4)$$

denominada la suma de cuadrados de las desviaciones. El método para encontrar una función, basado en la minimización de la ecuación 5.4 se denomina regresión por mínimos cuadrados.

5.1.1. Ajuste lineal por mínimos cuadrados

Suponga que se tiene un conjunto de N puntos $\{(x_k, y_k)\}_{k=1}^N$ donde todos los $\{x_k\}$ son distintos. La línea de **mínimos cuadrados** $y = f(x) = Ax + B$ es aquella que minimiza el error RMS de $E_2(f)$. Este error será mínimo si y solo si la cantidad $S^2 = \sum_{k=1}^N (Ax_k + B - y_k)^2$ es mínima. Se puede demostrar que los coeficientes A y B se pueden calcular mediante la solución del siguiente sistema lineal, conocido como las **ecuaciones normales**:

$$\begin{aligned} \left(\sum_{k=1}^N x_k^2 \right) A + \left(\sum_{k=1}^N x_k \right) B &= \sum_{k=1}^N x_k y_k, \\ \left(\sum_{k=1}^N x_k \right) A + NB &= \sum_{k=1}^N y_k \end{aligned} \quad (5.5)$$

Este sistema de ecuaciones puede ser resuelto mediante cualquiera de los métodos que estudiamos en el capítulo de solución de sistemas lineales. Sin embargo, el método utilizado en el programa que se muestra a continuación acomoda los datos de la forma que una matriz bien condicionada sea utilizada para la solución.


```

function [A,B]=lsline(X,Y)
% Entrada    - X es el vector de la abscisa
%            - Y es el vector de la ordenada
% Salida     - A es el coeficiente de x in Ax + B
%            - B es el coeficiente constante en Ax + B
xmean = mean(X);
ymean = mean(Y);
sumx2 = (X-xmean)*(X-xmean)';
sumxy = (Y-ymean)*(X-xmean)';
A = sumxy/sumx2;
B = ymean - A*xmean;

```

No vamos a demostrar la derivación de este algoritmo pero comprobaremos su funcionamiento.

5.1.2. Ajuste de ecuación de potencia $y = Ax^M$

Algunas situaciones involucran una función de la forma $y = Ax^M$, donde M es una constante conocida. Un ejemplo de esto es el movimiento planetario discutido al inicio del capítulo. En este caso solo existe un parámetro A que debe ser determinado.

Suponga que se tiene un conjunto de N puntos $\{(x_k, y_k)\}_{k=1}^N$ donde todos los $\{x_k\}$ son distintos. El coeficiente A para un ajuste de mínimos cuadrados a la curva de potencia $y = Ax^M$ está dado por

$$A = \frac{\left(\sum_{k=1}^N x_k^M y_k\right)}{\left(\sum_{k=1}^N x_k^{2M}\right)} \quad (5.6)$$

Ejercicio 5.1 Al examinar el comportamiento viscoso de un fluido es práctica común graficar la tasa de corte (gradiente de velocidad)

$$\frac{dv}{dy} = \dot{\gamma}$$

en la abscisa versus el esfuerzo cortante (τ) en las ordenadas. Cuando un fluido muestra una relación lineal entre esas dos variables, se denomina fluido Newtoniano, y la relación resultante es

$$\tau = \mu \dot{\gamma}$$

donde μ es la viscosidad del fluido. Ningún fluido es exactamente Newtoniano, aunque algunos fluidos, como por ejemplo el agua, se pueden aproximar con este comportamiento. Los fluidos que no se comportan de esa manera, se llaman no Newtonianos.

Un tipo de comportamiento no Newtoniano es el denominado como pseudoplástico y es típico de emulsiones o mezclas de líquidos inmiscibles. La relación entre τ y $\dot{\gamma}$ para un fluido pseudoplástico puede expresarse con la ecuación

$$\tau = \mu \dot{\gamma}^n$$

Los datos siguientes provienen de hidroxietilcelulosa en una solución de agua. Con el empleo de un ajuste por ley de potencias, encuentre los valores de μ y n .

Ejercicio 5.2 Problema Kepler.

$\dot{\gamma}, 1/s$	50	70	90	110	130
$\tau, N/m^2$	6.01	7.48	8.59	9.19	10.21

Verifique la constante calculada por Kepler mediante la implementación de un algoritmo para el cálculo de la constante para un ajuste a una función de potencia. ¿Cuál será el período para el quinto planeta Jupiter que se encuentra a 778 millones de kilómetros? Verifique con el valor real reportado en la literatura.

5.1.3. Cuantificación del error en la regresión lineal

Para calcular la *calidad* del ajuste a los datos, se utiliza un parámetro denominado el *coeficiente de correlación*, r :

$$r^2 = \frac{\text{variabilidad justificada alrededor de } y_m}{\text{variabilidad total alrededor de } y_m} \quad (5.7)$$

donde y_m es la media de los datos y_i :

$$y_m = \frac{1}{n} \sum_{i=1}^n y_i \quad (5.8)$$

La *variabilidad justificada* resulta de la variación de y con respecto a x

$$\sum_{i=1}^n [y(x_i) - y_m]^2 \quad (5.9)$$

La *variabilidad total* incluye también el error de precisión:

$$\sum_{i=1}^n (y_i - y_m)^2 = \dots = S^2 + \sum_{i=1}^n [y(x_i) - y_m]^2. \quad (5.10)$$

De tal forma que

$$r^2 = \frac{\sum [y(x_i) - y_m]^2}{S^2 + \sum [y(x_i) - y_m]^2} \quad (5.11)$$

Si asumimos que la suma cuadrada de las desviaciones S^2 es resultado solo de errores de precisión, entonces $S^2 \rightarrow 0$ y para un ajuste “perfecto” se tendría que $r^2 \rightarrow 1$. De esta forma, entre más cerca esté r de 1, “mejor” será el ajuste.

Desafortunadamente, cuando los datos se ven lineales, generalmente se obtiene que $r > 0,9$; el coeficiente de correlación no es un indicador muy sensible a la precisión de los datos.

En cambio consideremos el *error estándar* de los datos en y al rededor del ajuste, $s_{y/x}$:

$$s_{y/x} = \left(\frac{1}{(n-2)} \sum_{i=1}^n [y_i - y(x_i)]^2 \right)^{1/2} = \sqrt{\frac{S^2}{n-2}} \quad (5.12)$$

Esta cantidad aproxima la desviación estándar del error de precisión en y_i , esto es, si el error de precisión es pequeño también lo será $s_{y/x}$.

Mediante la introducción de la variación total de los datos, podemos hacer más obvio el significado de $s_{y/x}$, la variación total incluye tanto el error de precisión como la variación en y de los datos con x),

$$S_{yy}^2 = \sum_{i=1}^n (y_i - y_m)^2 \quad (5.13)$$

donde y_m es la media de los datos y_i :

$$y_m = \frac{1}{n} \sum_{i=1}^n y_i \quad (5.14)$$

Definiendo la variación total media como $s_{yy}^2 = S_{yy}^2/(n-1)$ y con un poco de álgebra se puede demostrar que,

$$\frac{s_{y/x}}{s_{yy}} = \left(\frac{n-1}{n-2} \right)^{1/2} (1-r^2)^{1/2} \quad (5.15)$$

La fracción de la mano izquierda de la ecuación 5.15 es la relación entre la desviación estándar de los datos al rededor de la línea y la variación total de los datos en y . Generalmente esta relación da una mejor idea de la calidad del ajuste. De esta forma se puede usar $(1-r^2)^{1/2}$ para caracterizar el ajuste.

Ejercicio 5.3 *Problema paracaidista.*

Se utiliza el método de la regresión lineal por mínimos cuadrados para resolver un problema de prueba de hipótesis relacionado con la caída de un paracaidista. Un modelo teórico matemático para la velocidad del paracaidista está dado por:

$$v(t) = \frac{gm}{c} \left(1 - e^{(-c/m)t} \right) \quad (5.16)$$

donde v = velocidad (m/s), g = constante gravitacional (9.8 m/s²), m = masa del paracaidista igual a 68.1 kg y c = coeficiente de arrastre de 12,5 kg/s. El modelo predice la velocidad del paracaidista en función del tiempo.

Un modelo empírico alternativo para la velocidad del paracaidista está dado por

$$v(t) = \frac{gm}{c} \left(\frac{t}{3,75 + t} \right) \quad (5.17)$$

Suponga que usted quiere probar y comparar la veracidad de esos dos modelos matemáticos. Esto se podría hacer al medir la velocidad real del paracaidista con valores conocidos de tiempo y al comparar estos resultados con las velocidades predichas de acuerdo con cada modelo.

Se implementó un programa para la recolección de datos experimentales, y los resultados se muestran en la tabla 5.2 mostrada a continuación. ¿Qué modelo predice de mejor manera el comportamiento medido experimentalmente?

5.1.4. Método de linealización de datos para $y = Ce^{Ax}$

Suponga que nos dan los datos $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$ y queremos ajustar una curva exponencial de la forma

$$y = Ce^{Ax} \quad (5.18)$$

Tiempo, s	v, m/s
1	10,00
2	16,30
3	23,00
4	27,50
5	31,00
6	35,60
7	39,00
8	41,50
9	42,90
10	45,00
11	46,00
12	45,50
13	46,00
14	49,00
15	50,00

Figura 5.2: Tabla de resultados experimentales de velocidad de caída como función del tiempo.

El primer paso consiste en sacar el logaritmo en ambos lados:

$$\ln(y) = Ax + \ln(C). \quad (5.19)$$

Después introducimos el cambio de variables:

$$Y = \ln(y), \quad X = x, \quad y \quad B = \ln(C) \quad (5.20)$$

De esta forma resultamos con la siguiente relación entre las nuevas variables X y Y :

$$Y = AX + B. \quad (5.21)$$

Los puntos originales (x_k, y_k) en el plano- xy se transforman en los puntos $(X_k, Y_k) = (x_k, \ln(y_k))$ en el plano- XY . Este proceso se denomina linearización. Una vez el sistema ha sido linearizado, la ecuación lineal 5.21 se ajusta a los puntos (X_k, Y_k) . Las ecuaciones normales para encontrar A y B son

$$\begin{aligned} \left(\sum_{k=1}^N X_k^2 \right) A + \left(\sum_{k=1}^N X_k \right) B &= \sum_{k=1}^N X_k Y_k, \\ \left(\sum_{k=1}^N X_k \right) A + NB &= \sum_{k=1}^N Y_k \end{aligned} \quad (5.22)$$

Una vez se han calculado los coeficiente A y B , el parámetro C en la ecuación 5.18 se puede calcular mediante la ecuación:

$$C = e^B. \quad (5.23)$$

Ejercicio 5.4 *Linearización de datos exponenciales.*

Use el método de linearización de datos para encontrar los coeficientes que den el mejor ajuste de la curva $y = Ce^{Ax}$ a los cinco datos $(0, 0,5), (1, 2,5), (2, 3,5), (3, 5,0)$, y, $(4, 7,5)$.

5.2. Regresión polinomial por mínimos cuadrados

En la sección 5.1 se desarrolló un procedimiento para obtener la ecuación de una línea recta por medio del criterio de mínimos cuadrados. En ingeniería, aunque algunos datos exhiben un comportamiento lineal marcado, son muchos los casos en los que los datos son pobremente representados por una línea recta, entonces, una curva podrá ser más adecuada para ajustarse a los datos. Como se analizó en la sección anterior, un método para lograr este objetivo es linealizando los datos. Otra alternativa es ajustar polinomios a los datos mediante *regresión polinomial*.

El procedimiento de mínimos cuadrados se puede extender fácilmente al ajuste de datos mediante polinomios de grado superior. Por ejemplo, suponga que ajustamos un polinomio de segundo grado o cuadrático:

$$y = a_0 + a_1x + a_2x^2 \quad (5.24)$$

En este caso la suma de los cuadrados de los residuos es:

$$S_r = \sum_{i=1}^n (y_i - a_0 - a_1x_i - a_2x_i^2)^2 \quad (5.25)$$

Al seguir el procedimiento de la sección anterior, obtenemos la derivada de la ecuación 5.25 con respecto a cada uno de los coeficientes desconocidos del polinomio,

$$\begin{aligned} \frac{\partial S_r}{\partial a_0} &= -2 \sum (y_i - a_0 - a_1x_i - a_2x_i^2) \\ \frac{\partial S_r}{\partial a_1} &= -2 \sum x_i (y_i - a_0 - a_1x_i - a_2x_i^2) \\ \frac{\partial S_r}{\partial a_2} &= -2 \sum x_i^2 (y_i - a_0 - a_1x_i - a_2x_i^2) \end{aligned}$$

Estas ecuaciones se igualan a cero y se reordenan para desarrollar el siguiente conjunto de ecuaciones normales:

$$\begin{aligned} (n)a_0 + (\sum x_i)a_1 + (\sum x_i^2)a_2 &= \sum y_i \\ (\sum x_i)a_0 + (\sum x_i^2)a_1 + (\sum x_i^3)a_2 &= \sum x_i y_i \\ (\sum x_i^2)a_0 + (\sum x_i^3)a_1 + (\sum x_i^4)a_2 &= \sum x_i^2 y_i \end{aligned} \quad (5.26)$$

donde todas las sumatorias van desde $i = 1$ hasta n . Observe que las tres ecuaciones anteriores son lineales y tienen tres incógnitas a_0 , a_1 y a_2 . Los coeficientes de las incógnitas se evalúan de manera directa, a partir de los datos observados.

En este caso, observamos que el problema de determinar un polinomio de segundo grado por mínimos cuadrados es equivalente a resolver un sistema de tres ecuaciones lineales simultáneas.

El caso anterior se puede extender con facilidad a un polinomio de m -ésimo grado. Se debe reconocer que la determinación de los coeficientes de un polinomio de m -ésimo grado es equivalente a resolver un sistema de $m+1$ ecuaciones simultáneas.

También es posible calcular un coeficiente de correlación usando la ecuación 5.7.

Ejercicio 5.5 Ajuste una ecuación cúbica a los datos siguientes:

x	3	4	5	7	8	9	11	12
y	1.6	3.6	4.4	3.4	2.2	2.8	3.8	4.6

Además de los coeficientes, determine r^2 y $s_{y/x}$.

5.3. Regresión lineal múltiple

Una extensión útil de la regresión lineal es el caso en el que y es una función lineal de dos o más variables independientes. Por ejemplo, y podría ser una función lineal de x_1 y x_2 , con en

$$y = a_0 + a_1x_1 + a_2x_2 \quad (5.27)$$

En particular tal ecuación es útil cuando se ajustan datos experimentales donde la variable objeto de estudio es una función de otras dos variables. En este caso bidimensional la “línea” de regresión se convierte en un “plano”.

Como en los casos anteriores, los coeficientes para el mejor ajuste se determinan al realizar la suma de los cuadrados de los residuos,

$$S_r = \sum_{i=1}^n (y_i - a_0 - a_1x_{1i} - a_2x_{2i})^2. \quad (5.28)$$

Derivando con respecto a cada uno de los coeficientes, igualando las ecuaciones resultantes a 0 y reordenando los términos, se obtiene el siguiente sistema de ecuaciones para los coeficientes:

$$\begin{bmatrix} n & \sum x_{1i} & \sum x_{2i} \\ \sum x_{1i} & \sum x_{1i}^2 & \sum x_{1i}x_{2i} \\ \sum x_{2i} & \sum x_{1i}x_{2i} & \sum x_{2i}^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_{1i}y_i \\ \sum x_{2i}y_i \end{bmatrix} \quad (5.29)$$

Ejercicio 5.6 *Las variables de diseño en ingeniería son a menudo dependientes de varias variables independientes. Por lo común, esta dependencia funcional se caracteriza mejor con ecuaciones de potencia multivariable. Una regresión lineal múltiple de datos transformados a logaritmos ofrece un recurso para evaluar tales relaciones.*

Por ejemplo, un estudio en ingeniería mecánica indica que el caudal de un líquido a través de una tubería está relacionado con el diámetro y la pendiente de una tubería según la relación

$$Q = a_0 D^{a_1} S^{a_2} \quad (5.30)$$

donde Q es el caudal, D es el diámetro de la tubería, S es la pendiente, y a_0, a_1 y a_2 son los coeficientes de ajuste.

Los datos experimentales se muestran en el cuadro 5.1. Use regresión lineal múltiple para analizar estos datos. Después con el modelo resultante prediga el flujo en una tubería de 2.5 ft de diámetro y con pendiente de 0.025 ft/ft.

Experimento	Diámetro, ft	Pendiente, ft/ft	Flujo, ft^3/s
1	1	0.001	1.4
2	2	0.001	8.3
3	3	0.001	24.2
4	1	0.01	4.7
5	2	0.01	28.9
6	3	0.01	84.0
7	1	0.05	11.1
8	2	0.05	69.0
9	3	0.05	200.0

Cuadro 5.1: Datos experimentales de diámetro, pendiente y flujo en una tubería circular de concreto.

Capítulo 6

Interpolación

6.1. Interpolación polinomial

Consideremos el problema de encontrar un polinomio que interpole un conjunto dado de datos:

$$\begin{array}{c|c|c|c|c} x & x_0 & x_1 & \dots & x_n \\ \hline y & y_0 & y_1 & \dots & y_n \end{array}$$

donde los x_i son todos distintos. Se dice que un polinomio $p(x)$ interpola estos datos si $p(x_i) = y_i$ para $i = 0, 1, \dots, n$. Los valores x_i se denominan los “nodos”.

Algunas veces es útil considerar una variable de este problema: tenemos una función desconocida (caja negra), $f(x)$, que queremos aproximar con un polinomio $p(x)$. Hacemos encontrando el polinomio que interpole los datos

$$\begin{array}{c|c|c|c|c} x & x_0 & x_1 & \dots & x_n \\ \hline f(x) & f(x_0) & f(x_1) & \dots & f(x_n) \end{array}$$

para alguna selección de nodos distintos x_i .

Un método estándar para interpolación de función procede reemplazando la función desconocida $f(x)$ con el polinomio de interpolación de grado N

$$P_N(x) = a_1x^N + a_2x^{N-1} + a_3x^{N-2} + \dots + a_Nx + a_{N+1} \quad (6.1)$$

que pase por los $N + 1$ datos, de tal forma que

$$f(x_i) = P_N(x_i) \quad (6.2)$$

para $i = 1, \dots, N + 1$. El problema se reduce a encontrar los $N + 1$ coeficientes a_1, a_2, \dots, a_{N+1} de tal forma que $P_n(x)$ satisfaga la restricción 6.2. Habiendo completado este cálculo podemos evaluar $P_N(x)$ para cualquier valor de x .

Puede pasar que $a_1 = 0$ en cuyo caso el orden del polinomio de interpolación se reduce de N a $N - 1$. Esto ocurrirá, por ejemplo, cuando la función $f(x)$ es un polinomio de grado menos que N . Si todos los $f(x_i)$ son iguales, entonces todos los coeficientes del polinomio serán cero excepto el último; $a_{N+1} = f(x_i)$ y el polinomio expresará una función constante.

6.1.1. Sistema de Vandermode

El método más simple de calcular los coeficientes del polinomio es forzando la restricción 6.2 en los $N + 1$ datos, derivando un sistema de $N + 1$ ecuaciones lineales algebraicas

$$\mathbf{V}^T \mathbf{a} = \mathbf{f} \quad (6.3)$$

donde

$$\mathbf{a} = (a_{N+1}, a_N, \dots, a_2, a_1)^T, \quad \mathbf{f} = (f(x_1), f(x_2), \dots, f(x_{N+1}))^T \quad (6.4)$$

\mathbf{V} es la matriz de *Vandermonde* de $(N + 1) \times (N + 1)$ con los elementos $V_{i,j} = x_j^{i-1}$, donde $i, j = 1, \dots, N + 1$. De forma explícita

$$\mathbf{V}^T = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{N-1} & x_1^N \\ 1 & x_2 & x_2^2 & \cdots & x_2^{N-1} & x_2^N \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_N & x_N^2 & \cdots & x_N^{N-1} & x_N^N \\ 1 & x_{N+1} & x_{N+1}^2 & \cdots & x_{N+1}^{N-1} & x_{N+1}^N \end{bmatrix} \quad (6.5)$$

Note que la abscisas x_i no tiene que ser ordenada en algún orden particular, por ejemplo, en orden de magnitud ascendente o descendiente.

Se puede demostrar que

$$\det \mathbf{V} = \prod_{j=1}^N \prod_{i>1} (x_i - x_j) \quad (6.6)$$

lo que garantiza que, siempre y cuando que los valores en la abscisa sean distintos, esto es $x_i \neq x_j$, la matriz de Vandermonde no es singular y la solución al sistema 6.5 es única.

Cuando $N = 1$ o 2 , la solución del sistema lineal 6.5 puede ser encontrada fácilmente en forma cerrada. Las fórmulas correspondientes serán derivadas más adelante. Para un mayor número de puntos, la solución deberá ser encontrada usando métodos numéricos.

Suponga ahora que mantenemos los puntos extremos x_1 y x_{N+1} fijos, y comenzamos a incrementar los valores de N desde su valor mínimo de 1 correspondiente a tan solo dos datos. Como la distancia entre puntos consecutivos se va reduciendo, los factores en la mano derecha de la ecuación 6.6 tienden a ser cada vez más pequeños. En la práctica, encontramos que el determinante de la matriz de Vandermonde se hace cada vez más pequeño, y por lo tanto el sistema lineal 6.3 se vuelve casi singular a valores relativamente pequeños de N . Este comportamiento limita la utilidad de este método.

6.1.2. Interpolación de Lagrange

El método de Lagrange produce un polinomio de interpolación de una manera en que se evita el cálculo explícito de los coeficientes del polinomio.

Polinomios de Lagrange

Para desarrollar este método, consideramos un conjunto de $N + 1$ datos e introducimos los i -ésimos polinomio de Lagrange de grado N , $l_{N,i}(x)$ que se define en términos de la abscisa de los

datos. Los polinomios $l_{N,i}(x)$ se construyen de tal forma que sean igual a cero en todos los puntos de datos, excepto en el i -ésimo punto donde es igual a uno, esto es,

$$l_{N,i}(x_j) = \delta_{i,j} \quad (6.7)$$

donde $\delta_{i,j}$ es el delta de Kronecker. Se puede fácilmente verificar que

$$l_{N,i}(x) = \frac{(x - x_1)(x - x_2) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_N)(x - x_{N+1})}{(x_i - x_1)(x_i - x_2) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_N)(x_i - x_{N+1})} \quad (6.8)$$

Polinomio de interpolación

El polinomio de interpolación está dado por la expansión lineal

$$P_N(x) = \sum_{i=1}^{N+1} f(x_i) l_{N,i}(x) \quad (6.9)$$

La ecuación 6.7 asegura que el lado derecho de la expresión 6.9 satisfaga las condiciones 6.2 y por lo tanto es el polinomio de interpolación deseado. Vale la pena anotar que hemos evitado la necesidad de resolver un sistema de ecuaciones.

Por ejemplo, el polinomio de interpolación correspondiente a los tres puntos $(0, 1)$, $(1, 2)$, $(2, 4)$ es

$$P_2(x) = 1 \frac{(x - 1)(x - 2)}{(0 - 1)(0 - 2)} + 2 \frac{(x - 0)(x - 2)}{(1 - 0)(1 - 2)} + 4 \frac{(x - 0)(x - 1)}{(2 - 0)(2 - 1)} = \frac{1}{2}(x^2 + x + 2) \quad (6.10)$$

El método de Lagrange tiene grandes ventajas: es fácil de programar, se comporta bien numéricamente con respecto al error de redondeo y no requiere solucionar un sistema de ecuaciones.

Una desventaja del método de Lagrange es que no calcula los coeficientes de manera explícita, fuera del último coeficiente, el cual puede ser calculado como $a_{N+1} = P_N(0)$. Otra desventaja es que si introducimos un dato adicional para mejorar la precisión, se hace necesario hacer todos los calcular nuevamente. Esta redundancia puede ser evitada mediante el uso de los métodos de interpolación de Neville, Aitken y Newton que serán discutidos a continuación. Sin embargo, a menos que se tengan gran cantidad de datos, esto no debería ser problemático y el método de Lagrange es recomendado para la mayoría de las aplicaciones.

Polinomio generador de Lagrange

Una expresión alternativa para los polinomios de Lagrange, la cual es útil para desarrollar la teoría de la interpolación, emerge mediante la introducción de un polinomio generados de grado $(N + 1)$

$$\Phi_{N+1}(x) = (x - x_1)(x - x_2) \cdots (x - x_N)(x - x_{N+1}) \quad (6.11)$$

Note que $\Phi_{N+1}(x)$ es cero en todos los datos. Con un poco de álgebra se puede demostrar que

$$l_{N,i}(x) = \frac{\Phi_{N+1}(x)}{(x - x_i)\Phi'_{N+1}(x)} \quad (6.12)$$

donde la prima representa derivación con respecto a x .

6.1.3. Interpolación de Neville

El algoritmo de Neville evalúa el polinomio de interpolación en una forma en la que fácilmente se puede incorporar un dato adicional. La idea básica es construir una secuencia de polinomios de orden cada vez mayor que ajuste un número de datos cada vez mayor. Para ilustrar esta lógica, hacemos una observación preliminar.

Sea $Q_{M-L}(x)$ el polinomio de grado $(M-L)$ cuya gráfica pasa por los $M-L+1$ datos,

$$x_L, x_{L+1}, \dots, x_{M-1}, x_M \quad (6.13)$$

los cuales no están distribuidos de ninguna manera en particular, y sea $S_{M-L}(x)$ el polinomio de interpolación de grado $(M-L)$ cuya gráfica para por los $M-L+1$ datos

$$x_{L+1}, \dots, x_{M-1}, x_M, x_{M+1} \quad (6.14)$$

Se puede verificar fácilmente que el polinomio de interpolación de grado $(M-L+1)$, $R_{M-L+1}(x)$ cuya gráfica pasa por todos los $M-L+2$ datos,

$$x_L, x_{L+1}, \dots, x_{M-1}, x_M, x_{M+1} \quad (6.15)$$

puede ser expresado como un promedio ponderado de los dos polinomios de orden menor $Q_{M-L}(x)$ y $S_{M-L}(x)$:

$$R_{M-L+1}(x) = \frac{(x_{M+1} - x)Q_{M-L}(x) + (x - x_L)S_{M-L}(x)}{x_{M+1} - x_L} \quad (6.16)$$

El lado derecho de esta ecuación puede ser escrito en la forma

$$R_{M-L+1}(x) = Q_{M-L}(x) + \frac{S_{M-L}(x) - Q_{M-L}(x)}{1 + \frac{x_{M+1} - x}{x - x_L}} \quad (6.17)$$

La piedra angular del algoritmo de Neville es que $Q_{M-L}(x)$ y $S_{M-L}(x)$ se pueden construir a partir de dos polinomios de grado $M-L-1$ y el proceso puede ser repetido para polinomios de grado descendiente hasta llegar a polinomios de grado cero, los cuales son sencillamente constantes. Los pasos son los siguientes:

- Comience considerando los $N+1$ polinomios de *grado cero* cuyas gráficas pasan por los $N+1$ datos individuales.
- Construya los polinomios de *grado uno* cuyas gráficas pasar por los dobletes de datos sucesivos.
- Construya los polinomio de *grado dos* cuyas gráficas pasen por los triplete de datos sucesivos.
- Continúe de esta manera hasta obtener el polinomio de interpolación $P_N(x)$.

Los polinomios intermedios pueden ser evaluados sistemáticamente actualizando N veces los componentes de un vector \mathbf{u} de dimensión $(N+1)$ como se muestra en el algoritmo 2, donde inicialmente $u_i = f(x_i)$. La matriz diagonal superior $A(i, m)$, donde el índice m cuenta el número de actualizaciones, forma lo que se conoce como la *tabla de Neville*. El valor requerido es $P_N(x) = A(1, N+1)$.

La primera fila de la tabla de Neville muestra sucesivamente mejores aproximaciones al valor de $f(x)$ calculado tomando cada vez un mayor número de datos. Por ejemplo, $A(1, 2)$ produce el valor $f(x)$ aproximando la gráfica de la función $f(x)$ mediante un línea recta que pasa por los primeros dos datos. Si un nuevo dato es introducido, el cálculo del nuevo valor de interpolación requiere tan solo valores a lo largo de la diagonal principal, y el resto de la tabla no necesita ser guardada.

Algoritmo 2 Evaluación del polinomio de interpolación usando el método de Neville.

```

do  $i = 1, N + 1$ 
   $u_i = f(x_i)$ 
   $A(i, 1) = u_i$  % Opcional, usado para construir la tabla de Neville
end do
do  $m = 2, N + 1$ 
  do  $i = 1, N - m + 2$ 
     $u_i \leftarrow u_i + \frac{u_{i+1} - u_i}{1 + \frac{x_{m+i-1} - x}{x - x_i}}$  ó  $u_i \leftarrow \frac{(x - x_i)u_{i+1} + (x_{m+i-1} - x)u_i}{x_{m+i-1} - x_i}$ 
     $A(i, m) = u_i$  % Opcional, usado para construir la tabla de Neville
  end do
end do
 $P_N(x) = u_1$ 

```

6.1.4. Interpolación de Newton y diferencias divididas

La interpolación de Newton es similar a la interpolación de Neville, en el sentido que la adición de un nuevo dato puede ser incluido con mínimo esfuerzo. Newton expreso el polinomio de interpolación en términos de una expansión lineal de una familia triangular de polinomios elementales en la forma

$$P_N(x) = \sum_{i=0}^N c_i \phi_i(x) \quad (6.18)$$

donde c_i son coeficientes constantes y $\phi_i(x)$ son polinomios de grado i definidos en términos de los datos de la siguiente manera

$$\begin{aligned}
\phi_0(x) &= 1 \\
\phi_1(x) &= x - x_1 \\
\phi_2(x) &= (x - x_1)(x - x_2) \\
&\dots \\
\phi_i(x) &= (x - x_1)(x - x_2) \dots (x - x_i) \\
&\dots \\
\phi_N(x) &= (x - x_1)(x - x_2) \dots (x - x_i) \dots (x - x_N)
\end{aligned} \quad (6.19)$$

Note que la abscisa del último dato x_{N+1} no aparece en estas expresiones.

En la práctica, $P_N(x)$ se calcula más eficientemente mediante el método de las multiplicaciones anidadas, escribiendo el polinomio como

$$P_N(x) = [(\dots \{c_N(x - x_N) + c_{N-1}\} \dots) c_3(x - x_3) - c_2](x - x_2) + c_1](x - x_1) + c_0 \quad (6.20)$$

y usando alguno de estos algoritmos

$$\begin{aligned}
P &= c_N \\
\text{Do } i &= N, 1, -1 \\
P &\leftarrow P(x - x_i) + c_{i-1} \\
\text{End Do} \\
P_N(x) &= P
\end{aligned} \quad (6.21)$$

o

$$\begin{aligned}
P &= c_0 \\
b &= 1 \text{ Do } i = 1, N \\
b &\leftarrow b(x - x_i) \\
P &\leftarrow P + c_i b \\
\text{End Do} \\
P_N(x) &= P
\end{aligned} \tag{6.22}$$

Los coeficientes c_i se pueden calcular mediante la construcción de la tabla de diferencias divididas:

$$\begin{array}{ccccccc}
f_1 & & & & & & \\
& f_{1,2} = (f_2 - f_1)/(x_2 - x_1) & & & & & \\
f_2 & & f_{1,2,3} = (f_{2,3} - f_{1,2})/(x_3 - x_1) & & & & \\
& f_{2,3} = (f_3 - f_2)/(x_3 - x_2) & & f_{1,2,3,4} = (f_{2,3,4} - f_{1,2,3})/(x_4 - x_1) & & & \\
f_3 & & f_{2,3,4} = (f_{3,4} - f_{2,3})/(x_4 - x_2) & & \dots & & \\
& f_{3,4} = (f_4 - f_3)/(x_4 - x_3) & & f_{2,3,4,5} = (f_{3,4,5} - f_{2,3,4})/(x_5 - x_2) & & & \\
f_4 & & f_{3,4,5} = (f_{4,5} - f_{3,4})/(x_5 - x_3) & & & & \\
\dots & \dots & \dots & \dots & \dots & &
\end{array} \tag{6.23}$$

Se puede demostrar mediante manipulaciones algebraicas que la línea diagonal superior no proporciona con los valores de c_i , esto es,

$$\begin{aligned}
c_0 &= f_1 \\
c_1 &= f_{1,2} \\
c_2 &= f_{1,2,3} \\
&\dots \\
c_N &= f_{1,2,3,4,\dots,N+1}
\end{aligned} \tag{6.24}$$

La adición de un dato adicional introduce una línea diagonal en la parte inferior de la tabla que no afecta el resto de la tabla.

Los valores de los coeficientes c_i pueden ser construidos de manera recursiva, actualizando N veces los componentes de un vector \mathbf{u} de dimensión $(N+1)$ siguiendo el siguiente algoritmo, donde, inicialmente $u_i = f(x_i)$. La matriz diagonal superior $\mathbf{A}(i, m)$ forma la tabla de diferencias divididas de Newton.

Ejercicio 6.1 *Todos los métodos de interpolación polinomial producen la misma respuesta.*

Use (a) el método de Vandermonde, (b) el método de la interpolación de Lagrange, (c) el algoritmo de Neville, y (d) el algoritmo de Newton para calcular el valor de la función $f(x)$ en $x = 0,25$ a partir de los datos $(0, 1.0)$, $(0.5, 2.0)$, $(1.0, 1.5)$, $(1.5, 1.0)$. Verifique que todos los métodos producen el mismo resultado.

Ejercicio 6.2 *Interpolación de tablas termodinámicas.*

La evaluación de funciones termodinámicas tales como el volumen molar, temperatura, presión y entropía es requerida para el diseño de varios procesos en ingeniería mecánica. Por ejemplo, el conocimiento de la temperatura de una cierta cantidad de gas que se comprime adiabáticamente

Algoritmo 3 Evaluación de los coeficientes c_i de la expansión de Newton 6.18 por el método de las diferencias divididas.

```

do  $i = 1, N + 1$ 
     $u_i = f(x_i)$ 
     $A(i, 1) = u_i$  % Opcional, usado para construir la tabla de Newton
end do
 $c_0 = u_1$ 
do  $m = 1, N$ 
    do  $i = 1, N - m + 1$ 
         $u_i \leftarrow \frac{u_{i+1} - u_i}{x_{i+m} - x_i}$ 
         $A(i, m + 1) = u_i$  % Opcional, usado para construir la tabla de Neville
    end do  $c_m = u_1$ 
end do

```

desde la presión atmosférica hasta alguna presión específica es necesario para el diseño de motores de combustión interna.

Muchas funciones termodinámicas han sido tabuladas en libros, textos y manuales. Por ejemplo, en Van Wylen, Sonntag y Borgnakke, *Fundamentals of classic thermodynamics*, Fourth Edition, Wiley, pg. 711–713 se da la presión de vapor de 0.01 a 374.14 °C. Note que la presión de vapor a 100 °C es igual a 1 atm, lo que significa que el punto de ebullición del agua es 100°C a nivel del mar.

Escriba una función en MATLAB que calcule la presión de vapor de vapor saturado a una temperatura específica entre 100 y 150 °C, usando interpolación de Lagrange de cuatro puntos. El programa deberá recibir la temperatura en °C y producir la presión de vapor en atm. Corra el programa para calcular la presión en 112, 114, 120, 136, 148 °C, y asegúrese que los resultados tengan sentido.

6.1.5. Errores en la interpolación polinomial

Consideremos dos preguntas:

- (1) Si queremos interpolar alguna función $f(x)$ en $n + 1$ nodos en algún intervalo cerrado, ¿cómo escogemos los nodos?
- (2) ¿Qué tan preciso podemos hacer un polinomio de interpolación sobre un intervalo cerrado?

La respuesta a la primera pregunta es que los nodos x_i no deben estar distribuidos uniformemente en el intervalo. Miremos esto con un ejemplo.

Ejemplo 6.1 *Función de Runge.* Para

$$f(x) = (1 + x^2)^{-1},$$

conocida como la Función de Runge, supongamos que $p_n(x)$ interpola f en n nodos uniformemente distanciados en $[-5, 5]$. Encuentre el error máximo entre el valor del polinomio de interpolación $p_n(x)$ y la función $f(x)$.

Solución: El error crece con el número de puntos. De hecho se puede demostrar que

$$\lim_{n \rightarrow \infty} \max_{x \in [-5, 5]} |p_n(x) - f(x)| = \infty$$

A menos que la función a interpolar $f(x)$ sea un polinomio de grado N o menor, el polinomio de interpolación $P_N(x)$ y $f(x)$ no corresponderán entre datos. La diferencia entre el valor aproximado y el valor exacto

$$e(x) \equiv P_N(x) - f(x) \quad (6.25)$$

es el error de interpolación. Por definición de los polinomios de interpolación, se tiene que

$$e(x_i) = 0 \quad (6.26)$$

para $i = 1, \dots, N + 1$. En general, sin embargo, $e(x) \neq 0$ si $x \neq x_i$. Estimar la magnitud del error de interpolación es útil por tres razones: (a) sirva para acreditar o desacreditar los resultados de la interpolación; (b) puede ser usado para obtener una idea de que tanto el error puede ser reducido mediante la inclusión de más datos; (c) es el punto de partida para encontrar la distribución optima de los datos en la abscisa que dará los mejores resultados (Chebyshev).

Se puede demostrar que cuando la función $f(x)$ es lo suficientemente suave, el error asociado con la interpolación polinomial está dado por

$$e(x) = -\frac{f^{(N+1)}(\xi)}{(N+1)!} \Phi_{N+1}(x) \quad (6.27)$$

donde $\Phi_{N+1}(x)$ es el polinomio generador de Lagrange de orden $(N+1)$ ecuación 6.11 y $f^{(N+1)}(\xi)$ es el valor de la $N+1$ derivada de la función f evaluada en el punto ξ que está en algún sitio entre x_1 y x_{N+1} .

6.2. Interpolación por partes y Trazas

Cuando el número de datos es alto, la interpolación polinomial puede introducir un error considerable en las ordenadas $f(x_i)$, especialmente cerca de los límites del intervalo de interpolación. Un remedio para esto es reemplazar el polinomio de interpolación con una colección de polinomios locales de interpolación de orden menor, cada uno involucrando un número pequeño de datos. En contraste, un polinomio de interpolación cuya gráfica pasa por todos los puntos es un polinomio global.

Los polinomios de interpolación local se definen sobre intervalos que involucran m datos sucesivos.

6.2.1. Interpolación lineal

En la interpolación lineal se usas dos puntos consecutivos y produce el polinomio de grado uno

$$\begin{aligned} P_1^{(i)} &= f(x_i) + (x - x_i) \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \\ &= f(x_i) \frac{x - x_{i+1}}{x_i - x_{i+1}} + f(x_{i+1}) \frac{x - x_i}{x_{i+1} - x_i} \end{aligned} \quad (6.28)$$

para $x_i < x < x_{i+1}$. La primera expresión es la representación familiar de una línea recta en términos del intercepto y la pendiente. La segunda expresión sigue el método de interpolación de Lagrange.

El error de interpolación está dado por la ecuación 6.27 con $N = 1$, en términos de el valor desconocido $\xi(x)$. Una expresión alternativa, y un poco más útil, resulta de expandir $f(x)$ y $f'(x)$ en series de Taylor alrededor del extremo izquierdo x_i , y notando que la diferencia $x - x_i$ está en el rango entre 0 y h_i , donde $h_i \equiv x_{i+1} - x_i$, para obtener

$$f(x) = f(x_i) + f'(x_i)(x - x_i) + \frac{1}{2}f''(x_i)(x - x_i)^2 + \mathcal{O}(h_i^3) \quad (6.29)$$

y

$$f'(x) = f'(x_i) + f''(x_i)(x - x_i) + \mathcal{O}(h_i^2) \quad (6.30)$$

Ahora, evaluamos la ecuación 6.30 en $x = x_{i+1}$, resolvemos para $f'(x_i)$, sustituimos el resultado en el segundo término de la mano derecha de la ecuación 6.29 y encontramos que el error está dado por

$$P_1^{(i)} - f(x) = -\frac{1}{2}f''(x_i)(x - x_i)(x - x_{i+1}) + \mathcal{O}(h_i^3) \quad (6.31)$$

Notamos otra vez que la diferencia $x - x_i$ está entre 0 y h_i , mientras que la diferencia $x - x_{i+1}$ está en el rango entre $-h_i$ y 0. Esto demuestra que la magnitud del primer término al lado derecho es del orden h_i^2 .

El primer término del error mostrado en la ecuación 6.31 pudo haberse deducido a partir de la ecuación 6.27, haciendo $N = 1$, expandiendo $f''(\xi_i)$ en una serie de Taylor alrededor del punto x_i , y reteniendo tan solo la contribución constante. El procedimiento seguido aquí es más conveniente, en el sentido que nos permite expresar correcciones de orden mayor en términos de las derivadas de $f(x)$ evaluadas en los datos de manera sistemática.

Ejercicio 6.3 Segunda contribución al error.

Deduzca la segunda contribución al error en la mano derecha de la ecuación 6.31. *Pista:* Retenga un término adicional en el lado derecho de las ecuaciones 6.29y 6.30, e introduzca la expansión correspondiente para la segunda derivada.

6.2.2. Interpolación cuadrática

En la interpolación cuadrática se ajusta una parábola por una trío de puntos consecutivos por medio de un polinomio de interpolación de grado dos que puede ser expresado de la siguiente manera

$$P_2^{(i)}(x) = f_i + (a_i(x - x_i) + b_i)(x - x_i) \quad (6.32)$$

para $x_{i-1} < x < x_{i+1}$. Los coeficientes a_i y b_i se pueden calcular de manera secuencial así

$$\begin{aligned} a_i &= \frac{1}{x_{i+1} - x_{i-1}} \left(\frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} - \frac{f(x_{i-1}) - f(x_i)}{x_{i-1} - x_i} \right) \\ b_i &= \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} - a_i(x_{i+1} - x_i) \end{aligned} \quad (6.33)$$

cuando los tres puntos están uniformemente distanciados a lo largo del eje x , esto es, $x_{i+1} - x_i = x_i - x_{i-1} = h$, las fórmulas 6.33 se simplifican a

$$\begin{aligned} a_i &= \frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1}))}{2h^2} \\ b_i &= \frac{f(x_{i+1}) - f(x_{i-1}))}{2h} \end{aligned} \quad (6.34)$$

Como en la subsección 6.2.1, encontramos que el error de interpolación puede ser expresado en la forma asintótica

$$P_2^{(i)} - f(x) = -\frac{1}{6}f''(x_i)(x - x_{i-1})(x - x_i)(x - x_{i+1}) + \mathcal{O}(h_i^4) \quad (6.35)$$

donde se asume que la magnitud de $h_{i+1} \equiv x_{i+2} - x_{i+1}$ es comparable con la de $h_i \equiv x_{i+1} - x_i$.

6.2.3. Interpolación por trazadores cúbicos.

En la interpolación por trazadores cúbicos (*cubic-spline interpolation*) se ajusta un polinomio de grado tres a través de dos datos consecutivos, ajustando la primera y segunda derivada de polinomios adyacentes. De esta forma la colección de polinomios cúbicos conforman una función con primeras y segundas derivadas continuas, pero discontinua tercera derivada y derivadas de orden más alto iguales a cero.

Por simplicidad en la notación, denotamos $f(x_i) \equiv f_i$. Procediendo con el cálculo de los polinomios, expresamos el i -ésimo polinomio cúbico en la forma

$$P_3^{(i)}(x) = [(a_i(x - x_i) + b_i)(x - x_i) + c_i](x - x_i) + f_i \quad (6.36)$$

para $x_i < x < x_{i+1}$ donde $i = 1, \dots, N$ y emprendemos el cálculo de los $3N$ coeficientes desconocidos a_i , b_i , y c_i .

Primero, requerimos que $P_3^{(N)}(x_{N+1}) = f_{N+1}$

$$\begin{aligned} P_3^{(i)}(x_{i+1}) &= P_3^{(i+1)}(x_{i+1}) = f_{i+1} \\ P_3^{(i)}(x_{i+1}) &= P_3^{(i+1)}(x_{i+1}) \\ P_3^{(i)}(x_{i+1}) &= P_3^{(i+1)}(x_{i+1}) \end{aligned} \quad (6.37)$$

para $i = 1, \dots, N-1$, y derivamos $2N$ expresiones para los coeficientes a_i y c_i en términos de b_i ,

$$a_i = \frac{1}{3} \frac{b_{i+1} - b_i}{h_i}, \quad c_i = \frac{f_{i+1} - f_i}{h_i} - \frac{1}{3} h_i (b_{i+1} + 2b_i) \quad (6.38)$$

donde $h_i = x_{i+1} - x_i$ y $i = 1, \dots, N$. Adicionalmente, encontramos que los $N+1$ valores desconocidos de b_i satisfacen el sistema tridiagonal de $N-1$ ecuaciones

$$\frac{h_i}{3} b_i + \frac{2}{3} (h_i + h_{i+1}) b_{i+1} + \frac{h_{i+1}}{3} b_{i+2} = \frac{f_{i+2} - f_{i+1}}{h_{i+1}} - \frac{f_{i+1} - f_i}{h_i} \quad (6.39)$$

para $i = 1, \dots, N-1$, donde hemos definido

$$b_{N+1} \equiv \frac{1}{2} P_3^{(N)}(x_{N+1}) \quad (6.40)$$

Cuando los puntos están uniformemente distribuidos a lo largo del eje x , separados por una distancia h , el sistema 6.39 adquiere la forma simplificada

$$b_i + 4b_{i+1} + b_{i+2} = 3 \frac{f_{i+2} - 2f_{i+1} + f_i}{h^2} \quad (6.41)$$

para $i = 1, \dots, N-1$.

Para cerrar el problema necesitamos dos más condiciones sobre la forma de las trazas que completa la formulación matemática del problema, tenemos varias posibilidades para estas:

- En un método, relacionamos b_1 con b_2 y b_3 , y b_{N+1} con b_N y b_{N-1} , mediante extrapolación y después solucionamos el sistema tridiagonal completado 6.39 para b_i .
- Para obtener una traza amarada (*clamped cubic spline*), especificamos la pendiente de la primera y última traza en los primeros y últimos puntos.
- Si la función a interpolar es periódica, requerimos que la primera y segunda derivada de la primera traza en el primer punto sea igual a las de la última traza en el último punto.
- Para obtener una traza cúbica natural (*natural cubic spline*), requerimos que la curvatura (segunda derivada) de la primera y última traza en el primer y último punto sea igual a cero.

Ejercicio 6.4 *Trazas cúbicas.*

Derive el sistema de ecuaciones 6.38 y 6.39.

Capítulo 7

Aproximando derivadas

Hemos usado ya varias veces aproximaciones de diferencias finitas para discretizar las ecuaciones de Laplace y Poisson. Sin embargo no hemos introducido formalmente el concepto y la formalidad de las diferencias finitas. En este capítulo derivamos algunas de las fórmulas de diferencias finitas que hemos usando y generalizamos el método para que usted pueda calcular cualquier derivada con cualquier orden de precisión.

7.1. Diferencias finitas

Suponga que tenemos una función $f(x)$ y deseamos calcular $f'(x)$ en algún punto x . Para este propósito usaremos el teorema de Taylor:

$$f(x+h) = f(x) + f'(x)h + \frac{f''(\xi)h^2}{2}. \quad (7.1)$$

Reacomodando esta ecuación obtenemos

$$f'(x) = \frac{f(x+h) - f(x)}{h} - \frac{f''(\xi)h}{2}. \quad (7.2)$$

Recuerde que ξ es algún número entre x y $x+h$ pero su valor exacto se desconoce. Si deseamos calcular $f'(x)$, no podemos evaluar $f''(\xi)$, entonces aproximamos la derivada ignorando el último término. En otras palabras, usamos $[f(x+h) - f(x)]/h$ como una aproximación de $f'(x)$. Al hacer esto hemos despreciado el último término. Si $f''(x)$ está acotado en el intervalo en cuestión entonces el término despreciado está acotado por una constante por h . Esto es

$$f'(x) = \frac{f(x+h) - f(x)}{h} + \mathcal{O}(h) \quad (7.3)$$

El error en que incurrimos al aproximar $f'(x)$ con $[f(x+h) - f(x)]/h$ se denomina *error de truncamiento*. No tiene nada que ver con el tipo de error que obtenemos cuando hacemos cálculos en un computador con precisión limitada; aún aunque desarrollara los cálculos en un computador con infinita precisión todavía tendría el error de truncamiento.

El error de truncamiento se puede hacer pequeño haciendo h pequeño. Sin embargo, a medida que h se hace más pequeño, la precisión se pierde en la ecuación 7.3 debido a la cancelación por resta. El error en el cálculo para h pequeño se denomina *error de redondeo*. Generalmente el error de redondeo aumentará cuando h decrece. De esta forma debe existir un valor de h diferente de cero para el cual la suma de estos dos errores se minimiza.

El error de truncamiento para esta aproximación es de $\mathcal{O}(h)$. Generalmente se desea una aproximación más precisa. Ya deberían saber que cualquier cálculo comienza con el Teorema de Taylor:

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

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

Restando estas dos líneas obtenemos:

$$f(x+h) - f(x-h) = 2f'(x)h + \frac{f'''(\xi_1) + f'''(\xi_2)}{3!}h^3.$$

Entonces se tiene,

$$2f'(x)h = f(x+h) - f(x-h) - \frac{f'''(\xi_1) + f'''(\xi_2)}{3!}h^3$$

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} - \frac{f'''(\xi_1) + f'''(\xi_2)}{2} \frac{h^2}{6}$$

Si $f'''(x)$ es continua, entonces existe un ξ entre ξ_1 y ξ_2 tal que $f'''(\xi) = \frac{f'''(\xi_1) + f'''(\xi_2)}{2}$. (Teorema del valor medio.) Asumiendo que existe $f'''(\cdot)$ está acotado uniformemente, obtenemos

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + \mathcal{O}(h^2) \quad (7.4)$$

En algunas situaciones será necesario evaluar la derivada en lugares ‘extraños’, entonces debemos encontrar nuevas fórmulas para la derivada. Esto se puede hacer fácilmente usando el teorema de Taylor. Esto se ilustra en el siguiente ejemplo.

Ejemplo 7.1 Utilizando evaluaciones de f en $x+h$ y $x+2h$ aproxime $f'(x)$. Asuma que $f(x)$ es una función analítica, esto es, que se puede derivar infinitamente.

Solución: Primero use el teorema de Taylor para expandir $f(x+h)$ y $f(x+2h)$

$$\begin{aligned} f(x+h) &= f(x) + f'(x)h + f''(x)\frac{h^2}{2} + f'''(x)\frac{h^3}{3!} + \dots \\ f(x+2h) &= f(x) + f'(x)2h + f''(x)\frac{(2h)^2}{2} + f'''(x)\frac{(2h)^3}{3!} + \dots \end{aligned}$$

Multiplicando la primera ecuación por 4 y restando la segunda de la primera obtenemos

$$4f(x+h) - f(x+2h) = 3f(x) + 2f'(x)h + 0f''(x) - 4f'''(x)\frac{h^3}{3!}$$

Ahora podemos despejar una expresión para $f'(x)$, suponiendo que $f'''(x)$ está acotada obtenemos

$$f'(x) = (-3f(x) + 4f(x+h) - f(x+2h))/2h + \mathcal{O}(h^2) \quad (7.5)$$

El cálculo de las segundas derivadas procede de manera similar.

Capítulo 8

Integración numérica

8.1. Cálculo de integrales en una dimensión

La interpolación de funciones tiene un gran campo de aplicación en el cálculo de integrales de funciones de una o más variables. En la versión unidimensional del problema, se desea calcular un número que exprese la integral de una función $f(x)$ entre dos límites especificados a y b ,

$$I = \int_a^b f(x)dx \quad (8.1)$$

En cálculo se demostró que I expresa el área entre la gráfica de la función $y = f(x)$, el eje x , y la líneas verticales $x = a$ y $x = b$. Cuando $b > a$, y por lo tanto $dx > 0$, el área por encima del eje x se suma y el área debajo del eje x se resta. Cuando $b < a$, lo opuesto es cierto. A la función $f(x)$ integrada se le denomina *integrando*.

8.1.1. ¿Cuándo necesitamos integración numérica?

Aunque tengamos acceso a tablas de integración (o a software especializado) algunas veces la integración analítica no es posible. La integración numérica se hace necesaria bajo las siguientes condiciones:

- El integrando $f(x)$ es lo suficientemente complicado, de tal forma que su integral indefinida o incluso definida en cierto rango no puede ser especificada. Recuerde que la integral indefinida es una función, mientras que la integral definida es un número.
- La integral indefinida o definida puede ser encontrada, pero su forma es tan intrincada que la evaluación se hace imposible.
- El integrando se da en términos de una tabla. Esto es, se nos proporcionan valores discretos de $f(x)$ en $N + 1$ puntos x_i , $i = 1, \dots, N + 1$ dentro del dominio de integración.
- La integración numérica es un componente indispensable de los métodos disponibles para la solución de ecuación diferenciales e integrales.

8.1.2. Estrategia general para la integración numérica

Un método típico de integración numérica involucra estos dos pasos:

- (1) Aproximación del integrando $f(x)$ ya sea con un polinomio global de interpolación definido en todo el dominio de integración, o con una colección de polinomios locales de interpolación definidos sobre intervalos que involucren un número menor de puntos.
- (2) Integración de los polinomios de interpolación sobre sus dominios individuales de definición.

Una posible modificación a este procedimiento básico se requiere cuando el integrando es singular en uno o más puntos en el dominio de integración, o cuando uno o dos límites de integración son infinitos.

El uso de polinomios globales de integración no es recomendable debido a los posibles problemas de convergencia resultantes del aumento en el número de puntos. De esta manera la integración en términos de polinomios locales es preferida en el desarrollo de métodos de integración de propósito general.

8.2. Integración mediante interpolación polinomial local: Reglas de Newton-Cotes

Comenzamos haciendo la suposición que ni el integrando $f(x)$ ni alguna de sus derivadas es infinito en ningún punto en el dominio de integración $[a, b]$, y que los límites de integración a y b son finitos. Más aún, asumimos que $f(x)$ puede ser calculado, o que su valor es conocido, en $N + 1$ puntos x_i , $i = 1, \dots, N + 1$ que están distribuidos de alguna manera en $[a, b]$, con $x_1 = a$ y $x_{N+1} = b$.

8.2.1. Regla del trapecioide

Aproximando la gráfica de $f(x)$ con una línea poligonal que pase por pares sucesivos de puntos nos llega a la regla del trapecioide. El área confinada entre la línea poligonal, el eje x , y las dos líneas verticales que pasan por dos datos sucesivos conforman un trapecioide.

Consideremos, en particular, la sección de la función $f(x)$ entre los puntos x_i y x_{i+1} , usamos la expresión para un polinomio de grado uno y la fórmula de error para calcular

$$\int_{x_i}^{x_{i+1}} f(x)dx \approx \int_{x_i}^{x_{i+1}} P_1^{(i)}(x)dx + \frac{1}{2}f''(x_i) \int_{x_i}^{x_{i+1}} (x - x_i)(x - x_{i+1})dx \quad (8.2)$$

Evaluando las integrales en el lado derecho de 8.2, encontramos

$$\int_{x_i}^{x_{i+1}} f(x)dx \approx \frac{1}{2}h_i(f_i + f_{i+1}) - \frac{1}{12}f''(x_i)h_i^3 \quad (8.3)$$

donde $h_i \equiv x_{i+1} - x_i$.

Regla del trapecioide compuesta

Sumando la contribución de todos los intervalos, obtenemos la regla del trapecioide compuesta cerrada expresada por

$$I \approx \frac{1}{2} [h_1 f_1 + (h_1 + h_2) f_2 + \cdots + (h_{N-1} + h_N) f_N + h_N f_{N+1}] - \frac{1}{12} \sum_{i=1}^N f''(x_i) h_i^3 \quad (8.4)$$

El calificativo *cerrada* significa que la regla de integración involucra los valores del integrando en ambos extremos.

Puntos uniformemente distribuidos

Cuando los puntos están uniformemente distribuidos $h_i = h = (b - a)/N$, la regla trapezoidal compuesta se reduce a la versión trapezoidal de la *regla cerrada de Newton-Cotes* expresada por

$$I \approx h \left(\frac{1}{2} f_1 + f_2 + \cdots + f_N + \frac{1}{2} f_{N+1} \right) - \frac{1}{12} h^3 \sum_{i=1}^N f''(x_i) \quad (8.5)$$

Una regla general de Newton-Cotes produce la aproximación numérica a una integral en términos de valores del integrando en una serie de puntos uniformemente distribuidos. El error asociado es de orden $\mathcal{O}(h^2)$.

8.2.2. Regla de Simpson 1/3

Aproximando la gráfica del integrando $f(x)$ con una colección de parábolas que pasan a través de tríos de puntos sucesivos, correspondiente a polinomios locales de interpolación de segundo orden, nos lleva a la regla de Simpson 1/3.

Usando el método de Lagrange para construir los polinomios de interpolación, encontramos que la contraparte de la ecuación 8.3 para puntos uniformemente distribuidos, separados por una distancia h es

$$\int_{x_{i-1}}^{x_{i+1}} f(x) dx \approx \frac{1}{3} h (f_{i-1} + 4f_i + f_{i+1}) - \frac{1}{90} f^{(4)}(x_i) h_i^5 \quad (8.6)$$

Regla de Simpson 1/3 compuesta

Suponiendo que el número de intervalos N es par, sumamos las contribuciones de las $\frac{1}{2}N$ parábolas y obtenemos la regla de Simpson compuesta cerrada

$$I \approx \frac{1}{3} h (f_1 + 4f_2 + 2f_3 + 4f_4 + \cdots + 2f_{N-1} + 4f_N + f_{N+1}) + E^{Simp}(h) \quad (8.7)$$

donde $E^{Simp}(h)$ es el error asociado de orden $\mathcal{O}(h^4)$.

8.2.3. Problemas

Ejercicio 8.1 *Integrales de línea.*

(a) Considere la siguiente integral de línea a lo largo del círculo de radio unitario centrado en el origen del plano xy

$$I(x_0, y_0) = \int_{\text{Circulo}} \ln [(x(l) - x_0)^2 + (y(l) - y_0)^2] dl \quad (8.8)$$

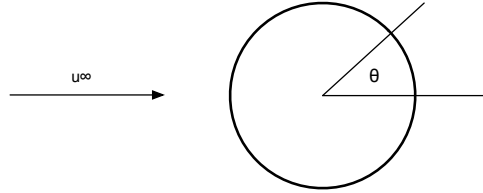
donde l es la longitud del arco a lo largo del círculo, y (x_0, y_0) es un punto específico. Evalúe la integral en dos puntos diferentes (x_0, y_0) localizados dentro del círculo, y dos puntos localizados por fuera del círculo y discuta sus resultados. Note que a lo largo del círculo de radio unitario $dl = d\theta$, donde θ es el ángulo medido en la dirección opuesta a las manecillas del reloj.

(b) Calcule la *integral de línea*

$$I = \int x dl = \int x \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx \quad (8.9)$$

donde l es la longitud de arco del segmento de línea $y = \sin x$ en el intervalo $0 \leq x \leq \frac{1}{2}\pi$, usando el método de su preferencia (ojalá lo haga con varios y compare).

Ejercicio 8.2 *Flujo potencial a través de una capa de vórtices.*



El flujo potencial en dos dimensiones debido a un cuerpo en movimiento puede ser expresado en términos de una capa de vórtices situada sobre el cuerpo. Específicamente, la velocidad \mathbf{u} puede ser expresada como la suma de un flujo incidente \mathbf{u}^∞ que domina a grandes distancias del cuerpo, y la velocidad inducida por la capa de vórtices \mathbf{u}^V . La fuerza de la capa de vórtices γ es igual al componente tangencial de la velocidad en la frontera, vista en un marco de referencia bajo el cual el cuerpo parece estar estacionario.

Si γ se conoce, \mathbf{u}^V puede ser calculada en cualquier punto en el flujo mediante integración numérica. En variables complejas, la velocidad \mathbf{u}^V en la posición compleja $z_0 = x_0 + iy_0$, donde i es la unidad imaginaria, está dada por

$$u^V - iv^V = \frac{1}{2\pi i} \int_C \frac{1}{z - z_0} \gamma dl \quad (8.10)$$

donde C es el contorno del cuerpo, l es la longitud de arco a lo largo del contorno del cuerpo, y $z = x + iy$.

Considere el flujo a lo largo del eje x con velocidad V que pasa por un cilindro estacionario de radio a con su centro en el origen. La fuerza de la capa de vórtices se conoce y es $\gamma = 2V \sin \theta$, donde θ es el ángulo polar.

(a) Desarrolle un método numérico para calcular la integral en la mano derecha de 8.10, y escriba un programa en MATLAB.

- (b) Calcule y grafique los perfiles de las componentes x y y de la velocidad en varios puntos a lo largo del segmento $x = 2$, $-2 < y < 2$.
- (c) Calcule la velocidad en varios puntos dentro del cilindro y discuta sus hallazgos.
- (Bono) Muestre las líneas de flujo alrededor del cilindro.

Capítulo 9

Solución de ecuaciones diferenciales

Consideremos la ecuación

$$\frac{dy}{dt} = 1 - e^{-t} \quad (9.1)$$

La función desconocida $y(t)$ se encuentra en forma de una derivada dy/dt . La ecuación 9.1 es una ecuación diferencial ordinaria y solución se puede encontrar de manera sencilla como

$$y(t) = t + e^{-t} + C \quad (9.2)$$

donde C es una constante de integración. Todas las funciones en la ecuación 9.2 son solución de la ecuación diferencial 9.1. Es decir, la familia de curvas mostradas en la figura son todas solución de la ecuación diferencial.

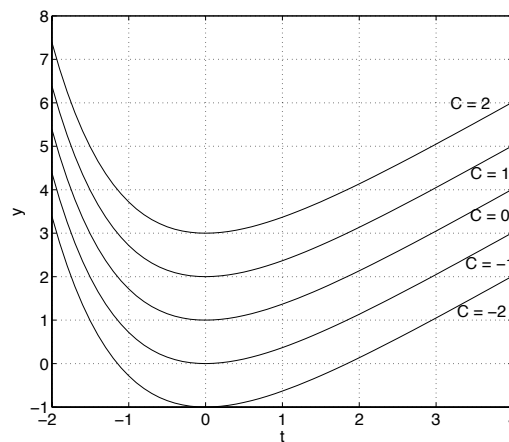


Figura 9.1: Curvas solución $y(t) = t + e^{-t} + C$.