

Elementos de Análisis Numérico algoritmos y aplicaciones

Liudmila Chainskaia
Elizabeth Doig



PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FONDO EDITORIAL 1999

LIUDMILA CHAINSKAIA
ELIZABETH DOIG

Elementos de
Análisis Numérico
algoritmos y aplicaciones



PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FONDO EDITORIAL 1999

Primera Edición, diciembre 1999

Elementos de Análisis Numérico
algoritmos y aplicaciones

Diagramación: Lourdes Rodríguez Matysek
Manuel Aristondo Gárate

Copyright © 1999 por Fondo Editorial de la Pontificia
Universidad Católica del Perú. Av. Universitaria, cuadra 18.
San Miguel. Telefax: 460-0872 Telfs.: 460-2870, 460-2291,
anexos 220-356.

Prohibida la reproducción de este libro por cualquier medio,
total o parcialmente, sin permiso expreso de los editores.

Derechos reservados
ISBN: 9972-42-152-4

Depósito legal
150105 2000-0066

Impreso en el Perú - Printed in Peru

A nuestras familias

Contenido

PREFACIO	XI
1. <i>Los Métodos Numéricos y la Aritmética de la Computadora</i>	1
1.1 Introducción	1
1.2 Fuentes del Error	3
1.3 Aritmética de Punto Flotante	7
1.4 Error Propagado	13
1.5 Convergencia de Sucesiones	19
1.6 Ejercicios	23
2. <i>Ecuaciones Algebraicas No Lineales</i>	27
2.1 Separación de Raíces	28
2.2 Método de Bisección	33
2.3 Método de Punto Fijo	40
2.4 Orden de Convergencia	51
2.5 Método de Newton-Raphson	53
2.6 Convergencia Acelerada	68
2.7 Método de Steffensen	73
2.8 Método de Punto Fijo para Sistemas	76
2.9 Método de Newton para Sistemas	82
2.10 Método de Bairstow para determinar el Factor Cuadrático de un Polinomio	89
2.11 Ejercicios	98
3. <i>Solución de Sistemas de Ecuaciones Lineales</i>	111
3.1 Métodos Directos de Solución de Sistemas Lineales	115
3.2 Cálculo de la Inversa de una Matriz	126
3.3 Factorización Matricial Triangular	134
3.4 Estrategias de Pivoteo	147
3.5 Métodos Iterativos para la Solución de Sistemas Lineales	154
3.6 Convergencia de los Métodos Iterativos	171
3.7 Error y Residual de una Solución Aproximada	181
3.8 Método de Potencia	184
3.9 Ejercicios	190

4.	<i>Aproximación de Funciones</i>	205
4.1	Aproximación Polinómica	207
4.2	Aproximación de Funciones Usando el Polinomio de Interpolación de Lagrange	209
4.3	Polinomio de Interpolación de Newton con Diferencias Divididas	217
4.4	Interpolación de una Función Basada en Puntos Igualmente Espaciados	229
4.5	Aplicaciones de los Polinomios de Interpolación: Derivación e Integración Numérica	240
4.6	Ajuste de Datos	263
4.7	Producto Escalar y Ortogonalidad de las Funciones	271
4.8	Aproximación de Mínimos Cuadrados por Polinomios Ortogonales	278
4.9	Cuadratura Gaussiana	283
4.10	Ejercicios	290
5.	<i>Ecuaciones Diferenciales Ordinarias</i>	301
5.1	El Método de Euler	301
5.2	Método de Taylor de Orden Superior	306
5.3	Métodos de Runge-Kutta	310
5.4	Métodos Predictor-Corrector	319
5.5	Sistemas de Ecuaciones Diferenciales	320
5.6	Ecuaciones Diferenciales de Orden Mayor	323
5.7	Ejercicios	327
APÉNDICE A:		
	Repaso de los Resultados Básicos del Cálculo	331
APÉNDICE B:		
	Programas en Pascal	338
APÉNDICE C:		
	Respuestas a ejercicios seleccionados	339
BIBLIOGRAFÍA		349
ÍNDICE		351

Prefacio

Este libro ha sido escrito como resultado de la experiencia adquirida durante la enseñanza de cursos sobre métodos numéricos para estudiantes de Ciencias e Ingeniería.

El libro presenta los conceptos teóricos fundamentales de los temas del Análisis Numérico y va acompañado de los algoritmos y programas en Pascal correspondientes. Con el fin de facilitar el entendimiento de estos conceptos fundamentales añadimos el apéndice A, como un repaso de los resultados básicos del cálculo; así mismo para aquellos lectores que tengan interés en profundizar algunos temas, se les dan las referencias bibliográficas apropiadas. Este libro podrá ser usado por un estudiante a nivel de pregrado, con conocimientos básicos de análisis matemático.

El libro está compuesto de 5 capítulos y 3 apéndices, que a su vez están divididos en secciones cada uno. Al inicio de cada capítulo se presenta una motivación con la finalidad de introducirlos en el tema correspondiente. Las demostraciones de los teoremas se presentan para ilustrar y reforzar los conceptos aprendidos. Cada tema está acompañado de gráficas y ejemplos escogidos para ayudar al entendimiento de la teoría, algunos de los cuales se utilizan para varios métodos con el propósito de ilustrar la eficiencia de cada método. Los algoritmos se presentan como un resumen o consecuencia del método desarrollado. En algunos casos se presentan comentarios en los cuales se dan algunas ideas teóricas y/o prácticas que ayudan a un mejor manejo de los métodos.

Al final de cada capítulo se coloca un conjunto de problemas, diseñados para reforzar y aplicar el material. En algunos casos se requiere del uso del Laboratorio de Computación con un Lenguaje de Programación. En el apéndice B, se muestran 15 programas en Pascal

en el disquete que se incluye en el libro, que corresponden a una parte de los algoritmos presentados. Estos se pusieron justo al final, para evitar interrumpir la exposición matemática en el cuerpo principal del texto.

En el apéndice C, se proporcionan las respuestas de los problemas seleccionados.

El estilo del texto es poco formal, para aumentar la facilidad de lectura y mejorar el flujo de las ideas, de manera que se las puede aplicar rápidamente, interrelacionándolas unas con otras.

Deseamos agradecer especialmente a José Inga, Magister en Informática, quien colaboró en la elaboración de los programas en Pascal; y así también a Katia Vigo, Licenciada en Matemáticas, por su apoyo en la solución de los problemas seleccionados.

Los Métodos Numéricos y la Aritmética de la Computadora

1.1 Introducción

Durante muchos años, la evolución de las matemáticas fue determinada por las necesidades que surgían en ciencias tales como la Física, la Ingeniería, la Economía, la Biología, la Astronomía, etc. En estas ciencias se presentaron problemas prácticos que las matemáticas debían hacer frente, obteniendo una solución numérica. Para los problemas que surgen en ellas, es necesario construir un modelo matemático que describa el problema que se está estudiando, para luego analizar sus propiedades, con el objeto de hallar una solución numérica confiable.

Las fórmulas matemáticas y científicas (métodos exactos o analíticos) pueden usarse para obtener respuestas numéricas a una variedad de problemas reales. Sin embargo, muchas veces afrontaremos situaciones, en las cuales los métodos exactos o analíticos no nos permitirán encontrar una solución, ya que no tendremos la fórmula adecuada, como por ejemplo en el caso de calcular la integral:

$$\int_0^b e^{-x^2} dx$$

Es en estos casos, que los procedimientos matemáticos que nos proporcionan una solución numérica aproximada del problema (métodos numéricos), surgen como una alternativa para la solución de estos problemas. En un método numérico será importante que su aplicación no requiera demasiado tiempo, ni que ocupe demasiado espacio en la memoria de la computadora; así también, que sea un procedimiento con un orden, convergente y estable. La materia que se ocupa de estudiar las propiedades relativas a los métodos y de los errores en la solución numérica, recibe el nombre de Análisis Numérico.

En la actualidad, con la multiplicación, el perfeccionamiento y la disponibilidad de las computadoras digitales a bajo costo, se ha facilitado enormemente el uso y el desarrollo del Análisis Numérico; ya que hace 5 ó 6 décadas atrás, el acceso a las computadoras estaba limitado por el costo de las mismas. De manera que para resolver los problemas se empleaban planteamientos analíticos simples. La mayor parte de las veces estas soluciones resultan útiles y facilitan la comprensión del comportamiento de algunos procedimientos matemáticos. Sin embargo, las soluciones analíticas tienen valor práctico limitado, porque la mayor parte de los problemas incluyen formas y procesos complejos.

Sobre la base de lo expuesto anteriormente, definimos el Análisis Numérico como,

una materia que se ocupa de la construcción, aplicación y análisis de las propiedades de los métodos numéricos, para la solución de distintos problemas matemáticos que se presentan en diferentes ramas del saber; tanto de ciencias como de ingeniería, realizándose los cálculos a partir de datos numéricos.

Los fines de cada método numérico son:

- a. Calcular una solución aproximada por medio de un número finito de pasos.
- b. Obtener una solución eficaz y confiable.

En el proceso de construir los métodos numéricos, se emplean procedimientos (algoritmos) que consisten de un conjunto de instrucciones para la ejecución de operaciones aritméticas o lógicas en un orden preestablecido, que nos conducirán a la solución aproximada de un problema dado.

Una vez elegido el algoritmo que se usará, también deberemos considerar todas las fuentes de error que puedan afectar los resultados, ya que esta solución será aproximada y requeriremos de determinada precisión.

1.2 Fuentes del Error

Generalmente, trabajamos con datos que son resultados de un proceso experimental, y al mismo tiempo, las medidas experimentales que obtenemos a lo más logran una precisión del 0.1%, esto es de 3 ó 4 cifras significativas. Sin embargo, una computadora al realizar las operaciones aritméticas con estos datos es capaz de producir resultados con al menos de 8, y a veces 16 ó más cifras, pero muchas veces estos valores no logran la precisión de los datos experimentales debido a los errores generados por el cálculo. En esta sección nos dedicaremos a estudiar las fuentes de error que afectan a la solución numérica hallada:

- a. En primer lugar, hay errores que surgen de los datos iniciales, debido a la imperfección de las mediciones.
- b. En segundo lugar, la naturaleza infinita de los números reales y la posibilidad de trabajarlos sólo en forma finita; ya sea por la capacidad de la calculadora o de la computadora, ocasiona que todas las operaciones que se realizan sean en forma aproximada, lo que da lugar a otro tipo de error conocido como error del cálculo o error del redondeo.
- c. En tercer lugar, todo proceso numérico debe ser un proceso finito, a pesar de que se use un concepto matemático infinito, como el concepto de límite, o de una serie infinita. Luego este proceso infinito se corta o se trunca, lo que da lugar a un error conocido como error de truncamiento o error del método.

Por ejemplo, en el problema de sumar una serie infinita, esto es:

$$e^x = 1 + x + x^2 / 2! + x^3 / 3! + \dots = \sum_{k=0}^{+\infty} x^k / k!$$

tendríamos que truncar la serie de Taylor después de un término de cierto orden, ya que es imposible incluir un número infinito de términos. Así, si la serie de Taylor para una función $f(x)$, en torno de $x = x^*$, se trunca después del término de orden N , se expresa como:

$$f(x) = f(x^*) + f'(x^*)h + f''(x^*)\frac{h^2}{2!} + \dots + f^{(N)}(x^*)\frac{h^N}{N!} + R_N(x) = \\ = P_N(x) + R_N(x);$$

donde:

$$h = x - x^*; \\ P_N(x) = \sum_{k=0}^N f^{(k)}(x^*)h^k / k!;$$

$$\text{y } R_N(x) = f^{(N+1)}(\xi(x))\frac{h^{N+1}}{(N+1)!} \text{ con } \xi(x) \text{ que está entre } x \text{ y } x^*;$$

a $P_N(x)$ se le llama el Polinomio de Taylor de grado N para $f(x)$ alrededor de x^* y a $R_N(x)$ se le llama el residuo (o error de truncamiento) asociado con $P_N(x)$.

Así por ejemplo, si se desea calcular aproximadamente $f(0.1) = e^{0.1}$, por medio de un polinomio de Taylor de grado 3, se tiene:

$$P_3(x) = 1 + x + x^2 / 2! + x^3 / 3! \\ R_3(x) = \frac{f^{(4)}(\xi(x))}{4!} x^4, \quad \xi(x) \in [0, x].$$

Entonces,

$$f(0.1) \approx P_3(0.1) = 1.10516666 \text{ y} \\ R_3(0.1) = f^{(4)}(\xi(0.1))(0.1)^4 / 4! = \\ = e^{\xi(0.1)}(0.1)^4 / 4! \leq \max_{[0, 0.1]} \{e^x\} 0.1^4 / 4! = 4.604878825 \times 10^{-6}$$

y si comparamos la aproximación obtenida con el valor exacto,

$$f(0.1) = 1.105170918$$

se observa que hay cuatro decimales que coinciden con el valor exacto, por lo que consideramos que es una solución confiable.

Todos estos errores se propagan a lo largo del proceso numérico, dando lugar a una respuesta aproximada, en vez de exacta y por lo tanto debe ser analizada, desde el punto de vista de qué tan cerca se encuentra esta respuesta del valor verdadero.

Nos dedicaremos ahora a analizar los errores del cálculo, ya que los errores del método se estudiarán en cada uno de los algoritmos que se obtengan más adelante.

Error del cálculo

Con el propósito de poder medir las cantidades aproximadas se pueden definir los conceptos de error absoluto y de error relativo.

Definición: Sea x un número real exacto, y x^* una aproximación de x , entonces se define:

el error absoluto como: $\Delta x = |x - x^*|$;

y el error relativo como: $\varepsilon_x = \frac{\Delta x}{|x|} = \frac{|x - x^*|}{|x|}$; $x \neq 0$

Ejemplo 1:

Si

$$x = \pi = 3.14159265\dots \text{ y } x^* = 3.14,$$

entonces el error absoluto es:

$$\Delta x = |\pi - 3.14| = 0.00159265\dots \leq 0.0016,$$

y el error relativo es:

$$\varepsilon_x = \Delta x / \pi = \frac{|\pi - 3.14|}{\pi} \leq \frac{0.0016}{3.14} \approx 5.095 \times 10^{-4} \quad (\text{ó } 0.0509\%),$$

ya que $x^* < x = \pi$.

Ejemplo 2:

a. Si

$$x = 2/3 = 0.6666666... \quad \text{y} \quad x^* = 0.667$$

entonces,

$$\Delta x = \text{error absoluto} = |x - x^*| = 0.001 / 3 = 10^{-3}/3$$

$$\varepsilon_x = \text{error relativo} = \frac{|x - x^*|}{|x|} = 0.0005 = 5 \times 10^{-4} \text{ (ó } 0.05\%).$$

b. Si

$$x = 0.003125 \quad \text{y} \quad x^* = 0.00312$$

entonces los errores:

$$\Delta x = 0.000005, \quad \varepsilon_x = 0.0016 \text{ ó } 0.16\%.$$

En la práctica es imposible trabajar con estos conceptos ya que generalmente no se conoce a x como cantidad exacta, entonces debemos usar una cota cercana (una TOLERANCIA), ya sea para el error absoluto o para el error relativo.

La necesidad de definir además del error absoluto, al error relativo, es por lo que sólo el error absoluto no mide suficientemente bien la cercanía buscada, ya que si los números son grandes en magnitud, la diferencia entre el valor exacto y el valor aproximado puede ser muy grande, sin que la aproximación sea “mala”.

Así, si

$$x = \pi \times 10^5 = 314159.265... \quad \text{y} \quad x^* = 3.14 \times 10^5 = 314000,$$

el error absoluto es $\Delta x = |x - x^*| = 159.265... < 160$ valor muy grande, pero el error relativo es $\varepsilon_x = \Delta x/x \leq \Delta x/x^* \leq 5 \times 10^{-3}$, una cantidad pequeña.

Por esta razón es preferible trabajar con el error relativo.

El conocimiento del error relativo además permite indicar la medida de la cercanía dentro de la misma presentación de la cantidad aproximada, usando la idea de los dígitos significativos o cifras exactas.

Definición: Se dice que el número aproximado x^* se aproxima a x , con k dígitos significativos (ó k cifras exactas) si k es el entero más grande no negativo, para el cual:

$$\varepsilon_x = \frac{|x - x^*|}{|x|} \leq 0.5 \times 10^{-k+1} = 5 \times 10^{-k}$$

Así, si

$$x = \pi, x^* = 3.14$$

$$\varepsilon_x = \frac{|x - x^*|}{|x|} \leq 5 \times 10^{-3},$$

nos indica que en la aproximación $x^* = 3.14$, las tres cifras son exactas; o si

$$x = \pi \times 10^5 \text{ y } x^* = 3.14 \times 10^5 = 314000,$$

entonces el error relativo

$$\varepsilon_x = \frac{|x - x^*|}{|x|} \leq 5 \times 10^{-3},$$

nos indica que la aproximación $x^* = 3.14 \times 10^5$ tiene 3 cifras exactas.

1.3 Aritmética de Punto Flotante

Un número de máquina tiene un número finito de dígitos, esto es *t-dígitos*; por lo que al realizar los cálculos en la computadora, los números reales se truncan, lo que origina los errores del cálculo.

Sea A el conjunto de números que se pueden representar en una máquina dada, esto es un conjunto finito, la pregunta que surge es ¿Cómo aproximar un número x que no pertenece al conjunto A , por medio de un número g que esté en A ? Este problema se presenta no sólo con los datos iniciales, sino también cuando se realizan los cálculos:

$$x \pm y, x \times y, x/y; \quad \text{con } x, y \in A$$

Bien, lo que se quiere es aproximar a cualquier número x que no esté en A , por un número de máquina $fl(x)$ que esté en A que satisfaga:

$$|x - fl(x)| \leq |x - g| \quad \forall g \in A$$

Dicho número se conoce como el número de punto flotante de x , y puede ser obtenido en la mayoría de los casos por redondeo o por corte.

Ejemplo 1:

Sí $t = 4$ dígitos, entonces:

- $fl(0.24185 \times 10^0) = 0.2419 \times 10^0$
- $fl(3.13161) = fl(0.313161 \times 10^1) = 0.3132 \times 10^1$
- $fl(-1.97631) = fl(-0.197631 \times 10^1) = -0.1976 \times 10^1$

En general se puede proceder como sigue a fin de encontrar $fl(x)$ para una aritmética de t -dígitos:

- Si $x \notin A$ se le representa en la forma $a \times 10^b$, en donde $|a| \geq 10^{-1}$ y la representación decimal de $|a|$ está dada por:

$$|a| = 0.\alpha_1\alpha_2\alpha_3\dots\alpha_r\alpha_{r+1}\dots$$

tal que $0 \leq \alpha_i \leq 9, \forall i \geq 2, \alpha_1 \neq 0$

- Se elige a $fl(x)$ como:

$$a^* = \begin{cases} 0.\alpha_1\alpha_2 \dots \alpha_r, & \text{si } \alpha_{r+1} \leq 4 \\ 0.\alpha_1\alpha_2 \dots \alpha_r + 10^{-r}, & \text{si } \alpha_{r+1} \geq 5 \end{cases}$$

Por lo tanto:

$$fl(x) = \text{signo}(x) \times a^* \times 10^b$$

Resumiendo:

Un número de punto flotante de t -dígitos en base β tiene la forma:

$$x = \pm (0.d_1d_2d_3\dots d_t)_\beta \times \beta^e$$

donde $(.d_1d_2d_3\dots d_t)_\beta$ es una fracción llamada la mantisa y e es un entero llamado el exponente o la característica. Tal número de punto flotante se dice normalizado en el caso $d_1 \neq 0$ y $0 \leq d_i < \beta$, para $i = 2, \dots, n$.

Generalmente la base β puede ser:

$$\beta = 10, \text{ ó } \beta = 2, \text{ ó } \beta = 8, \text{ ó } \beta = 16.$$

El exponente e se limita a un recorrido $m \leq e \leq M$, donde m y M son enteros, generalmente $m = -M$.

Comentario

En el caso de que la base sea decimal ($\beta = 10$), se encuentra que el error relativo admite la siguiente cota:

$$|fl(x) - x|/|x| \leq 5 \times 10^{-(t+1)} / |a| \leq 5 \times 10^{-t}; \text{ puesto que } |a| \geq 10^{-1},$$

y si denotamos como,

$$\varepsilon = \frac{fl(x) - x}{x} \Rightarrow x(1 + \varepsilon) = fl(x),$$

nos permite definir a fl como la función:

$$\begin{aligned} fl: \mathfrak{R} &\rightarrow A \\ x &\rightarrow fl(x) = x(1 + \varepsilon), \\ \text{donde } |\varepsilon| &\leq 5 \times 10^{-t}. \end{aligned}$$

Por otro lado, ya observamos que los resultados de las operaciones aritméticas:

$$x \pm y, \quad x \times y, \quad x / y,$$

no son necesariamente números de máquina. De ahí, basándose en el resultado anterior, se pueden definir operaciones sustitutas como:

$$\begin{aligned}x \pm^* y &= (x \pm y)(1 + \varepsilon_1); \\x \times^* y &= (x \times y)(1 + \varepsilon_2); \\x /^* y &= (x / y)(1 + \varepsilon_3); \\ \text{donde cada } |\varepsilon_i| &\leq 5 \times 10^{-t}, \quad i = 1, 2, 3\end{aligned}$$

Ejemplo 2:

La aritmética de punto flotante no es necesariamente asociativa o distributiva.

Sí $t = 8$ dígitos con,

$$\begin{aligned}a &= 0.63381158 \times 10^{-4} \\b &= 0.73688329 \times 10^2 \\c &= -0.73687711 \times 10^2\end{aligned}$$

luego,

$$\begin{aligned}a +^*(b +^*c) &= a +^* 6.18 \times 10^{-4} = 6.8138116 \times 10^{-4} = 0.68138116 \times 10^{-3} \\(a +^*b) +^*c &= 73.688392 +^*c = 6.8100 \times 10^{-4} = 0.68100000 \times 10^{-3}\end{aligned}$$

mientras que el resultado exacto es:

$$a + b + c = 0.68138116 \times 10^{-3},$$

notamos que no se satisface la propiedad asociativa.

De aquí, concluimos que hay operaciones que pueden ocasionar errores significativos, algunas de las más frecuentes son:

- (i) La división de un resultado con dígitos finitos entre un número de magnitud pequeña δ .

Supongamos que el número exacto fuese r y que se aproxima por su número de punto flotante:

$$r^* = r(1 + \varepsilon) = r + \varepsilon^*,$$

entonces la división será

$$(r + \varepsilon^*) / \delta = r / \delta + \varepsilon^* / \delta = r / \delta + \varepsilon'',$$

donde el error final $\varepsilon'' = \varepsilon^*/\delta$, crece.

(ii) La resta de números casi iguales.

Sean r_1 , r_2 y sus números de punto flotante r_1^* , r_2^* , luego la resta de éstos,

$$r_1^* - r_2^* = (r_1 + \varepsilon_1^*) - (r_2 + \varepsilon_2^*) = r_1 - r_2 + (\varepsilon_1^* - \varepsilon_2^*),$$

y por lo tanto el error relativo final será:

$$\frac{|r_1^* - r_2^* - (r_1 - r_2)|}{|r_1 - r_2|} = \frac{|\varepsilon_1^* - \varepsilon_2^*|}{|r_1 - r_2|}$$

cantidad que puede ser muy grande ya que el denominador es pequeño.

Ejemplo 3:

Supongamos que se tiene la ecuación cuadrática y sus raíces:

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

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

En el caso de que b sea positivo, y que la magnitud de $\sqrt{b^2 - 4ac}$ sea aproximadamente igual a la de b entonces $-b + \sqrt{b^2 - 4ac} \approx 0$, lo que podría generar mayor error.

Sea la ecuación cuadrática:

$$\frac{1}{3}x^2 - \frac{123}{4}x + \frac{1}{6} = 0$$

y se sabe que sus raíces exactas son:

$$x_1 = 92.24457963 \text{ y } x_2 = 5.420372688 \times 10^{-3}.$$

Trabajando con una aritmética de $t = 4$ dígitos, se obtiene:

$$0.3333x^2 - 0.3075 \times 10^2 x + 0.1667 = 0$$

$$x_{1,2} = \frac{0.3075 \times 10^2 \pm \sqrt{(0.3075 \times 10^2)^2 - 4(0.3333)0.1667}}{2 \times 0.3333}$$

$$\Rightarrow fl(x_1) = fl\left(\frac{0.3075 \times 10^2 + 0.3075 \times 10^2}{2 \times 0.3333}\right) = 92.26,$$

$$fl(x_2) = fl\left(\frac{0.3075 \times 10^2 - 0.3075 \times 10^2}{2 \times 0.3333}\right) = 0,$$

lo que lleva a una respuesta incorrecta para la raíz x_2 .

Una forma de evitar el problema de restar estos dos números casi iguales, consiste en “racionalizar” el numerador:

$$x_2 = \left(\frac{-b - \sqrt{b^2 - 4ac}}{2a} \right) \left(\frac{-b + \sqrt{b^2 - 4ac}}{-b + \sqrt{b^2 - 4ac}} \right) = \frac{2c}{-b + \sqrt{b^2 - 4ac}}$$

y aplicando este resultado al cálculo de x_2 empleando la aritmética de 4 dígitos, se tendrá:

$$fl(x_2) = fl\left(\frac{2(0.1667)}{61.5}\right) = 0.005421 = 5.421 \times 10^{-3}$$

y si comparamos este último resultado con el valor exacto, observamos de que se mejora muchísimo en la precisión.

Comentarios

Si estamos interesados en la precisión absoluta, ésta se determina mejor si los resultados se expresan con un cierto número de cifras decimales. Concretamente si redondeamos un número hasta t cifras decimales el error absoluto es a lo sumo 5 en la $(t+1)$ -ésima cifra decimal (por ejemplo 5 en la quinta cifra decimal para tener 4 decimales en el ejemplo anterior)

- Vemos así que el error absoluto es comparable al número de cifras decimales utilizadas.
- Sin embargo, si estamos interesados, en la precisión relativa, esta se podrá determinar más fácilmente si los resultados se expresan con un cierto número de cifras significativas. Concretamente, si redondeamos un número hasta t cifras significativas el error relativo es a lo sumo 5 en la t -ésima cifra significativa. Vemos por tanto que el error relativo es comparable al número de cifras significativas utilizadas.

1.4 Error Propagado

Se puede definir el error propagado como el error de salida provocado por un error en la entrada, asumiendo que todos los cálculos intermedios se efectúan exactamente (esto es, sin error de redondeo). Si consideramos una situación real, el error de salida estaría incluyendo el error propagado más todos los tipos de errores que se generen en el procedimiento.

Ejemplo 1:

Sea $f(x) = 4x^3 - 6x^2 - x + 1$ y consideremos a $x = 3.1617$ y a $x^* = 3.16$ un valor aproximado de x , luego efectuando los cálculos con una aritmética de 10 dígitos, tendremos:

$$f(3.1617) = 64.28201851 \quad \text{y} \quad f(x^*) = f(3.16) = 64.144384,$$

lo que nos lleva a un error propagado o de salida:

$$|f(x) - f(x^*)| = 0.13763451$$

y observamos que el error absoluto de entrada es:

$$|x - x^*| = 0.0017$$

por lo tanto el error de entrada crece en 80.9615 veces su valor.

Ejemplo 2:

Sea $f(x) = \pi x(1+x^2)^{-1}$ y consideremos a los valores $x = 15.4796$ y a su valor aproximado $x^* = 15.47$, luego evaluando a $f(x)$ en cada uno de estos valores, empleando una aritmética de 10 dígitos, tenemos:

$$f(15.4796) = 0.2021070509 \text{ y } f(15.47) = 0.2022314262$$

lo que nos lleva al error propagado

$$|f(15.4796) - f(15.47)| = 1.243753 \times 10^{-4}$$

y el error de entrada es:

$$|x - x^*| = 9.6 \times 10^{-3},$$

por lo que vemos que el error de salida ha disminuido en 0.0129557604 veces el valor de entrada.

Error propagado para las funciones de una variable

Para determinar el efecto de un error individual sobre la respuesta final, se puede, en principio proceder como sigue: sea x el número en el cual ocurre el error y consideramos la respuesta final como dependiente (una función) de x , es decir $f(x)$. Si $f(x)$ es una función suave entonces se puede usar la Serie de Taylor:

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{1}{2} f''(x^*)(x - x^*)^2 + \dots$$

$$\Rightarrow f(x) - f(x^*) \approx f'(x^*)(x - x^*)$$

$$\therefore |f(x) - f(x^*)| \approx |f'(x^*)| |x - x^*|$$

Entonces el error propagado: $|f(x) - f(x^*)|$ en $f(x^*)$ es proporcional al error de entrada $|x - x^*|$ en x^* , con el factor de proporcionalidad igual a $f'(x^*)$.

Así, sí

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}} \Rightarrow$$

$$\text{el error relativo} = \frac{|f(x) - f(x^*)|}{|f(x)|} \approx |f'(x^*)| \frac{|x - x^*|}{|f(x)|} \approx \frac{1}{2} \frac{|x - x^*|}{\sqrt{x}\sqrt{x^*}} \approx \frac{1}{2} \frac{|x - x^*|}{|x|}$$

es decir que el error relativo en $f(x^*)$ es aproximadamente la mitad del error relativo en x^* , sin tener en cuenta la magnitud de x . Entonces podemos afirmar que la operación “raíz cuadrada” es una operación “segura”(estable) desde el concepto del error relativo.

En contraste con la función:

$$f(x) = \frac{10}{1 - x^2} \Rightarrow f'(x) = \frac{20x}{(1 - x^2)^2} \text{ y el error relativo es :}$$

$$\frac{|f(x) - f(x^*)|}{|f(x)|} \approx \frac{2|x^*||x|}{|1 - (x^*)^2|} \frac{|x - x^*|}{|x|}$$

En este caso el error relativo en $f(x^*)$ es considerablemente mayor que el error relativo en x^* , si $|x|$ está cercano de 1.

Ejemplo 3:

Sea

$$f(x) = 4x^3 - 6x^2 - x + 1$$

y su derivada

$$f'(x) = 12x^2 - 12x - 1,$$

si $x = 3.1617$ y $x^* = 3.16$ es un valor aproximado entonces se tiene:

$$f'(x^*) = f'(3.16) = 12(3.16)^2 - 12(3.16) - 1 = 80.9072$$

y el error propagado aproximado es:

$$|f(x) - f(x^*)| \approx 80.9072 |x - x^*| = 0.13754224$$

Ejemplo 4:

Sea $f(x) = (\pi x)/(1+x^2)$, y su primera derivada es:

$$f'(x) = \pi \left[\frac{1-x^2}{(1+x^2)^2} \right],$$

luego, si $x = 15.4796$ y $x^* = 15.47$ es un valor aproximado entonces se tiene:

$$f'(x^*) = f'(15.47) = -0.01296369843$$

el error propagado aproximado es:

$$|f(x) - f(x^*)| \approx |f'(15.47)| |x - x^*| = 1.24451504 \times 10^{-4}$$

y el error relativo aproximado:

$$\begin{aligned} \frac{|f(x) - f(x^*)|}{|f(x)|} &\approx \frac{|f'(x^*)|}{|f(x)|} |x - x^*| = \frac{\left| \pi \left[\frac{1 - (x^*)^2}{(1 + (x^*)^2)^2} \right] \right|}{\left| \pi \left[\frac{x}{1 + x^2} \right] \right|} |x - x^*| \\ &\approx \frac{|1 - (x^*)^2|}{(1 + (x^*)^2)} \frac{|x - x^*|}{|x|} \end{aligned}$$

En este caso el error relativo en $f(x^*)$ es aproximadamente igual que el error relativo en x^* si x toma valores grandes, y por esto consideramos que es una operación estable.

Crecimiento del error

Basándonos en los ejemplos anteriores, observamos que los errores independientemente de la fuente de la cual provengan, pueden propagarse de diferente forma. En algunos casos realmente disminuirán y afectarán muy poco la exactitud, mientras que en otros casos, pueden

crecer de tal manera que los resultados serán incorrectos y por lo tanto no válidos.

Representemos al crecimiento del error ε , después de n operaciones como $C_n(\varepsilon)$.

Si asumimos que $C_n(\varepsilon)$ toma la forma:

$$C_n(\varepsilon) = an\varepsilon, \text{ donde } a \text{ es una constante que no depende de } n,$$

diremos que el crecimiento del error ε , es de tipo lineal; si más bien tomase la forma:

$$C_n(\varepsilon) = ab^n\varepsilon, \text{ donde } b \text{ es una constante mayor que } 1,$$

diremos que el crecimiento del error es de tipo exponencial.

Como podemos apreciar, en el caso de ser $C_n(\varepsilon)$ de tipo lineal, los resultados serán aproximados al valor exacto y por lo tanto aceptables, en este caso diremos que este procedimiento es estable; si más bien $C_n(\varepsilon)$ toma la forma de tipo exponencial, el crecimiento del error será grande y por lo tanto afectará a los resultados de tal manera que no serán válidos, en este caso diremos que es un procedimiento inestable.

Ejemplo 5:

Veamos el caso de la sucesión $\left\{ \frac{1}{3^n} \right\}_{(n \geq 0)}$, que se puede construir a partir

de las siguientes relaciones recurrentes:

(a) Sea la relación:

$$x_n = \frac{5}{6}x_{n-1} - \frac{1}{6}x_{n-2}; \quad n = 2, 3, 4, \dots$$

$$\text{con } x_0 = 1, \quad x_1 = \frac{1}{3}$$

redondeando a 5 dígitos, tenemos:

n	x_n	$\frac{1}{3^n}$
2	0.11110	0.11111
3	0.37027×10^{-1}	0.37037×10^{-1}
4	0.12339×10^{-1}	0.12346×10^{-1}
5	0.41107×10^{-2}	0.41152×10^{-2}
6	0.13691×10^{-2}	0.13717×10^{-2}
7	0.45577×10^{-3}	4.5725×10^{-4}
8	0.15162×10^{-3}	1.5242×10^{-4}

Tabla 1

Apreciamos que los resultados son aproximados a los valores exactos, y consideramos a este proceso estable.

(b) En el caso de la relación recurrente,

$$x_n = \frac{5}{3}x_{n-1} - \frac{4}{9}x_{n-2}; \quad n = 2, 3, 4, \dots$$

$$\text{con } x_0 = 1, \quad x_1 = \frac{1}{3},$$

n	x_n	$\frac{1}{3^n}$
2	0.11112	0.11111
3	0.37050×10^{-1}	0.37037×10^{-1}
4	0.12365×10^{-1}	0.12346×10^{-1}
5	0.4142×10^{-2}	0.41152×10^{-2}
6	0.14080×10^{-2}	0.13717×10^{-2}
7	0.5058×10^{-3}	4.5725×10^{-4}
8	0.21725×10^{-3}	1.5242×10^{-4}

Tabla 2

Notamos que el error crece demasiado, por lo que es un procedimiento inestable.

1.5 Convergencia de Sucesiones

El concepto de convergencia que es un proceso infinito, está presente en el cálculo y más aún en el análisis, a través de las nociones básicas como la derivada, la integral, la continuidad y la representación de funciones por medio de series convergentes. Pero, para los problemas de la ciencia y la ingeniería, rara vez se necesitan respuestas numéricas exactas. Más bien se requiere de una aproximación a la respuesta que sea precisa hasta un cierto número de decimales, o precisa dentro de una tolerancia ϵ .

Por lo tanto, si tenemos un método numérico que genera una sucesión $\alpha_1, \alpha_2, \alpha_3, \dots$ de números que converja a la respuesta pedida α de un problema dado, según la tolerancia que queramos bastará con calcular hasta un número determinado de términos.

Recordemos la definición de convergencia de sucesiones:

Una sucesión $\{\alpha_n\}_{(n \geq 0)}$ de números converge a α si y sólo si, $\forall \varepsilon > 0$, existe un entero $N(\varepsilon)$ tal que $\forall n \geq N(\varepsilon)$, $|\alpha_n - \alpha| < \varepsilon$.

Por lo que podemos acercarnos a una solución α , tan cerca como quisiéramos.

Por ejemplo, sean las sucesiones:

(i) $\alpha_n = (n-3)/n, \forall n \geq 1$

n	α_n
1	-2
2	-1/2
3	0
4	0.25
5	0.4
6	0.5
7	0.571428
8	0.625
9	0.66666666

Tabla 3

observamos que converge lentamente a su límite 1.

(ii) $\beta_n = \frac{\text{sen}(1/n)}{(1/n)}, \forall n \geq 1$.

n	β_n
1	0.84147
2	0.9588510772
3	0.9815840904
4	0.989615837
5	0.993346654
6	0.9953767962
7	0.9966021085
8	0.99739786

Tabla 4

y descubrimos que converge más rápidamente a su límite 1, que la sucesión $\{\alpha_n\}_{(n \geq 0)}$.

(iii) Consideremos a la siguiente sucesión que se define recursivamente:

$$\gamma_{n+1} = \frac{2}{3}\gamma_n + \frac{1}{\gamma_n^2}, \quad \forall n \geq 1$$

con $\gamma_1 = 3$

n	γ_n
1	2.1111111
2	1.631784139
3	1.463411989
4	1.442554125
5	1.442249635
6	1.44224957
7	$\sqrt[3]{3} = 1.44224957$

Tabla 5

y vemos que la rapidez de convergencia de esta sucesión es aún más rápida que la de β_n , ya que en 6 iteraciones encontramos su límite.

Estos ejemplos nos permiten apreciar que hay sucesiones que convergen con mayor rapidez a su límite que otras. Este concepto de rapidez de convergencia, tiene un papel importante en el cálculo numérico, ya que una de las cualidades con las que deberían contar los métodos numéricos es la de que su aplicación proporcione una solución aproximada y confiable; y si consideramos el error de redondeo por la aritmética de la computadora, al calcular un número mayor de términos para lograr la precisión deseada, lo que se podría lograr es que el error de crecimiento aumente y por lo tanto el método numérico no sería el conveniente.

Para analizar este tipo de problemas; es decir medir cuán rápido converge la sucesión, lo haremos por comparación con otras sucesiones. Así, diremos que una sucesión $\{\alpha_n\}_{(n \geq 0)}$, tiene rapidez de convergencia

del orden de $\{\beta_n\}_{(n \geq 0)}$, donde β_n es otra sucesión, tal que $\beta_n \neq 0, \forall n \geq 0$, si se satisface:

$$\frac{|\alpha_n - \alpha|}{|\beta_n|} \leq k,$$

para n suficientemente grande, donde k es una constante positiva que no depende de n , y lo denotamos como:

$$\alpha_n = \alpha + O(\beta_n).$$

Por ejemplo, en el caso de la sucesión dada en (i), sabemos que su límite es 1, y si definimos a la sucesión $\beta_n = \frac{1}{n}, \forall n \geq 1$, tenemos:

$$\frac{|\alpha_n - 1|}{|\beta_n|} = \frac{\left| \frac{n-3}{n} - 1 \right|}{\left| \frac{1}{n} \right|} \leq 3, \quad \forall n \geq 1$$

Por lo tanto, $\alpha_n = 1 + O(1/n)$, es decir la sucesión $\{\alpha_n\}_{(n \geq 1)}$ tiene una rapidez de convergencia equivalente a la de $\{1/n\}_{(n \geq 1)}$.

En el caso de la sucesión $\{\beta_n\}_{(n \geq 1)}$ del ítem (ii), notamos que su límite es 1. De manera de que si expresamos a la función $f(x) = \text{sen}(x)$, a través de la fórmula de Taylor, en torno de $x = 0$, para $N = 2$, tendremos:

$$\begin{aligned} \text{sen } x &= x - f'''(\xi(x)) \frac{x^3}{3!} = x - \text{sen}(\xi(x)) \frac{x^3}{3!}; \text{ con } \xi(x) \in]0, x[\\ &\Rightarrow \frac{|\text{sen}(x) - x|}{|x^3|} = \frac{|\text{sen}(\xi(x))|}{3!} \\ &\Rightarrow \frac{\left| \frac{\text{sen}(x)}{x} - 1 \right|}{x^2} = \frac{|\text{sen}(\xi(x))|}{3!} \leq \frac{\max_{[0,x]} |\text{sen } x|}{3!} \leq \frac{1}{6}, \quad \forall x \in]0, x[; \end{aligned}$$

si definimos $x = 1/n$, y asociamos a este último resultado con la definición de rapidez de convergencia, tenemos:

$$\frac{|\beta_n - 1|}{|\hat{\beta}_n|} \leq 1/6,$$

donde $\hat{\beta}_n = \frac{1}{n^2}$, por lo que concluimos que:

$$\beta_n = 1 + O(1/n^2).$$

De esto, deducimos de que la sucesión $\{\beta_n\}_{(n \geq 0)}$, tiene mayor rapidez de convergencia que la sucesión $\{\alpha_n\}_{(n \geq 0)}$.

1.6 Ejercicios

1. a. Obtenga la serie de Taylor en $x = 0$ para la función $f(x) = \ln(x+1)$. Emplee el polinomio de tercer grado para aproximar $f(0.5)$ y encuentre una cota para el error en esta aproximación.
- b. Obtenga la serie de Taylor en $x=0$ para la función $f(x) = \cos x$. Emplee el polinomio de cuarto grado para aproximar $f(\pi/3)$ y encuentre una cota para el error en esta aproximación.

- c. Obtenga la serie de Taylor en torno a $x = 1$ para la función $f(x) = e^{x-1}$. Emplee el polinomio de quinto grado para aproximar $f(2)$ y encuentre una cota para el error en esta aproximación.
2. Resuelva los siguientes sistemas de ecuaciones:
- a. $21.7x + 24.34y = 1.24$ b. $1.013x - 6.099y = 14.22$
 $14.16x + 15.84y = 1.15$ $-18.11x + 112.2y = -0.1376$
- empleando una aritmética de punto flotante de $t=3$ dígitos.

3. a. Resuelva la siguiente ecuación cuadrática:

$$x^2 - 140.5x + 2.5 = 0$$

empleando una aritmética de punto flotante de cuatro dígitos. Analizar los resultados obtenidos.

- b. Resuelva la siguiente ecuación cuadrática:

$$2x^2 - 435.6x + 6.1 = 0$$

empleando una aritmética de punto flotante de cuatro dígitos. Analizar los resultados obtenidos.

4. Considere las ecuaciones:

$$31.69x + 14.31y = 45.00 \quad \dots \quad (1)$$

$$13.11x + 5.89y = 19.00 \quad \dots \quad (2)$$

La solución única a este sistema de ecuaciones es

$$x = 7.2, \text{ e } y = -12.8.$$

Un método que se presenta generalmente en cursos elementales de álgebra para resolver problemas de este tipo es multiplicar la

ecuación (1) por el coeficiente de x en la ecuación (2), y la ecuación (2) por el coeficiente de x en la ecuación (1), y después restar las ecuaciones resultantes.

- a. Encuentre los valores para x e y , haciendo uso de la aritmética de punto flotante con 4 dígitos.
 - b. Analice por qué las respuestas obtenidas en (a) difieren considerablemente de los valores reales de x e y .
5. a. Halle la serie de Taylor en torno a $x = 0$, para la función,

$$f(x) = \tan(x)$$

Emplee el polinomio de Taylor de grado 3, para estimar $f(0.5)$ y encuentre una cota para el error en esta aproximación.

- b. Sea la sucesión:

$$\alpha_n = \frac{\tan(1/n)}{(1/n)}, \quad \forall n \geq 1$$

haciendo uso de la serie de Taylor del ítem (a), halle la rapidez de convergencia de la sucesión $\{\alpha_n\}$.

6. a. Halle la serie de Taylor en torno a $x=0$, para la función,

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

Emplee el polinomio de Taylor de grado 4, para estimar $f(0.5)$ y encuentre una cota para el error en esta aproximación.

- b. Sea la sucesión:

$$\alpha_n = \frac{1 - \cos(1/n)}{(1/n)}, \quad \forall n \geq 1$$

haciendo uso de la serie de Taylor del ítem (a), halle la rapidez de convergencia de la sucesión $\{\alpha_n\}$.

7. Encuentre el valor óptimo para k en la afirmación

$$\cos(1/n) - 1 + (1/n)^2/2 = O((1/n)^k)$$

cuando $n \rightarrow \infty$.

8. Demuestre que para cualquier $r > 0$, $n^r = O(e^n)$ cuando $n \rightarrow \infty$.
9. Demuestre que para toda $r > 0$, $\ln(n) = O(n^r)$ cuando $n \rightarrow \infty$.
10. a. Demuestre que si $x_n = O(\alpha_n)$, entonces $cx_n = O(\alpha_n)$.
- b. Demuestre que si $x_n = O(\alpha_n)$, e $y_n = O(\alpha_n)$, entonces $x_n + y_n = O(\alpha_n)$.

Ecuaciones Algebraicas No Lineales

Este capítulo se ocupa del problema de determinar raíces de ecuaciones (o ceros de funciones). Es un problema que aparece con frecuencia tanto en el área de ciencias como de ingeniería, ya sea como un fin en sí mismo, o bien como un paso intermedio al resolver un problema más complejo.

Así, se pueden presentar situaciones en las que necesitamos resolver ecuaciones no lineales, desde su forma más elemental como una ecuación cuadrática de la forma:

$$ax^2 + bx + c = 0, (a \neq 0)$$

en la cual, sus raíces exactas son:

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

o ecuaciones simples de la forma:

- (i) $e^x - 3x = 0$
- (ii) $\tan x - 2x = 0$
- (iii) $x^9 - 7x^4 - 1 = 0$

o por ejemplo en Física, sabemos que la Ley de los gases ideales se comporta de acuerdo a la relación:

$$PV = nRT$$

y una generalización de ésta es:

$$\left[P + \frac{a}{V^c} \right] (V - b) = nRT$$

donde R es la constante de los gases y a , b , c , son parámetros que dependen del gas, P es la presión y T es la temperatura absoluta. Si se conociesen P y T , ¿Cómo encontraríamos el volumen V de moles de gas?

Todas las ecuaciones mencionadas anteriormente no pueden ser resueltas en un número finito de pasos y en realidad, esto es cierto para las ecuaciones algebraicas no lineales en general. En estos casos, la única alternativa es una técnica de solución aproximada.

Un método para obtener una solución aproximada consiste en graficar la función $f(x)$ y determinar donde corta al eje X , este punto será la raíz de $f(x) = 0$. Aunque los métodos gráficos son útiles en la obtención de estimaciones aproximativas de las raíces, están limitadas por la carencia de precisión.

Los métodos para obtener soluciones aproximadas, consisten generalmente en comenzar con un valor o rango de valores estimados de la solución y determinar posteriormente un cierto número de aproximaciones a la solución cada vez mejores.

El problema general, planteado en el caso más sencillo de una función definida en los números reales y cuya imagen está en el conjunto de los reales, es el siguiente:

Dada una función $f: \mathcal{R} \rightarrow \mathcal{R}$, encontrar los valores de x para los cuales $f(x) = 0$.

2.1 Separación de Raíces

El problema de hallar las raíces de ecuaciones o ceros de funciones, se resuelve en varias etapas:

La 1^{ra} Etapa se ocupa de separar los ceros de $f(x)$, y consiste en determinar el intervalo $[a, b]$, dentro del cual se encuentra un solo cero de $f(x)$. Esta etapa se realiza con ayuda del Teorema del Valor Intermedio (ver Apéndice A), que nos asegura de que si $f(a)f(b) < 0$, entonces hay un punto $\alpha \in]a, b[$, tal que $f(\alpha)=0$ y por lo tanto $x = \alpha$ es un cero de $f(x)$ ó una raíz de la ecuación $f(x) = 0$.

La 2^{da} Etapa consiste en construir una sucesión $x_n \in [a, b]$ tal que $\lim_{n \rightarrow \infty} x_n = \alpha$ ($x_n \rightarrow \alpha$). Truncando luego este proceso infinito, podemos obtener $x_n \approx \alpha$, es decir x_n para algún n suficientemente grande, como una aproximación de α ; introduciendo el error del truncamiento en el método.

Como el concepto de límite a través de su unicidad, determina un sólo cero de $f(x)$, es importante tener los ceros de $f(x)$, separados entre sí, a través de los intervalos $[a, b]$, cada uno de los cuales contiene un sólo cero de $f(x)$.

Esta tarea se realiza, generalmente en la 1^{ra} Etapa, previa a la construcción y aplicación del método numérico; y una de las formas más efectivas de realizarla, es tener la gráfica de $f(x)$, ya que los ceros de $f(x)$ son los puntos de intersección de $f(x)$ con el eje X .

En muchos casos es más conveniente presentar la ecuación $f(x) = 0$ en su forma equivalente: $f_1(x) = f_2(x)$. Entonces los ceros de $f(x)$ se convierten en las abscisas de los puntos de intersección de las dos curvas $f_1(x)$ y $f_2(x)$.

Aplicaciones

Ejemplo 1

Si $f(x) = e^x - \text{sen } x$

y α es un cero de $f(x)$, entonces α satisface a la ecuación

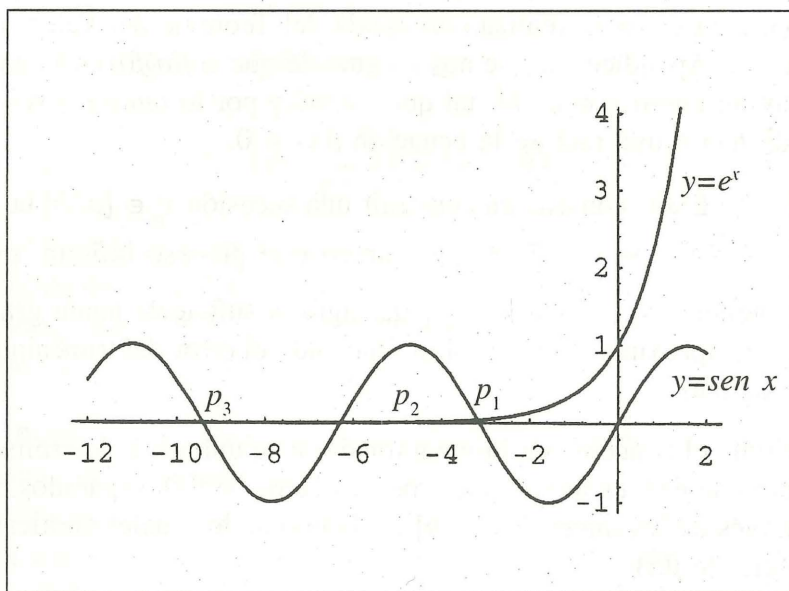
$$e^x - \text{sen } x = 0 \quad \text{ó} \quad e^x = \text{sen } x.$$

Si construimos a las gráficas de las dos curvas

$$y = f_1(x) = e^x \quad \text{e} \quad y = f_2(x) = \text{sen } x;$$

uno de sus puntos de intersección es precisamente (α, β) , con α cero de $f(x)$.

Graficando:



Gráfica 1

De la gráfica podemos observar que las curvas se interceptan sólo en el segundo cuadrante, y que hay infinitos puntos de intersección: $P_1(\alpha_1, \beta_1)$; $P_2(\alpha_2, \beta_2)$; ... lo que implica que nuestra función tiene infinitos ceros: $\alpha_1, \alpha_2, \dots$, todos negativos. No hay ningún cero positivo.

La gráfica también nos indica como podemos separar los ceros de $f(x)$. Así:

$$\alpha_1 \in \left(-\frac{3\pi}{2}, -\pi\right);$$

$$\alpha_2 \in \left(-2\pi, -\frac{3\pi}{2}\right);$$

$$\alpha_3 \in \left(-\frac{7\pi}{2}, -3\pi\right); \dots etc.$$

y si nos fijamos en el intervalo $[-8, -4]$, vemos que está incluido en el intervalo $[-3\pi, -3\pi/2]$, en el cual hay un sólo cero de $f(x)$.

Ejemplo 2:

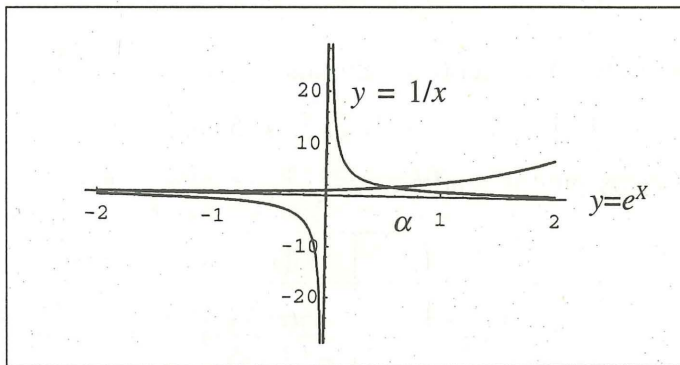
Sea la ecuación,

$$f(x) = xe^x - 1 = 0 \Rightarrow e^x = 1/x$$

y de la intersección de las gráficas de las curvas:

$$y = f_1(x) = e^x \quad e \quad y = f_2(x) = 1/x$$

se observa que hay un sólo cero α de $f(x)$.



Gráfica 2

Se puede también analizar el comportamiento de la función en el intervalo de interés, una vez que aseguramos que hay por lo menos un cero en este intervalo.

Así, sí

$$\begin{aligned} f(x) &= xe^x - 1; \\ f(0) &= -1; \\ f(1) &= e - 1 > 0, \\ \Rightarrow f(0) f(1) &< 0, \text{ hay por lo menos un cero en } [0, 1] \end{aligned}$$

Para asegurar que este cero es único vemos el comportamiento de la derivada en este intervalo.

Como

$$f'(x) = e^x + xe^x > 0, \forall x \in]0, 1[$$

$f(x)$ es una función estrictamente creciente en este intervalo y por lo tanto no puede tener otro punto de intersección con el eje X . Lo que comprueba que en el intervalo $]0, 1[$, hay un sólo cero de $f(x)$.

Ejemplo 3:

Otra forma de ubicar la raíz de $f(x)$, consiste en tabular un conjunto de puntos, observando los cambios de signo de la función $f(x)$.

Sea la ecuación,

$$f(x) = -19(x - 1/2)(x - 1) + e^x - e^{-2x},$$

y seleccionamos un conjunto de puntos:

$$\{-3, -2, -1, 0, 1, 2, 3, 4, 5, 6, 7, \dots\}$$

Tabulando estos puntos tendremos:

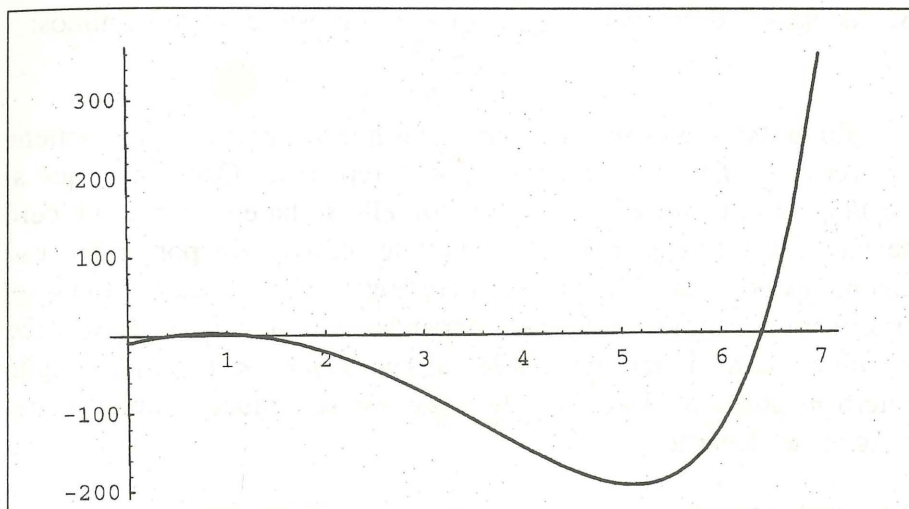
x	$f(x)$
-3	-669.38
-2	-196.96
-1	-64.02
0	-9.5
1	2.583
2	-21.129
3	-74.917
4	-144.9
5	-193.59
6	-119.07
7	355.63

Tabla 1

Observamos que la función cambia de signo en los intervalos:

$$[0, 1], [1, 2], [6, 7],$$

y si hacemos uso del Teorema del Valor Intermedio, concluiremos que en cada uno de estos intervalos hay por lo menos una raíz. Lo que podemos verificar a través de la siguiente gráfica:



Gráfica 3

Para realizar la 2da. etapa, necesitamos seleccionar y aplicar un método numérico. A continuación veremos algunos de ellos.

2.2 Método de Bisección

Si $f(x)$ es una función continua sobre el intervalo $[a_0, b_0]$ y si $f(a_0)f(b_0) < 0$, la función $f(x)$ cambia de signo en el intervalo $[a_0, b_0]$ y, por lo tanto, tiene por lo menos un cero en el intervalo. El método de bisección hace uso de esta idea de la siguiente forma:

Calculamos

$$x_0 = (a_0 + b_0) / 2$$

y averiguamos si

$$f(a_0)f(x_0) < 0.$$

Si lo es entonces $f(x)$ tiene un cero en $[a_0, x_0]$. A continuación redefinimos:

$$\begin{aligned} a_1 &= a_0 \\ b_1 &= x_0 \end{aligned}$$

y comenzamos con el nuevo intervalo $[a_1, b_1]$, cuya longitud es igual a la mitad del intervalo original.

Si $f(a_0)f(x_0) > 0$ entonces $f(x_0)f(b_0) < 0$ en este caso redefinimos:

$$\begin{aligned} a_1 &= x_0 \\ b_1 &= b_0 \end{aligned}$$

En ambos casos se ha generado un nuevo intervalo que contiene un cero de $f(x)$, y el proceso puede repetirse. Claro está que si $f(a_0)f(x_0) = 0$, entonces $f(x_0) = 0$ y con ello se ha encontrado un cero de $f(x)$. Sin embargo por los errores de redondeo o por el proceso mismo, es poco factible que en la computadora $f(x_0)$ sea cero. Así, el criterio para concluir no deberá depender de que $f(x_0) = 0$. Se debe permitir una tolerancia razonable, tal como $|f(x_0)| < 10^{-m}$, para algún entero m positivo. El método de bisección se conoce también como método de Bipartición.

Algoritmo de Bisección

Objetivo: Hallar una raíz α con una tolerancia ε especificada de una función continua $f(x)$ en un intervalo $[a, b]$.

Paso 1: Encontrar un intervalo $[a_0, b_0]$ incluido en $[a, b]$ tal que

$$f(a_0)f(b_0) \leq 0.$$

Para $n = 0, 1, 2, 3, \dots$ hacer los pasos (2) y (3):

Paso 2: Calcular $f(x_n)$, siendo $x_n = (a_n + b_n)/2$.

Definir:

$$\begin{aligned} a_{n+1} &= a_n \\ b_{n+1} &= x_n \\ \text{si } f(a_n)f(x_n) &< 0; \end{aligned}$$

en caso contrario definir:

$$\begin{aligned} a_{n+1} &= x_n \\ b_{n+1} &= b_n \end{aligned}$$

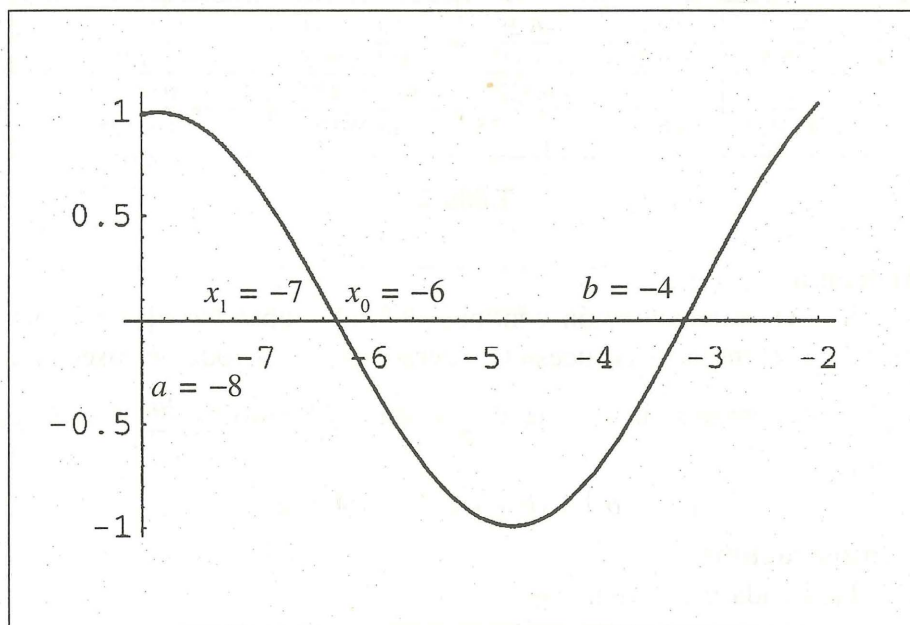
Paso 3: Averiguar si se satisface la tolerancia ε especificada*, si fuese así se toma como solución aproximada $\alpha \approx x_{n+1}$.

***Nota:** Para detener el algoritmo se pueden utilizar cualesquiera de los siguientes criterios:

1. $|x_{n+1} - x_n| < \varepsilon$
2. $\frac{|x_{n+1} - x_n|}{|x_{n+1}|} < \varepsilon$
3. $|f(x_n)| < \varepsilon$
4. Fijar un número máximo de iteraciones $n = N$, para interrumpir el algoritmo.

En la siguiente gráfica se ilustra el método, para la función:

$$f(x) = e^x - \text{sen}x, x \in [-8, -4]$$



Gráfica 4

Como $f(-8) = 0.98969371$, y $f(-4) = -0.7384868564$ entonces:

- Si $n = 0$, se tiene que $a_0 = -8$, $b_0 = -4$ y el punto medio $x_0 = -6$, evaluando $f(-6) = -0.276936746$, y como $f(a_0) f(x_0) < 0$ entonces:

$$a_1 = a_0 = -8, b_1 = x_0 = -6$$

- Si $n = 1$, $x_1 = -7$ y $f(-7) = 0.6578984807$, y como $f(x_1) f(b_1) < 0$ entonces:

$$a_2 = x_1 = -7 \quad y \quad b_2 = b_1 = -6.$$

- Si $n = 2$, se tiene que $x_2 = -6.5$ y $f(-6.5) = 0.2166234273$, y como $f(x_2) f(b_2) < 0$, entonces:

$$a_3 = -6.5, \quad b_3 = -6,$$

y así sucesivamente, en el siguiente cuadro se muestran algunas iteraciones más de la aplicación de este algoritmo:

n	a_n	b_n	x_n	$f(x_n)$	$TOL = x_{n+1} - x_n $
3	-6.5	-6	-6.25	-0.031248	0.25
4	-6.5	-6.25	-6.375	0.0933893	0.125
5	-6.375	-6.25	-6.3125	0.0311239	0.0625
6	-6.3125	-6.25	-6.28125	-6.42×10^{-5}	0.03125

Tabla 2

Teorema

Si $f(x)$ es una función continua definida sobre $[a, b]$ y tal que $f(a)f(b) < 0$, entonces la sucesión generada por el método de Bisección $\{x_n\}_{(n>0)}$ converge a un cero α , de $f(x)$ en $[a, b]$; esto es $\lim_{n \rightarrow \infty} x_n = \alpha$ y además:

$$|x_n - \alpha| < (b - a) / 2^{n+1}, \quad \forall n \geq 0$$

Demostración:

Para cada $n \geq 0$, tenemos:

$$b_n - a_n = \frac{1}{2^n} (b - a), \quad y \quad \alpha \in]a_n, b_n[.$$

Ya que $x_n = (a_n + b_n)/2$, para toda $n \geq 0$, se sigue que

$$|x_n - \alpha| \leq (b_n - a_n)/2 = \frac{b-a}{2^{n+1}} \quad \square$$

Aplicaciones:

Ejemplo 1:

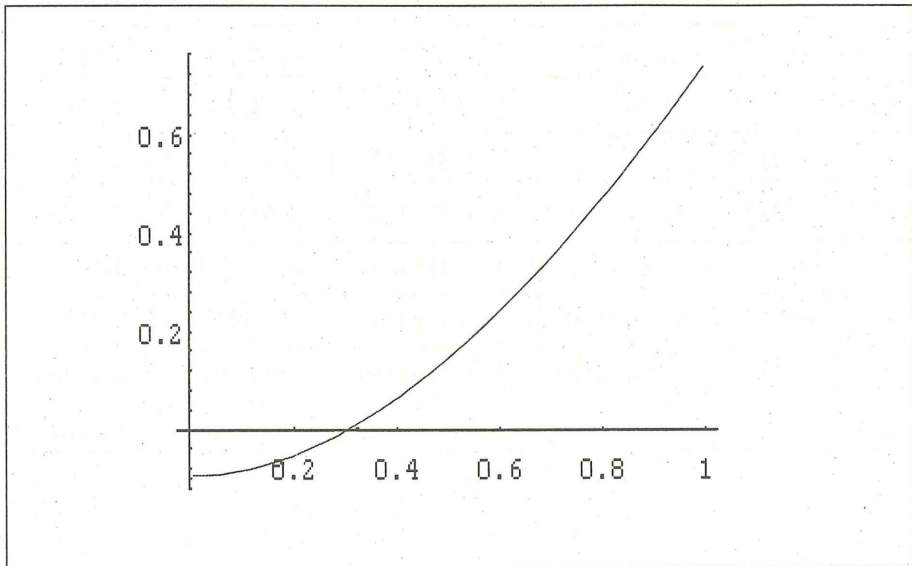
Encontrar un cero de:

$$x \operatorname{sen} x - 0.1 = 0$$

en el intervalo $[0, 1]$, mediante el método de bisección, con una tolerancia de 0.001.

Solución:

La gráfica de $f(x) = x \operatorname{sen} x - 0.1$ es:



Gráfica 5

- Si $n = 0$, se tiene:

$x_0 = 0.5$, entonces $f(0.5) = 0.139713 > 0$, y como:

$$f(a_0) = f(0) = -0.1 \text{ y } f(b_0) = f(1) = 0.741471$$

se eligen: $a_1 = 0$ y $b_1 = 0.5$.

- Si $n = 1$, $x_2 = 0.25$, y como $f(0.25) = -0.038149 < 0$ y $f(0.5) > 0$, entonces:

$$a_2 = 0.25 \quad \text{y} \quad b_2 = 0.5$$

y de esta forma se calculan los siguientes intervalos, que se muestran en la tabla a continuación:

n	a_n	b_n	x_n	$f(x_n)$	$TOL = \frac{b-a}{2^{n+1}}$
0	0	1	0.5	0.139713	1/2
1	0	0.5	0.25	-0.038149	1/4
2	0.25	0.5	0.375	0.0373523	1/8
3	0.25	0.375	0.3125	-3.92×10^{-3}	1/16
4	0.3125	0.375	0.34375	0.01585	1/32
5	0.3125	0.34375	0.328125	5.74×10^{-3}	1/64
6	0.3125	0.328125	0.3203125	8.54×10^{-4}	1/128
7	0.3125	0.3203125	0.3164062	-1.54×10^{-3}	1/256
8	0.3164062	0.3203125	0.3183593	-3.50×10^{-4}	1/512
9	0.3183593	0.3203125	0.3193359	2.51×10^{-4}	1/1024 = 9.77×10^{-4}

Tabla 3

Como $|x_9 - \alpha| \leq \frac{b-a}{2^{n+1}} = 9.77 \times 10^{-4}$, x_9 será una solución aproximada con dos decimales exactos por lo menos.

Ejemplo 2:

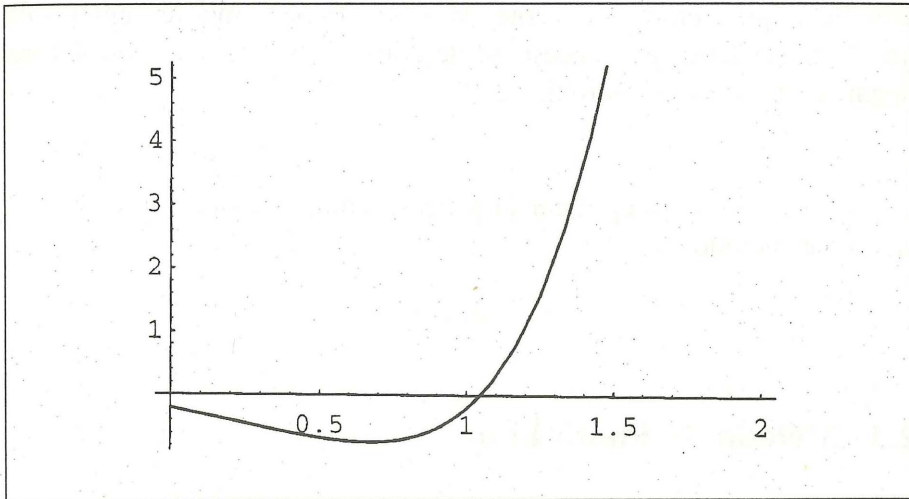
La ecuación

$$x^5 - x - 0.2 = 0,$$

tiene un cero real en $[1, 2]$. Se quiere hallar este cero con 4 cifras exactas, utilizando el método de bisección ¿Cuántas iteraciones se requiere para tal efecto?

Solución:

La gráfica de $f(x) = x^5 - x - 0.2$ es:



Gráfica 6

Por el Teorema para el método de bisección, se sabe que el error absoluto satisface:

$$|e_n| = |x_n - \alpha| \leq (b - a) / 2^{n+1},$$

entonces el error relativo satisface:

$$\begin{aligned} |x_n - \alpha| / |\alpha| &\leq (b - a) / 2^{n+1} = 1 / 2^{n+1} < \varepsilon = 5 \times 10^{-4}, \text{ ya que } 1 < \alpha < 2 \\ &\Rightarrow n + 1 > (\ln 5 - 4 \ln 10) / (-\ln 2) \\ &\Rightarrow n > 9.9657 \\ &n \geq 10 \end{aligned}$$

Entonces $n \geq 10$, y se observa que la convergencia del método de bisección es lenta, puesto que $(b - a) / 2^{n+1} \approx 1 / 2^{n+1}$; lo que en muchos casos nos llevará a realizar un número grande de iteraciones, incluyendo los cálculos correspondientes de las cuatro sucesiones siguientes:

$$\{a_n\}, \{b_n\}, \{x_n\}, \{f(x_n)\}$$

lo cual puede resultar muy tedioso.

El método de bisección tiene como ventajas, la de ser un método sencillo y que siempre converge. Más bien tiene como desventajas el hecho de ser lento y la necesidad de construir varias sucesiones para llegar a una solución aproximada.

Con el fin de simplificar el proceso computacional, se define el siguiente método:

2.3 Método de Punto Fijo

En este método tratamos de obtener una raíz de la ecuación:

$$f(x) = 0 \dots \quad (1)$$

expresando dicha ecuación en la forma equivalente

$$x = g(x) \dots \quad (2)$$

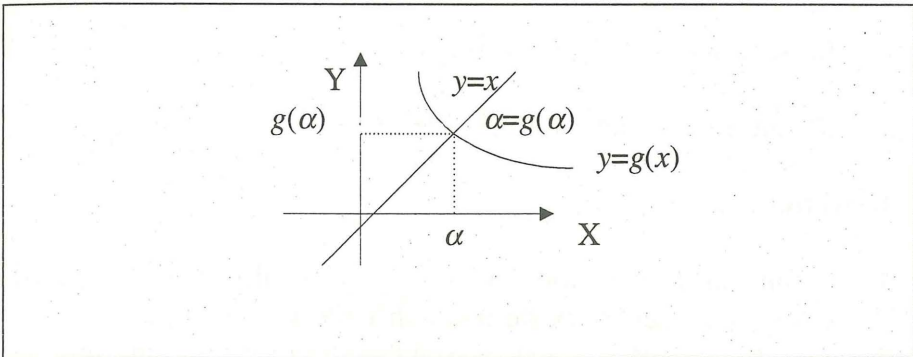
es decir, una raíz α de la ecuación (1):

$$f(\alpha) = 0$$

también es solución de la ecuación (2):

$$\alpha = g(\alpha)$$

lo que se aprecia, al graficar a las funciones x y $g(x)$:



Gráfica 7

de modo que a cualquier solución α de la ecuación se le llama punto fijo de g , ya que g asocia a α consigo misma y a su vez debe ser solución de la ecuación (1).

La ecuación (2) se puede lograr de muchas formas, por ejemplo:

Sea la ecuación,

$$f(x) = x^4 - 3x^2 - 1 = 0$$

entonces entre las posibles formas tendríamos:

- (a) $x = g_1(x) = (1 + 3x^2)^{1/4}$
- (b) $x = g_2(x) = [(x^4 - 1)/3]^{1/2}$
- (c) $x = g_3(x) = 1/(x^3 - 3x)$
- (d) $x = g_4(x) = [1/(x^2 - 3)]^{1/2}$

Cada una de estas funciones $g(x)$ recibe el nombre de función de iteración para resolver la ecuación (1).

Una vez que se escoge una función de iteración se define el algoritmo de Punto Fijo, como:

$$x_{n+1} = g(x_n), \quad \forall n \geq 0 \dots \quad (3)$$

donde x_0 es una aproximación inicial.

Para que este algoritmo sea eficaz, se debe asegurar que:

1. La sucesión $\{x_n\}_{(n \geq 0)}$ converge al punto α
2. El límite α es un punto fijo de $g(x)$.

Teorema

Sea la función de iteración $g(x) \in C_{[a, b]}$ y es tal que $g(x) \in [a, b]$, $\forall x \in [a, b]$ y además existe una constante $k < 1$ tal que $\forall x \in [a, b], |g'(x)| \leq k$, entonces $g(x)$ tiene exactamente un Punto Fijo $\alpha \in [a, b]$ y comenzando con cualquier punto x_0 en $[a, b]$, la sucesión generada por el algoritmo de punto fijo converge a α , y además se satisface:

$$|x_n - \alpha| \leq k^n \max\{x_0 - a, b - x_0\}$$

$$|x_n - \alpha| \leq \frac{k^n}{1-k} |x_1 - x_0|, \quad \forall n \geq 0$$

Demostración:

Si $g(a) = a$ ó $g(b) = b$, la existencia del punto fijo es obvia.

Supongamos que no, entonces debe ser cierto que $g(a) > a$ y $g(b) < b$.

Definamos

$$h(x) = g(x) - x$$

$h(x)$ es continua en $[a, b]$, y $h(a) = (g(a) - a) > 0$, $h(b) = (g(b) - b) < 0$, por el Teorema del Valor Intermedio existe $\alpha \in]a, b[$ tal que $h(\alpha) = 0$. Por lo tanto $g(\alpha) - \alpha = 0$ y α es un punto fijo de $g(x)$.

Supongamos además, que la desigualdad

$$|g'(x)| \leq k < 1, \quad \forall x \in]a, b[,$$

se satisface, y que α y β son puntos fijos en $[a, b]$ con $\alpha \neq \beta$

Por el Teorema del Valor Medio, existe un número ξ entre α y β , por lo tanto está en $[a, b]$ tal que:

$$|\alpha - \beta| = |g(\alpha) - g(\beta)| = |g'(\xi)| |\alpha - \beta| \leq k |\alpha - \beta| < |\alpha - \beta|, \dots \quad (4)$$

lo cual es una contradicción. Esta contradicción viene de la única suposición de que $\alpha \neq \beta$. Por lo tanto, $\alpha = \beta$ y el punto fijo en $[a, b]$ es único.

Veamos la convergencia, como $g(x) \in [a, b], \forall x \in [a, b]$, la sucesión $\{x_n\}_{(n \geq 0)} \in [a, b]$ para toda n .

Basándose en el Teorema del Valor Medio,

$$|x_n - \alpha| = |g(x_{n-1}) - g(\alpha)| \leq |g'(\xi)| |x_{n-1} - \alpha| \leq k |x_{n-1} - \alpha| \leq \dots \leq k^n |x_0 - \alpha| \quad \dots(5)$$

Como $k < 1$, tenemos

$$\lim_{n \rightarrow \infty} |x_n - \alpha| \leq \lim_{n \rightarrow \infty} k^n |x_0 - \alpha| = 0,$$

y $\{x_n\}_{(n \geq 0)}$ converge a α

De la desigualdad (5),

$$|x_n - \alpha| \leq k^n |x_0 - \alpha| \leq k^n \max \{x_0 - a, b - x_0\}, \text{ ya que } \alpha \in [a, b].$$

Por otro lado, sobre la base del procedimiento anterior para $n > 0$, se tiene que:

$$|x_{n+1} - x_n| = |g(x_n) - g(x_{n-1})| \leq k |x_n - x_{n-1}| \leq \dots \leq k^n |x_1 - x_0| \quad \dots(6)$$

Por lo tanto, para $m > n \geq 1$,

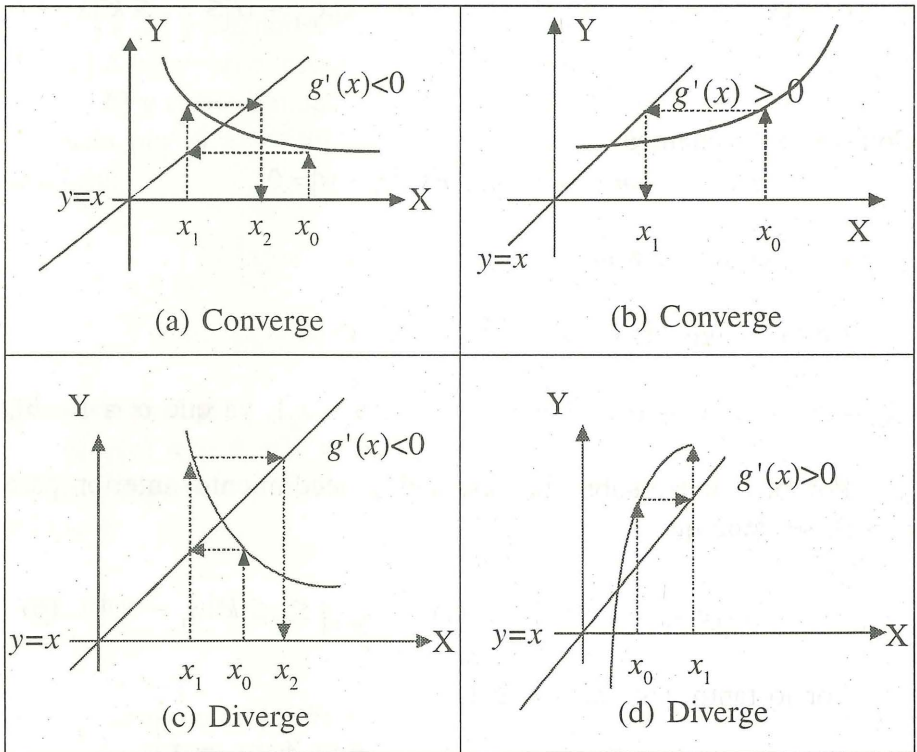
$$\begin{aligned} |x_m - x_n| &= |x_m - x_{m-1} + x_{m-1} - \dots + x_{n+1} - x_n| \\ &\leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \dots + |x_{n+1} - x_n| \\ &\leq k^{m-1} |x_1 - x_0| + k^{m-2} |x_1 - x_0| + \dots + k^n |x_1 - x_0| \\ &= k^n (1 + k + k^2 + \dots + k^{m-n-1}) |x_1 - x_0| \end{aligned}$$

Además, $\lim_{m \rightarrow \infty} x_m = \alpha$, así que

$$|x - x_n| = \lim_{m \rightarrow \infty} |x_m - x_n| \leq k^n |x_1 - x_0| \sum_{i=0}^{\infty} k^i = \frac{k^n}{1-k} |x_1 - x_0| \quad \square$$

Comentarios:

- (1) Podemos observar que en las cotas para el error absoluto aparece k^n , por lo que se deduce que en cuanto más pequeño sea k , más rápida será la convergencia. La convergencia puede ser muy lenta si k está próximo de 1.
- (2) La interpretación geométrica de la convergencia o divergencia del método para una función $g(x)$ creciente o decreciente, se puede apreciar a través de las siguientes gráficas:

**Gráfica 8**

Observamos que en los casos (a), (b) el método converge y en los casos (c) y (d) el método diverge, ya que en estos dos últimos la pendiente $|g'(x)| > 1$.

Algoritmo de Punto Fijo

Objetivo: Hallar el punto fijo α de una función $g(x)$ cero de la función $f(x)$, continua en un intervalo $[a_0, b_0]$ tal que $f(a_0)f(b_0) < 0$, con una tolerancia ε especificada.

Paso 1: Hallar una aproximación inicial x_0 de α en el intervalo $[a_0, b_0]$.

Para $n = 0, 1, 2, 3, \dots$, hacer los pasos (2) y (3).

Paso 2: Calcular $x_{n+1} = g(x_n)$.

Paso 3: Averiguar si se satisface la tolerancia especificada (empleando cualesquiera de los criterios dados), si fuese así se toma como solución $\alpha \approx x_{n+1}$.

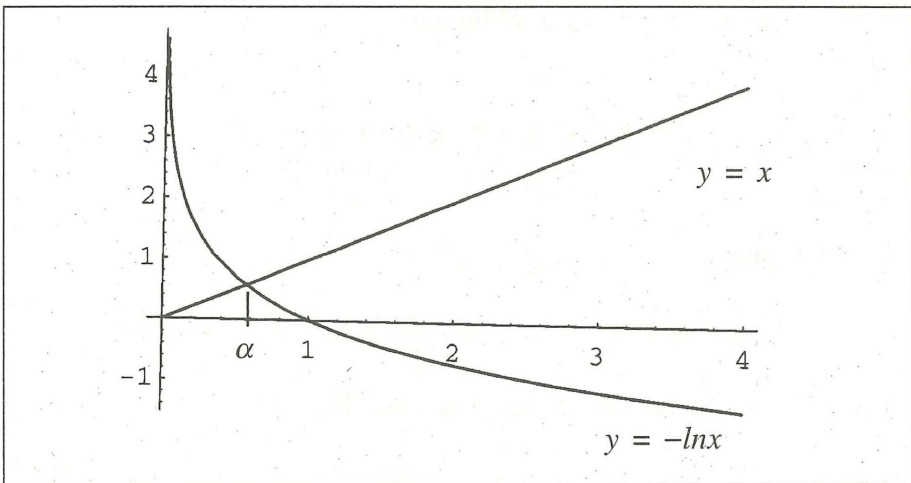
Aplicaciones

Ejemplo 1:

Usando el método de Punto Fijo, resolver la ecuación:

$$\ln x + x = 0$$

Solución: Presentando la ecuación gráficamente:



Gráfica 9

Observamos que la ecuación tiene una solución que es la intersección de las curvas: $y = x$ e $y = -\ln x$; y la solución $\alpha \in [0.5, 0.65]$.

Tomando $x = -\ln x = g(x)$ observamos que

$$|g'(x)| = |-1/x| \geq 1 \quad \forall x \in [0.5, 0.65],$$

y así elegida $g(x)$ no satisface las condiciones del teorema.

Para hallar otra función $g(x)$, la ecuación dada:

$$x + \ln x = 0$$

se puede transformar en

$$x - e^{-x} = 0 \Rightarrow x = e^{-x} = g(x).$$

Analizamos si $g(x)$ satisface las condiciones del Teorema de convergencia del método de Punto Fijo:

- a. Veamos si $g(x) \in [0.5, 0.65] \quad \forall x \in [0.5, 0.65]$, observamos que $g(x)$ es una función decreciente, por lo que calculamos sólo en los extremos del intervalo, obteniendo:

$$0.5 < e^{-0.6} \leq g(x) \leq e^{-1/2} < 0.65, \quad \forall x \in [0.5, 0.65], \\ \Rightarrow g(x) \in [0.5, 0.65].$$

- b. Además:

$$|g'(x)| = |-e^{-x}| \leq e^{-1/2} \leq 0.61 < 1, \quad \forall x \in [0.5, 0.65]$$

Lo que verifica las dos condiciones del Teorema.

Entonces se define el algoritmo de Punto Fijo:

$$x_{n+1} = g(x_n) = e^{-x_n},$$

tomando $x_0 = 0.5$ como aproximación inicial, se tiene:

n	x_n	$TOL = \frac{ x_{n+1} - x_n }{ x_{n+1} }$	n	x_n	$TOL = \frac{ x_{n+1} - x_n }{ x_{n+1} }$
0	0.5		8	0.56640945	3.581×10^{-3}
1	0.60653065	0.17563936	9	0.56755963	2.026×10^{-3}
2	0.54523921	0.11241203	10	0.56690721	1.15×10^{-3}
3	0.57970309	0.0594509	11	0.56727719	6.522×10^{-4}
4	0.56006462	0.03506464	12	0.56706735	3.700×10^{-4}
5	0.57117214	0.01944688	13	0.56718636	2.098×10^{-4}
6	0.56486294	0.01116943	14	0.56711886	1.190×10^{-4}
7	0.56843804	6.289×10^{-3}			

Tabla 4

Analizando los resultados, después de parar el algoritmo con $n = 14$, se tiene:

$$|e_{14}| = |\alpha - x_{14}| \leq (k^{14}/(1-k)) |x_1 - x_0|,$$

con $k = 0.6065 < 0.61 \Rightarrow |\alpha - x_{14}| \leq 0.000269792$,

esto nos indica una precisión de por lo menos 3 decimales exactos. Además, también se puede ver de la tabla que $|x_{14} - x_{13}| = 0.000068$, lo que nos proporciona la misma información, que

$$TOL = \frac{|x_{n+1} - x_n|}{|x_{n+1}|} = 1.190 \times 10^{-4}$$

obtenida en la tabla, en relación con la precisión obtenida.

Si continuásemos aplicando el algoritmo de Punto Fijo, se obtendrían los siguientes resultados:

n	x_n	$TOL = \frac{ x_{n+1} - x_n }{ x_{n+1} }$
15	0.5671571437	6.75×10^{-5}
20	0.5671424776	3.96×10^{-6}
25	0.5671433381	2.324×10^{-7}
30	0.5671432876	1.3576×10^{-8}
35	0.5671432906	8.82×10^{-10}
38	0.5671432904	1.76×10^{-10}
39	0.5671432904	0

Tabla 5

se puede concluir que la convergencia del método de Punto Fijo es lenta ya que se precisan de 39 iteraciones para lograr una solución casi exacta, y la rapidez de convergencia está en relación al valor $k=0.6065$, que en este caso es relativamente alto.

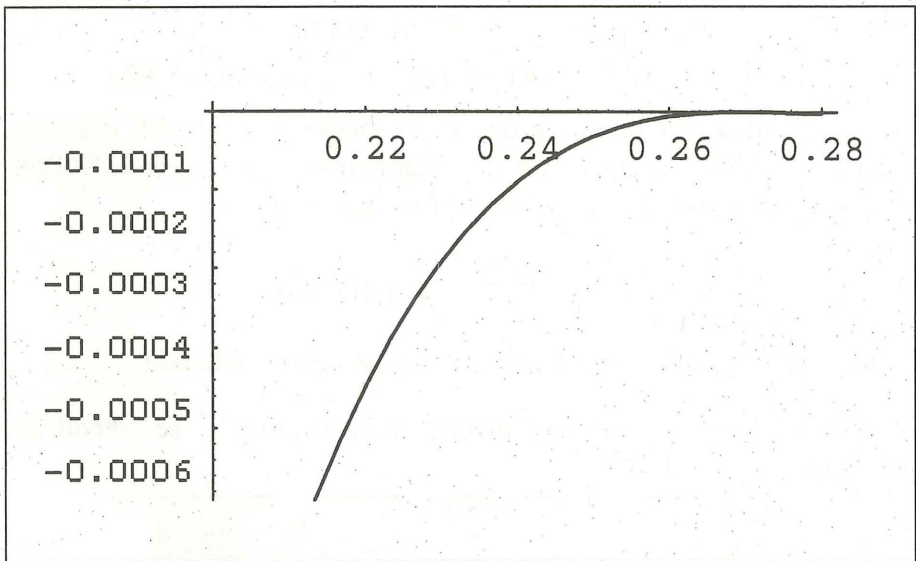
Ejemplo 2:

Empleando el método de Punto Fijo, resolver la ecuación cúbica:

$$f(x) = 10x^3 - 8.3x^2 + 2.295x - 0.21141 = 0.$$

Solución:

La gráfica de esta función se muestra a continuación:



Gráfica 10

A partir de la gráfica reparamos que $f(x)$ tiene una raíz en el intervalo $[0.26, 0.29]$.

Se halla una función $g(x)$:

$$x = g(x) = \left[\frac{(8.3x^2 - 2.295x + 0.21141)}{10} \right]^{1/3}, \text{ con } x \in [0.26, 0.29],$$

y su derivada:

$$g'(x) = \frac{16.6x - 2.295}{30 \left[\frac{8.3x^2 - 2.295x + 0.21141}{10} \right]^{2/3}}, \quad \forall x \in]0.26, 0.29[$$

Verificando las condiciones:

(a) La primera condición exige que:

$$g: [0.26, 0.29] \rightarrow B \subseteq [0.26, 0.29]$$

Ya que g es una función creciente, es suficiente con analizar las imágenes de g en los extremos del intervalo, así:

$$g(0.26) = 0.2600147921 \quad \text{y} \quad g(0.29) = 0.29,$$

(b) Y faltaría verificar la segunda condición:

$$|g'(x)| < k \leq 1, \quad \forall x \in]0.26, 0.29[$$

en este caso como $g'(x)$ es función creciente, también será suficiente analizar los valores de $g'(x)$, en los extremos del intervalo, así:

$$g'(0.26) = 0.996435 \quad \text{y} \quad g'(0.29) = 0.998415 = k.$$

Luego como se satisface esta segunda condición, se define el algoritmo:

$$x_{n+1} = g(x_n) = \left[\frac{(8.3x_n^2 - 2.295x_n + 0.21141)}{10} \right]^{1/3}, \quad \forall n \geq 0,$$

y tomando $x_0 = 0.265$ como aproximación inicial, se obtienen las iteraciones que mostramos en la siguiente tabla:

n	x_n	TOL
0	0.265	
1	0.2650029666	2.966×10^{-6}
2	0.2650059293	2.9627×10^{-6}
3	0.2650088881	2.9588×10^{-6}
4	0.2650118429	2.9548×10^{-6}
5	0.2650147938	2.9509×10^{-6}
10	0.26502949	2.93×10^{-6}
20	0.26505859	2.89×10^{-6}
30	0.26508732	2.86×10^{-6}
40	0.26511567	2.82×10^{-6}
50	0.26514367	2.79×10^{-6}
60	0.26517130	2.75×10^{-6}
70	0.26519859	2.71×10^{-6}
80	0.26522554	2.68×10^{-6}
90	0.26525215	2.65×10^{-6}
100	0.26527843	2.61×10^{-6}
110	0.26530439	2.58×10^{-6}
120	0.26533004	2.55×10^{-6}
130	0.26535537	2.52×10^{-6}
140	0.26538040	2.49×10^{-6}
150	0.26540514	2.46×10^{-6}
160	0.26542958	2.43×10^{-6}
170	0.26545373	2.4×10^{-6}
180	0.26547761	2.37×10^{-6}
190	0.2655012	2.35×10^{-6}
200	0.26552453	2.32×10^{-6}

Tabla 6

Si nos fijamos en la tabla de iteraciones, después de 200 iteraciones se logra una solución aproximada 0.26552453, con una $TOL = |x_{n+1} - x_n| = 2.32 \times 10^{-6}$, y $|x_{n+1} - x_n| / |x_{n+1}| = 8.7374 \times 10^{-6}$, el valor exacto de la raíz es de 0.27, nuevamente observamos que el algoritmo de Punto Fijo es muy lento, y esto debido a que el valor de k en este caso es muy próximo de 1, por lo que la sucesión de valores converge muy lentamente.

Comentarios:

El método de Punto Fijo tiene como ventajas: la de tener un algoritmo sencillo, ya que usa en cada paso una evaluación de la función de iteración, generando de esta forma una sola sucesión $\{x_n\}_{(n \geq 0)}$.

Este método presenta como desventajas la necesidad de encontrar una función de iteración que asegure la convergencia del proceso, y además esta convergencia no debe ser demasiado lenta, lo que ocurre en el caso de $k \approx 1$, ya que daría lugar a un error propagado grande.

2.4 Orden de Convergencia

Con el fin de medir la rapidez de convergencia se pueden dar los siguientes conceptos:

Definición:

Sea $\lim_{n \rightarrow \infty} x_n = \alpha$ y $e_n = x_n - \alpha$, $n \geq 0$.

Si existen constantes positivas λ y k , tal que

$$\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{|e_n|^k} = \lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^k} = \lambda$$

se dice entonces que la convergencia de $\{x_n\}$ a α es de orden k , con una constante de error asintótico λ .

Casos Particulares:

La constante λ afecta a la rapidez de convergencia, pero no es tan importante como la constante k , orden de convergencia,

- Si $k = 1$ la convergencia es de orden 1 ó lineal y

$$\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{|e_n|} = \lambda$$

lo que implica que para n suficientemente grande se satisface:

$$e_{n+1} \approx \lambda e_n.$$

- Si $k = 2$ la convergencia es de orden 2 ó cuadrática y se cumple

que: $\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{e_n^2} = \lambda$, y para n suficientemente grande se satisface:

$$e_{n+1} \approx \lambda e_n^2$$

por lo tanto la convergencia es mucho más rápida.

Caso 1: Convergencia Lineal

Analizaremos que la convergencia del método de Punto Fijo es de orden 1, es decir la convergencia de $x_{n+1} = g(x_n) \xrightarrow{n \rightarrow \infty} \alpha$ es lineal, ya que:

$$e_{n+1} = x_{n+1} - \alpha = g(x_n) - g(\alpha)$$

y por el Teorema del Valor Medio, se sabe que existe un valor ξ_n que está entre x_n y α , tal que:

$$g(x_n) - g(\alpha) = g'(\xi_n)(x_n - \alpha) = g'(\xi_n) e_n$$

entonces:

$$\lim_{n \rightarrow \infty} \left| \frac{e_{n+1}}{e_n} \right| = \lim_{n \rightarrow \infty} |g'(\xi_n)| = |g'(\alpha)|, \text{ ya que } \xi_n \text{ está entre } x_n \text{ y } \alpha,$$

y $g'(x)$ es una función continua en una vecindad de α y la convergencia es lineal si $g'(\alpha) \neq 0$, que es el caso de Punto Fijo.

Caso 2: Convergencia Cuadrática

Ahora, construiremos una función de iteración que da lugar a una sucesión de convergencia de orden 2. Tomaremos a una función $g(x)$, tal que $g'(\alpha) = 0$, y usando el Teorema de Taylor, tendremos:

$$\begin{aligned} e_{n+1} &= x_{n+1} - \alpha = \\ &= g(x_n) - g(\alpha) = g'(\alpha)(x_n - \alpha) + g''(\xi_n)(x_n - \alpha)^2/2! = \end{aligned}$$

como $g'(\alpha) = 0$, se tiene:

$$e_{n+1} = g''(\xi_n)e_n^2/2!$$

por lo tanto:

$$\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{e_n^2} = \lim_{n \rightarrow \infty} \frac{|g''(\xi_n)|}{2!}$$

y si este límite es diferente de cero, tendremos

$$\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{e_n^2} = \lambda \neq 0,$$

esto es, la convergencia sería de orden 2.

Es fácil ver que tomando

$$g(x) = x - f(x)/f'(x)$$

con

$$g'(x) = 1 - \{[f'(x)]^2 - f(x)f''(x)\} / [f'(x)]^2 = f(x)f''(x) / [f'(x)]^2$$

para $x = \alpha$,

$$g'(\alpha) = [f(\alpha)f''(\alpha)] / [f'(\alpha)]^2 = 0 \text{ si } f'(\alpha) \neq 0.$$

Esta elección de $g(x)$, nos lleva a una función de iteración de la forma:

$$x_{n+1} = g(x_n) = x_n - f(x_n)/f'(x_n),$$

que se conoce como el método de Newton. La convergencia del método de Newton es cuadrática.

2.5 Método de Newton-Raphson

Hemos definido el método de Newton como el método de Punto Fijo con

$$g(x) = x - f(x) / f'(x).$$

Entonces se puede analizar la convergencia del método de Newton a través del siguiente teorema:

Teorema 1:

Sea $f(x) \in C^2[a, b]$, y sea $\alpha \in [a, b]$ un cero de $f(x)$, con $f'(\alpha) \neq 0$, entonces $\exists \delta > 0$ tal que el método de Newton genera una sucesión $\{x_n\}_{n \geq 0}$, que converge a α , para cualquier aproximación inicial $x_0 \in [\alpha - \delta, \alpha + \delta]$.

Demostración:

Probaremos que se cumplen las condiciones del Teorema de Punto Fijo. Como $f'(x)$ es continua en $[a, b]$, y $\alpha \in [a, b]$, entonces $f'(x)$ es continua en $x = \alpha$.

Además $f'(\alpha) \neq 0$, entonces $\exists \delta > 0$ tal que

$$f'(x) \neq 0 \quad \forall x \in [\alpha - \delta, \alpha + \delta] \subset [a, b].$$

Entonces $g(x) = x - f(x)/f'(x)$ está definida y es continua en $[\alpha - \delta, \alpha + \delta]$. Por otro lado, se sabe que:

$$g'(\alpha) = 0$$

y por la continuidad de $g'(x)$, podemos afirmar que $|g'(x)| \leq k < 1$ en la cercanía de $x = \alpha$, por lo tanto $\exists \delta > 0$ tal que $g(x)$ está definida y $|g'(x)| \leq k < 1$, $\forall x \in [\alpha - \delta, \alpha + \delta]$ y como:

$$|g(x) - \alpha| = |g(x) - g(\alpha)| = |g'(\xi)| |x - \alpha| \leq k |x - \alpha| < |x - \alpha|$$

$$\forall x \in [\alpha - \delta, \alpha + \delta],$$

lo que prueba que el método de Newton como el método de Punto Fijo, converge $\forall x_0 \in [\alpha - \delta, \alpha + \delta]$. \square

Ejemplo 1:

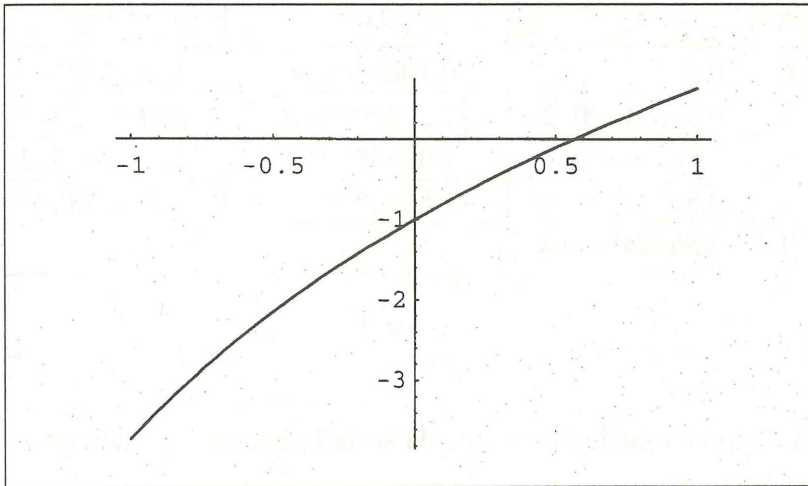
Hallar la raíz de la ecuación

$$f(x) = x - e^{-x} = 0,$$

empleando el método de Newton.

Solución:

La gráfica de esta función se muestra en el siguiente cuadro:



Gráfica 11

Como se cumple que

$$f(0) = -1, f(1) = 0.6321206,$$

entonces hay un cero de $f(x)$ en este intervalo y de la gráfica observamos que este cero está próximo a 0.5.

Evaluando la primera y la segunda derivada para definir el algoritmo se tiene:

$$f'(x) = 1 + e^{-x}, \quad y \quad f''(x) = -e^{-x}$$

Notamos que son funciones continuas en este intervalo.

El algoritmo se define como

$$x_{n+1} = x_n + \frac{x_n - e^{-x_n}}{1 + e^{-x_n}}$$

y puesto que la raíz está próxima a 0.5 y $f'(0.5) = 1.6065 \neq 0$, usaremos $x_0 = 0.5$ como aproximación inicial.

En la siguiente tabla se muestran las iteraciones que corresponden al algoritmo:

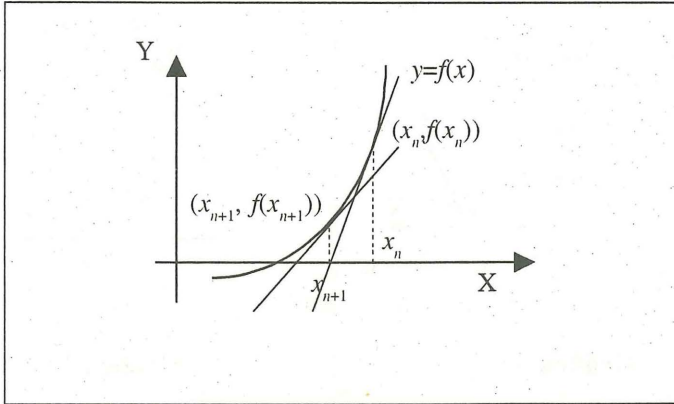
n	x_n	$f(x_n)$	$f'(x_n)$
0	0.5	-0.1065306597	1.60653066
1	0.5663110032	$-1.3045098 \times 10^{-3}$	1.567615513
2	0.567143165	-1.965351×10^{-7}	1.567143362
3	0.5671432904	-1.53×10^{-11}	1.56714329
4	0.5671432904	—	—

Tabla 7

Luego la solución aproximada sería la que se obtiene con la 4ta. Iteración.

Comentarios:

1. El Teorema 1 establece que la aproximación inicial x_0 , debe ser muy cercana a α , pero no nos da una idea práctica para determinarla.
2. El método de Newton se conoce también como el método de las tangentes, ya que $x_{n+1} = x_n - f(x_n)/f'(x_n)$, determina x_{n+1} como el punto de intersección de la recta tangente trazada a la curva $y = f(x)$ en el punto $(x_n, f(x_n))$, con el eje X :



Gráfica 12

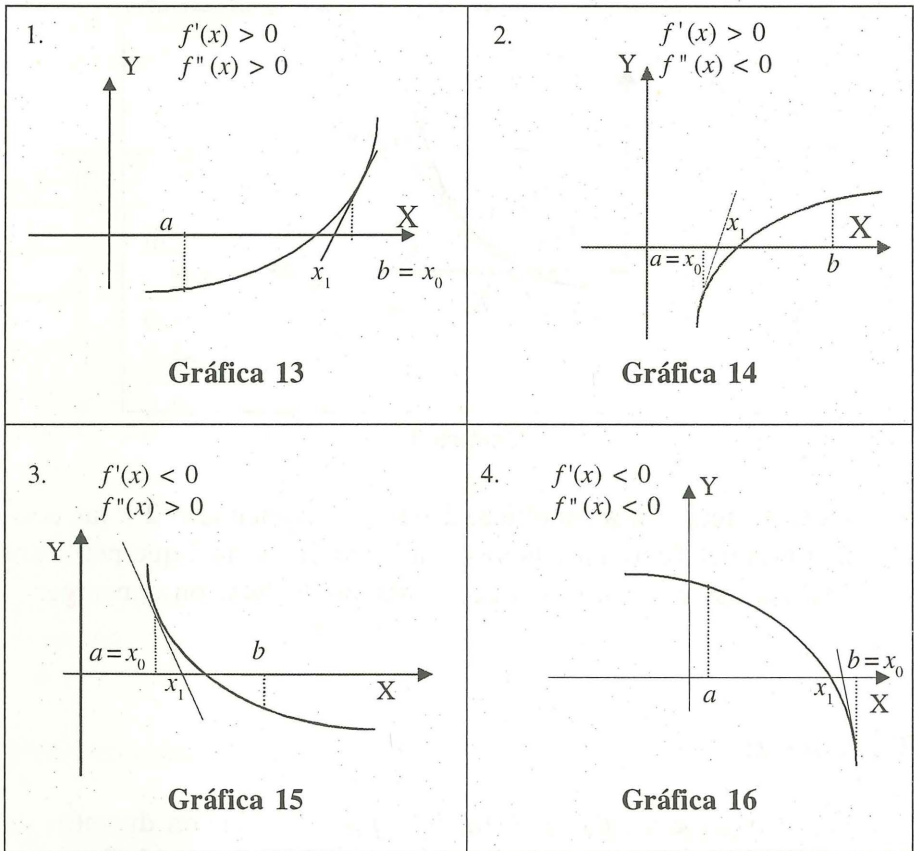
3. La interpretación geométrica, dada en el comentario 2 y las condiciones del Teorema 1 nos conducen al Teorema 2 que permitirá analizar los casos en los que el método de Newton converge:

Teorema 2:

Sea $f(a)f(b) < 0$, $f(x) \in C^2[a, b]$ y $f'(x)$, $f''(x)$ son distintos de cero en $[a, b]$ y cada uno de ellas conserva el signo en $[a, b]$. Entonces tomando $x_0 \in [a, b]$ tal que $f(x_0)f''(x_0) > 0$, el método de Newton determina una sucesión que converge a α , único cero de $f(x)$ en $[a, b]$.

Comentario:

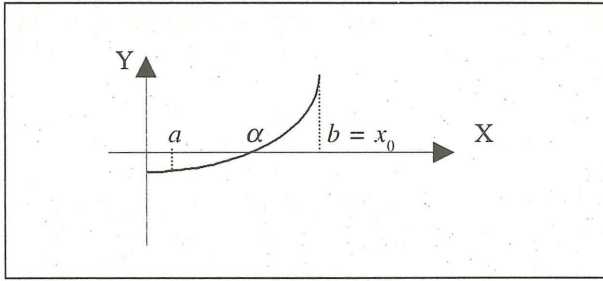
Geoméricamente el Teorema 2 se puede interpretar a través de los cuatro casos siguientes:



Demostración:

Probaremos sólo el caso 1. Los casos restantes se pueden demostrar en forma parecida.

Tomemos $f'(x) > 0$, $f''(x) > 0$



Gráfica 17

Se puede observar que $f(a) < 0$, $f(b) > 0$, y como $f'(x) > 0$, $\forall x \in [a, b]$, entonces la función $f(x)$ tiene un sólo cero en el intervalo.

Tomando $x_0 = b$ y $x_0 > \alpha$, probaremos que:

$$x_{n+1} = x_n - f(x_n) / f'(x_n)$$

determina una sucesión monótona decreciente.

Usando inducción probaremos que:

$$x_{n+1} > \alpha, \forall n.$$

Para $n = 0$, sabemos que $x_0 > \alpha$, supongamos que $x_n > \alpha$, para algún n .

Tomando

$$\alpha = x_n + (\alpha - x_n)$$

y usando el Teorema de Taylor (ver Apéndice A), tenemos:

$$0 = f(\alpha) = f(x_n) + f'(x_n)(\alpha - x_n) + f''(\xi_n)(\alpha - x_n)^2/2!$$

con

$$\alpha < \xi_n < x_n \text{ y como } f''(x) > 0 \Rightarrow$$

$$x_{n+1} = x_n - f(x_n) / f'(x_n) > \alpha$$

Además, como

$$f'(x_n) > 0 \text{ y } f(x_n) > 0 \Rightarrow x_{n+1} < x_n, \quad n \geq 0$$

lo que implica que x_n es una sucesión monótona decreciente y por lo tanto existe $\lim_{n \rightarrow \infty} x_n = \tilde{\alpha}$

Pero

$$\lim_{n \rightarrow \infty} x_n = \tilde{\alpha} = \lim_{n \rightarrow \infty} [x_{n+1}] = \lim_{n \rightarrow \infty} \left[x_n - \frac{f(x_n)}{f'(x_n)} \right]$$

tenemos:

$$\tilde{\alpha} = \tilde{\alpha} - f(\tilde{\alpha}) / f'(\tilde{\alpha}) \Rightarrow f(\tilde{\alpha}) = 0 \Rightarrow \tilde{\alpha} = \alpha,$$

lo que prueba que $x = \alpha$ es el único cero de $f(x)$ en $[a, b]$ \square

Algoritmo de Newton-Raphson:

Objetivo: Hallar una raíz α con una tolerancia ε especificada, de una función $f(x) \in C^2$ en $[a, b]$.

Paso 1: Hallar una aproximación inicial $x_0 \in [a, b]$

Para $n = 0, 1, 2, \dots$ hacer los pasos (2) y (3).

Paso 2. Calcular:

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

Paso 3. Averiguar si se satisface la tolerancia especificada, si fuese así tomar como la solución

$$\alpha \approx x_{n+1}$$

Ejemplo 2:

Hallar el recíproco de un número positivo c , sin emplear la división. Esto es, si $c > 0$, hay que determinar $1/c$.

Solución:

Con tal propósito definiremos la función $f(x) = 1/x - c$, cuyo cero es $1/c$. Utilizando el Teorema 2, se ve:

$$f'(x) = -1/x^2 < 0, \text{ y } f''(x) = 2/x^3 > 0, \forall x > 0.$$

Entonces el algoritmo de Newton se define como:

$$x_{n+1} = x_n - f(x_n) / f'(x_n);$$

y toma la forma:

$$x_{n+1} = x_n - (1/x_n - c) / (-1/x_n^2) = x_n + (1 - cx_n) x_n$$

$$\text{ó } x_{n+1} = x_n(2 - cx_n), \text{ con } x_0 > 0.$$

Por ejemplo, para hallar e^{-1} , para $e = 2.7182818$, se tiene:

algoritmo: $x_{n+1} = x_n(2 - ex_n), \forall n \geq 0,$

con $x_0 = 0.3$, ya que la raíz se ubica en el intervalo $[0, 1]$.

Las iteraciones se muestran en la siguiente tabla:

n	x_n	$TOL = x_{n+1} - x_n $
1	0.3553546354	0.0553546355
2	0.3674530222	0.0120983868
3	0.3678789469	0.0004259247
4	0.3678794412	4.942×10^{-7}
5	0.3678794412	0

Tabla 8.

Dificultades:

Como los teoremas 1 y 2 exigen que $f'(x) \neq 0$ en la cercanía de α , las dificultades surgen cuando $f'(x) = 0$ para algún x cercano a α ó $f'(\alpha) = 0$, que es el caso de un cero múltiple de $f(x)$, condiciones que no siempre son fáciles de detectar.

Ejemplo 3:

Sea

$$f(x) = 10x^3 - 8.3x^2 + 2.295x - 0.21141 = 0.$$

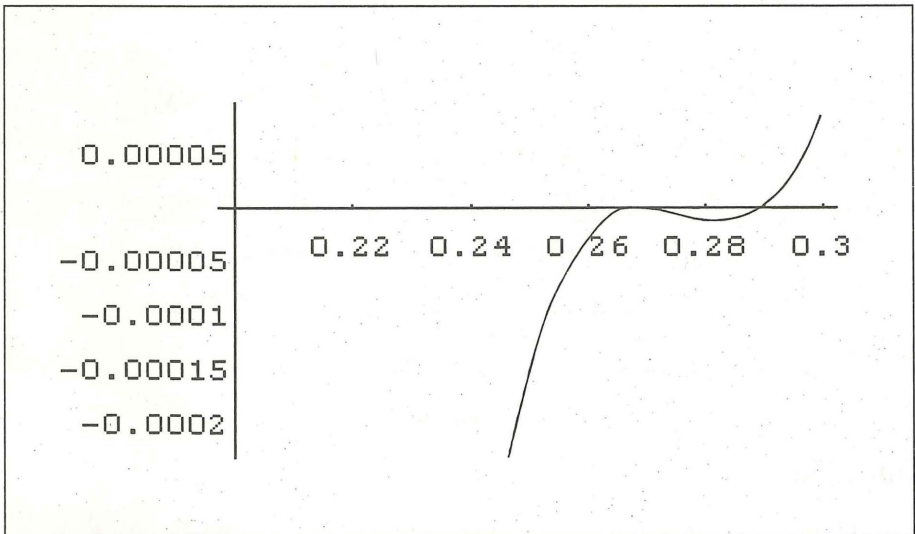
Se sabe que la función $f(x)$ tiene una raíz en $\alpha = 0.27$. Emplearemos el método de Newton para tratar de encontrar esta raíz, comenzando con las aproximaciones iniciales:

- (i) $x_0 = 0.265$
- (ii) $x_0 = 0.275$
- (iii) $x_0 = 0.28$
- (iv) $x_0 = 0.285$
- (v) $x_0 = 0.3$

y analizaremos los resultados obtenidos.

Solución:

En el siguiente cuadro se muestra la gráfica de la función cúbica dada:



Gráfica 18

de la cual se puede apreciar que tiene dos raíces en $\alpha_1 = 0.27$ (revisar el ejemplo 2 de Punto Fijo) y en 0.29.

Construiremos el algoritmo de Newton:

$$f(x) = 10x^3 - 8.3x^2 + 2.295x - 0.21141 = 0,$$

$$f'(x) = 30x^2 - 16.6x + 2.295,$$

$$f''(x) = 60x - 16.6,$$

algoritmo:

$$x_{n+1} = x_n - \frac{10x_n^3 - 8.3x_n^2 + 2.295x_n - 0.21141}{30x_n^2 - 16.6x_n + 2.295}, \text{ para } n \geq 0$$

y aplicándolo a las diferentes aproximaciones iniciales, tenemos:

(i) Si $x_0 = 0.265$

n	x_n	$f(x_n)$	$f'(x_n)$
0	0.265	-6.25×10^{-6}	2.75×10^{-3}
1	0.2672	-1.69045×10^{-6}	1.314049×10^{-3}
2	0.2685591768	-4.451054×10^{-7}	$6.38608425 \times 10^{-4}$
3	0.2692561695	-1.147723×10^{-7}	$3.14130714 \times 10^{-4}$
4	0.2696215343	-2.91893×10^{-8}	$1.55683368 \times 10^{-4}$
5	0.2698090257	-7.3639×10^{-9}	7.7483855×10^{-5}
6	0.2699040636	-1.8497×10^{-9}	3.8650674×10^{-5}
7	0.2699519205	-4.634×10^{-10}	1.9301149×10^{-5}
8	0.2699759294	-1.16×10^{-10}	9.645622×10^{-6}
9	0.2699879556	-2.9×10^{-11}	4.822112×10^{-6}
10	0.2699939696	-7.3×10^{-12}	2.413251×10^{-6}
11	0.2699969946	-1.8×10^{-12}	1.202431×10^{-6}
12	0.2699984916	0	6.03428×10^{-7}
13	0.2699984916	-	-

Tabla 9

(ii) Si $x_0 = 0.275$, se obtienen las siguientes iteraciones:

n	x_n	$f(x_n)$	$f'(x_n)$
0	0.275	3.75×10^{-6}	-1.25×10^{-3}
1	0.272	-7.2×10^{-7}	-6.8×10^{-4}
2	0.2709411765	-1.68825×10^{-7}	$-3.4989620 \times 10^{-4}$
3	0.2704586746	-4.1×10^{-8}	$-1.7715836 \times 10^{-4}$
4	0.2702266139	-1.01549×10^{-8}	-8.91049×10^{-5}
5	0.2701126483	-2.536×10^{-9}	$-4.4678631 \times 10^{-5}$
6	0.2700561649	-6.292×10^{-10}	$-2.2371325 \times 10^{-5}$
7	0.2700280396	-1.57×10^{-10}	$-1.1192254 \times 10^{-5}$
8	0.270014012	-3.92×10^{-11}	-5.59891×10^{-6}
9	0.2700070106	-9.8×10^{-12}	-2.802766×10^{-6}
10	0.2700035141	-2.4×10^{-12}	-1.405×10^{-6}
11	0.2700018062	0	-7.2×10^{-7}
12	0.2700018062	0	0

Tabla 10

vemos que la convergencia en los casos (i) y (ii) es lenta, (a pesar de que el método de Newton tiene convergencia cuadrática.)

(iii) Si $x_0 = 0.28$, obtenemos las siguientes iteraciones:

n	x_n	$f(x_n)$	$f'(x_n)$
0	0.28	-1×10^{-5}	-1×10^{-3}
1	0.27	0	0
2	no está definido ya que $f'(0.27) = 0$		

Tabla 11

en este caso no se puede aplicar el algoritmo, notar de que $f(0.27) = f'(0.27) = 0$.

(iv) Si $x_0 = 0.285$, obtenemos las siguientes iteraciones:

n	x_n	$f(x_n)$	$f'(x_n)$
0	0.285	-1.125×10^{-5}	7.5×10^{-4}
1	0.3	9×10^{-5}	0.015
2	0.294	2.304×10^{-5}	7.68×10^{-3}
3	0.291	4.41×10^{-6}	4.83×10^{-3}
4	0.2900869565	3.508571×10^{-7}	4.0697920×10^{-3}
5	0.2900007464	2.9858×10^{-9}	4.0006×10^{-3}
6	0.29	0	4×10^{-3}
7	0.29	—	—

Tabla 12

La convergencia es hacia la raíz $\alpha_2 = 0.29$, luego $f(x)$ tiene dos raíces próximas entre si $\alpha_1 = 0.27$ y $\alpha_2 = 0.29$.

(v) Si $x_0 = 0.3$, obtenemos las siguientes iteraciones:

n	x_n	$f(x_n)$	$f'(x_n)$
0	0.3	9×10^{-5}	0.015
1	0.294	2.304×10^{-5}	7.68×10^{-3}
2	0.291	4.41×10^{-6}	4.83×10^{-3}
3	0.2900869565	3.508571×10^{-7}	4.06979×10^{-3}
4	0.2900007464	2.9858×10^{-9}	4.0006×10^{-3}
5	0.29	0	4×10^{-3}
6	0.29	—	—

Tabla 13

Por el lado izquierdo o derecho de 0.29, el algoritmo converge hacia dicha raíz.

¿Qué pasa?

En el caso de la raíz 0.27, si se inicia con $x_0 = 0.275$ ó 0.265 ó 0.28 el algoritmo, converge lentamente o no es aplicable.

Se puede concluir que $f(x)$ es un polinomio cúbico con raíces $\alpha_1 = 0.27$ y $\alpha_2 = 0.29$, entonces:

$$f(x) = 10(x - 0.27)^2(x - 0.29)$$

ya que en 0.27 es una raíz o cero de multiplicidad 2.

Cuando hay un cero con multiplicidad, el algoritmo de Newton se vuelve lento ya que $f'(0.27) = 0$ (ver los Teorema 1 y 2), luego los errores por la aritmética de Punto Flotante tienden a aumentar, puesto que de acuerdo al algoritmo en cada paso se estaría dividiendo entre una cantidad de magnitud pequeña.

¿Cómo se podría mejorar la rapidez de la convergencia?

En el caso de que una función presente un cero en $x = \alpha$ con multiplicidad m ; esto es $f(x)$ se puede escribir como $f(x) = (x-\alpha)^m q(x)$, para $x \neq \alpha$, donde $\lim_{x \rightarrow \alpha} q(x) \neq 0$. Se puede enfrentar este problema definiendo la función:

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

luego:

$$\begin{aligned} f'(x) &= m(x-\alpha)^{m-1}q(x) + q'(x)(x-\alpha)^m, \\ \Rightarrow h(x) &= \frac{(x-\alpha)q(x)}{mq(x) + q'(x)(x-\alpha)} \end{aligned}$$

entonces $x = \alpha$ es raíz de $f(x)$ de multiplicidad m y es raíz de multiplicidad uno de $h(x)$. Luego aplicando el método de Newton a $h(x)$, se tiene:

$$\begin{aligned} x_{n+1} &= x_n - \frac{h(x_n)}{h'(x_n)}, \quad \forall n \geq 0 \\ x_{n+1} &= x_n - \frac{f'(x_n)f(x_n)}{[f'(x_n)]^2 - f''(x_n)f(x_n)} \end{aligned}$$

y este algoritmo se conoce como el método modificado de Newton, y se puede probar que tiene convergencia cuadrática. La desventaja principal en este método está en el cálculo adicional de $f''(x)$, y en la laboriosidad del algoritmo lo que puede ocasionar serios errores de redondeo.

En nuestro caso:

$$\begin{aligned}f(x) &= 10x^3 - 8.3x^2 + 2.295x - 0.21141 \\f'(x) &= 30x^2 - 16.6x + 2.295 \\f''(x) &= 60x - 16.6\end{aligned}$$

entonces:

$$h(x) = \frac{f(x)}{f'(x)} = \frac{x^2 - 0.56x + 0.0783}{3x - 0.85}$$

y el algoritmo del método modificado es:

$$\begin{aligned}x_{n+1} &= x_n - \frac{h(x_n)}{h'(x_n)} \\x_{n+1} &= x_n - \frac{3x_n^3 - 2.53x_n^2 + 0.7109x_n - 0.066555}{0.2411 - 1.7x_n + 3x_n^2}\end{aligned}$$

- (a) En el caso (iii) se utilizó $x_0 = 0.28$, como aproximación inicial para el algoritmo de Newton, observándose que no se podía aplicar el algoritmo. Si aplicamos el algoritmo modificado de Newton, obtenemos las siguientes iteraciones:

n	x_n	$h(x_n)$	$h'(x_n)$
0	0.28	1×10^{-6}	3×10^{-4}
1	0.276666	1.7777×10^{-6}	3.99999×10^{-4}
2	0.272222	1.3168725×10^{-6}	6.37×10^{-4}
3	0.2701550388	1.2164×10^{-7}	7.8767×10^{-4}
4	0.2700006104	4.883×10^{-10}	7.99×10^{-4}
5	0.27	0	8×10^{-4}
6	0.27	—	—

Tabla 14

Apreciamos que hay una convergencia rápida hacia la raíz 0.27.

- (b) Si $x_0 = 0.275$, con el método modificado de Newton se obtienen las siguientes iteraciones:

n	x_n	$h(x_n)$	$h'(x_n)$
0	0.275	1.875×10^{-6}	4.75×10^{-4}
1	0.2710526316	7.35×10^{-7}	7.1911×10^{-4}
2	0.2700308166	2.45×10^{-8}	7.97×10^{-4}
3	0.2700000239	1.91×10^{-11}	7.9×10^{-4}
4	0.27	0	8×10^{-4}
5	0.27	—	—

Tabla 15

en el caso (ii) se precisaban realizar 11 iteraciones para que converja el método y ahora se reducen a 5 iteraciones, es decir se recupera la convergencia cuadrática del método de Newton.

2.6 Convergencia Acelerada

En el caso del Método de Punto Fijo sabemos que tiene convergencia lineal, por lo que en ciertas circunstancias es necesario realizar muchas iteraciones lo que aumenta el error de redondeo. En esta sección desarrollaremos técnicas que aumentan la rapidez de convergencia de este método.

Consideremos una técnica llamada el método Δ^2 de Aitken, la cual puede ser usada para acelerar la convergencia de cualquier sucesión que converja linealmente, independientemente de su origen:

Sea $\{x_n\}_{(n \geq 0)}$ una sucesión que converge linealmente a su límite α ; entonces $e_n = x_n - \alpha$, es una sucesión que converge a cero,

$$\Rightarrow \lim_{n \rightarrow +\infty} \frac{|e_{n+1}|}{|e_n|} = \lambda \quad (\lambda > 0)$$

Sea el $(n+2)$ -ésimo error que se comete con el término x_{n+2} :

$$\begin{aligned} e_{n+2} &= x_{n+2} - \alpha \\ \Rightarrow x_{n+2} &= e_{n+2} + \alpha \end{aligned}$$

supongamos que n sea lo suficientemente grande y que los errores e_n tienen el mismo signo, entonces se tiene:

$$e_{n+1} \approx \lambda e_n \quad \text{y} \quad e_{n+2} \approx \lambda e_{n+1}$$

pero entonces,

$$\begin{aligned} x_{n+2} &\approx \lambda e_{n+1} + \alpha \\ x_{n+2} &\approx \lambda(x_{n+1} - \alpha) + \alpha \quad \dots(1) \end{aligned}$$

lo que nos proporciona una relación entre α y λ ; en forma similar para $(n+1)$, se obtiene:

$$x_{n+1} \approx \lambda(x_n - \alpha) + \alpha \quad \dots(2)$$

y eliminando λ en las ecuaciones (1) y (2), para despejar α , se tiene:

$$\frac{x_{n+2} - \alpha}{x_{n+1} - \alpha} \approx \frac{x_{n+1} - \alpha}{x_n - \alpha}$$

$$\Rightarrow (x_{n+2} - \alpha)(x_n - \alpha) \approx (x_{n+1} - \alpha)^2$$

$$x_{n+2} x_n - \alpha x_{n+2} - \alpha x_n + \alpha^2 \approx x_{n+1}^2 - 2\alpha x_{n+1} + \alpha^2$$

$$x_{n+2} x_n - x_{n+1}^2 \approx \alpha (x_{n+2} + x_n - 2x_{n+1})$$

$$\Rightarrow \alpha \approx \frac{x_n x_{n+2} - x_{n+1}^2}{x_{n+2} + x_n - 2x_{n+1}}$$

y arreglando los términos:

$$\alpha \approx \frac{x_n (x_{n+2} + x_n - 2x_{n+1}) - x_n^2 + 2x_n x_{n+1} - x_{n+1}^2}{x_{n+2} + x_n - 2x_{n+1}}$$

$$\alpha \approx x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n}$$

$$\alpha \approx \hat{x}_n = x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n} \quad \dots(3)$$

y la sucesión $\{\hat{x}_n\}_{(n \geq 0)}$ que se ha generado, corresponde al método Δ^2 de Aitken y está basada en la suposición de que converge más rápidamente a α , en el sentido de que $\lim_{n \rightarrow \infty} \frac{\hat{x}_n - \alpha}{x_n - \alpha} = 0$.

Definición:

Dada la sucesión $\{x_n\}_{(n \geq 0)}$ se define la diferencia progresiva o finita Δx_n mediante:

$$\Delta x_n = x_{n+1} - x_n, \quad \forall n \geq 0$$

Las potencias mayores $\Delta^k x_n$ (que son diferencias finitas o progresivas de orden k) se definen recursivamente mediante

$$\Delta^k x_n = \Delta^{k-1}(\Delta x_n), \quad \forall k \geq 2$$

En base a la definición dada, para $k = 2$:

$$\Delta^2 x_n = \Delta(\Delta x_n) = \Delta(x_{n+1} - x_n) = \Delta x_{n+1} - \Delta x_n = x_{n+2} - 2x_{n+1} + x_n$$

Por lo tanto la fórmula para el método Δ^2 de Aitken, sería:

$$\hat{x}_n = x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n}, \quad \forall n \geq 0 \quad \dots(4)$$

Algoritmo de Δ^2 de Aitken:

Objetivo: Hallar una sucesión $\{x_n\}_{(n \geq 0)}$ que converge a α , cero de la función $f(x)$, de convergencia lineal.

Paso 1. Usando el algoritmo del Punto Fijo construir la sucesión

$$\{x_n\}_{(n \geq 0)}$$

Paso 2. Para $n \geq 0$ tomar:

$$\hat{x}_n = x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n}$$

donde

$$\begin{aligned}\Delta x_n &= x_{n+1} - x_n \\ \Delta^2 x_n &= x_{n+2} - 2x_{n+1} + x_n\end{aligned}$$

Paso 3. Verificación de la tolerancia.

Ejemplo:

Sea la ecuación:

$$f(x) = x - 2^{-x} = 0.$$

Hallar la raíz de $f(x)$ en $[0, 1]$.

Solución:

Construiremos el algoritmo de Punto Fijo, para luego aplicar el método Δ^2 de Aitken.

Observamos que: $f(0) = -1 < 0$ y $f(1) = 1/2 > 0$.

Sea

$$x = g(x) = 2^{-x} = 1/2^x$$

Verifiquemos si se satisfacen las condiciones de convergencia del método de Punto Fijo:

(i) Veamos la primera condición:

$$g: [0, 1] \rightarrow B \subset [0, 1], \forall x \in [0, 1]$$

notamos que $g(x)$ es una función estrictamente decreciente, por lo que es suficiente analizar a $g(x)$ en sus extremos:

$$g(0) = 1, g(1) = 1/2.$$

por lo tanto $B = [1/2, 1]$.

(ii) Con relación al comportamiento de $g'(x)$:

$$g'(x) = (e^{x \ln(1/2)})' = \ln(1/2)e^{x \ln(1/2)} = \ln(1/2)(1/2)^x$$

Observamos que es continua y decreciente, por lo tanto:

$$|g'(0)| = |\ln(1/2)| = 0.693147 = k$$

$$|g'(1)| = |(1/2)\ln(1/2)| = 0.3465735$$

y de ahí $g'(x)$ está acotada:

$$|g'(x)| \leq k = 0.693147, \text{ para todo } x \text{ en } [0, 1]$$

Trabajaremos con los siguientes algoritmos:

$$\text{Punto Fijo: } x_{n+1} = g(x_n) = (1/2)^{x_n}$$

$$\Delta^2 \text{ de Aitken: } \hat{x}_n = x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n}, \forall n \geq 0.$$

mostrando las iteraciones obtenidas en el siguiente cuadro:

n	x_n (Punto Fijo)	\hat{x}_n (Δ^2 de Aitken)
0	0.5	0.6421876657
1	0.7071067812	0.641386131
2	0.61254732	0.6412250795
3	0.6540408601	0.6411935347
4	0.6354978458	0.6411872814
5	0.6437186417	0.6411860482
6	0.6400610212	0.6411858045
7	0.641685807	
8	0.6409635372	
9	0.6412845091	
10	0.6411418515	
11	0.6412052524	
12	0.6411770746	
13	0.6411895977	
14	0.641184032	
15	0.6411865056	
16	0.6411854062	

Tabla 16

Observamos que mejora la rapidez de convergencia con el método Δ^2 de Aitken, parecería más razonable comenzar la iteración de Punto Fijo con

\hat{x}_n como tentativa inicial, ya que \hat{x}_n será mejor aproximación inicial que x_n . Esta idea nos conduce a un nuevo método que se conoce como el algoritmo de Steffensen.

2.7 Método de Steffensen

El método de Steffensen consiste en una combinación del Método de Punto Fijo y el de Δ^2 de Aitken de la siguiente forma:

Con el Método de Punto Fijo (MPF), se genera con x_0 dado los puntos:

$$x_1 = g(x_0), \quad x_2 = g(x_1);$$

Con el Método Δ^2 de Aitken (Δ^2), se calcula: $y_1 = \hat{x}_0 = x_0 - \frac{(x_1 - x_0)^2}{x_2 - 2x_1 + x_0}$

luego, se toma a \hat{x}_0 como mejor aproximación que x_0 , para aplicar nuevamente el método de Punto Fijo, esto es:

$$x_0^{(1)} = \hat{x}_0, \quad x_1^{(1)} = g(\hat{x}_0) \quad x_2^{(1)} = g(x_1^{(1)})$$

y con el método Δ^2 de Aitken, se calcula:

$$y_2 = \hat{x}_0^{(1)} = x_0^{(1)} - \frac{(\Delta x_0^{(1)})^2}{\Delta^2 x_0^{(1)}}$$

y así sucesivamente, se continúan alternando los dos métodos, obteniéndose la sucesión $\{y_n\}_{(n \geq 0)}$, de Steffensen.

Ejemplo:

En el ejemplo anterior se resolvió la ecuación:

$$f(x) = x - 2^{-x} = 0, \text{ en el intervalo } [0, 1],$$

y se definió el algoritmo para el método de Punto Fijo como:

$$x_{n+1} = g(x_n) = (1/2)^{x_n} \quad \forall n \geq 0$$

Aplicando el método de Steffensen, tendremos:

$$n = 1$$

$$MPF: y_0 = x_0^{(0)} = 0.5, \quad x_1^{(0)} = g(0.5) = 0.70710678$$

$$x_2^{(0)} = g(x_1^{(0)}) = 0.612547$$

$$\Delta^2 \text{ de Aitken: } y_1 = x_0^{(0)} - \frac{(x_1^{(0)} - x_0^{(0)})}{x_2^{(0)} - 2x_1^{(0)} + x_0^{(0)}} = 0.6421876657$$

$$n = 2$$

$$MPF: y_1 = x_0^{(1)}, \quad x_1^{(1)} = g(y_1) = 0.64074061$$

$$x_2^{(1)} = g(x_1^{(1)}) = 0.64138361$$

$$\Delta^2 \text{ de Aitken: } y_2 = x_0^{(1)} - \frac{(x_1^{(1)} - x_0^{(1)})}{x_2^{(1)} - 2x_1^{(1)} + x_0^{(1)}} = 0.641185792$$

$$n = 3$$

$$MPF: y_2 = x_0^{(2)}, \quad x_1^{(2)} = g(y_2) = 0.64118572$$

$$x_2^{(2)} = g(x_1^{(2)}) = 0.64118575$$

$$\Delta^2 \text{ de Aitken: } y_3 = x_0^{(2)} - \frac{(x_1^{(2)} - x_0^{(2)})}{x_2^{(2)} - 2x_1^{(2)} + x_0^{(2)}} = 0.6411857445$$

$$n = 4$$

$$MPF: y_3 = x_0^{(3)}, \quad x_1^{(3)} = g(y_3) = 0.64118574$$

$$x_2^{(3)} = g(x_1^{(3)}) = 0.64118574$$

$$\Delta^2 \text{ de Aitken: } y_4 = x_0^{(3)} - \frac{(x_1^{(3)} - x_0^{(3)})}{x_2^{(3)} - 2x_1^{(3)} + x_0^{(3)}} = y_3$$

Lo que se resume en la siguiente tabla:

n	y_n
0	0.5
1	0.6421876657
2	0.6411857921
3	0.6411857445
4	0.6411857445

Tabla 17

se observa una mejoría en el proceso computacional, ya que con el método de Punto Fijo la convergencia era lenta.

Algoritmo de Steffensen:

Objetivo: Mejorar computacionalmente el algoritmo Δ^2 de Aitken. Usando en cada paso de iteración de Punto Fijo el valor obtenido en el algoritmo de Δ^2 de Aitken como valor inicial.

Hacer para $n = 0, 1, 2, \dots$

Paso 1. Con valor inicial tomar:

$$\begin{aligned}x_0 &= y_n \\x_1 &= g(x_0) \\x_2 &= g(x_1)\end{aligned}$$

Paso 2. Calcular:

$$y_{n+1} = \hat{x}_0 = x_0 - \frac{(x_1 - x_0)^2}{x_2 - 2x_1 + x_0}$$

y redefinir $x_0 = \hat{x}_0$

Paso 3. Verificación de la tolerancia.

2.8 Método de Punto Fijo para Sistemas

Consideremos un sistema de ecuaciones no lineales:

$$f_1(x_1, x_2, \dots, x_n) = 0$$

$$f_2(x_1, x_2, \dots, x_n) = 0$$

$$f_3(x_1, x_2, \dots, x_n) = 0$$

$$\cdot \quad \cdot \quad \cdot$$

$$\cdot \quad \cdot \quad \cdot$$

$$\cdot \quad \cdot \quad \cdot$$

$$f_n(x_1, x_2, \dots, x_n) = 0$$

donde cada función $f_i(x)$ puede pensarse como un mapeo de un vector $x = (x_1, \dots, x_n)^t$ (vector transpuesto) del espacio n -dimensional \mathfrak{R}^n , en la recta real \mathfrak{R} . Alternativamente, el sistema puede representarse definiendo una función vectorial F , de \mathfrak{R}^n en \mathfrak{R}^n por:

$$F(x_1, x_2, \dots, x_n) = (f_1(x_1, x_2, \dots, x_n), f_2(x_1, x_2, \dots, x_n), \dots, f_n(x_1, x_2, \dots, x_n))^t.$$

Usando notación vectorial para representar las variables x_1, x_2, \dots, x_n , el sistema dado asume la forma:

$$F(x) = 0$$

Las funciones f_1, f_2, \dots, f_n se llaman funciones coordenadas de F .

En el caso de una variable, en el proceso iterativo para resolver la ecuación $f(x) = 0$, se transforma primero esta ecuación en una ecuación de la forma $x = g(x)$. La función $g(x)$ tiene sus puntos fijos precisamente en las soluciones de la ecuación original. Aquí se estudiará un procedimiento similar para funciones de \mathfrak{R}^n en \mathfrak{R}^n , en donde el sistema se transforma en la ecuación $x = G(x)$.

Definición:

Se dice que una función G de $D \subset \mathfrak{R}^n$ a \mathfrak{R}^n tiene un **Punto Fijo** en

$\alpha \in D$ si $G(\alpha) = \alpha$.

El siguiente teorema extiende los teoremas de Punto Fijo al caso n -dimensional. Este teorema es un caso especial del bien conocido

Teorema de Contracción de Mapeos

Sea $D = \{(x_1, x_2, \dots, x_n)' / a_i \leq x_i \leq b_i \text{ para cada } i = 1, 2, \dots, n\}$ para alguna colección de constantes a_1, a_2, \dots, a_n y b_1, b_2, \dots, b_n . Supongamos que $G = (g_1(x), g_2(x), \dots, g_n(x))$ es una función continua y sus componentes tienen las primeras derivadas parciales continuas en $D \subset \mathfrak{R}^n$ con la propiedad de que $G(x) \in D$ para $x \in D$. Entonces G tiene un punto fijo en D .

Además, supóngase que existe una constante $K < 1$ con

$$\left| \frac{\partial g_i(x)}{\partial x_j} \right| \leq \frac{K}{n}, \text{ siempre que } x \in D,$$

para cada $j = 1, 2, \dots, n$ y cada función componente $g_i(x)$. Entonces la sucesión $\{x^{(k)}\}_{(k \geq 0)}$ generada por

$$x^{(k)} = G(x^{(k-1)}) \text{ para cada } k \geq 1,$$

y por una $x^{(0)}$ aproximación inicial seleccionada arbitrariamente en D , converge al punto fijo único $\alpha \in D$ y

$$\|x^{(k)} - \alpha\|_{\infty} \leq \frac{K^k}{1 - K} \|x^{(1)} - x^{(0)}\|_{\infty}.$$

Demostración:

Leer en el libro de Ortega [18].

Comentario:

Observar que a partir de este resultado del teorema, se puede concluir:

$$\sum_{j=1}^n \left| \frac{\partial g_i(x)}{\partial x_j} \right| \leq K < 1, \forall i = 1, 2, \dots, n$$

lo que puede ser usado como criterio de convergencia.

Ejemplo:

Resolver el sistema:

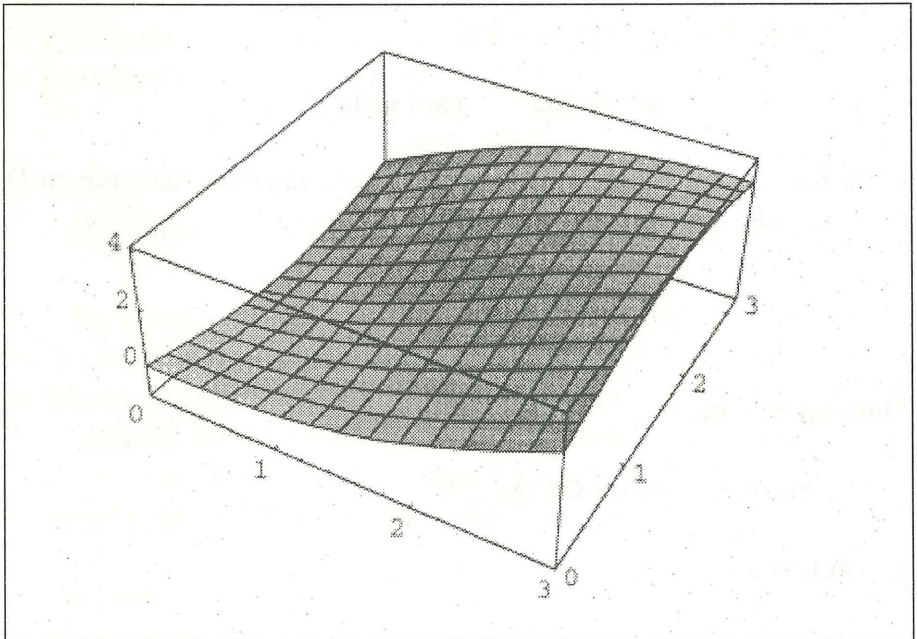
$$f(x,y) = x - \text{sen}(x + y) = 0$$

$$g(x,y) = y - \text{cos}(x - y) = 0$$

si $(x_0, y_0) = (0.75, 0.75)$, empleando el método de Punto Fijo.

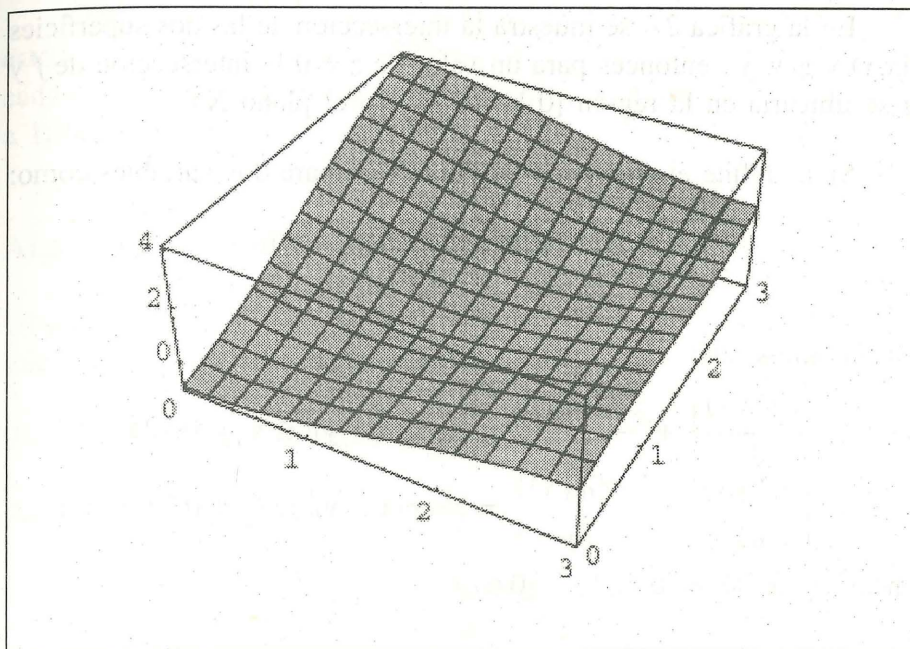
Solución:

La gráfica en \mathcal{R}^3 de las funciones f y g se muestra a continuación:

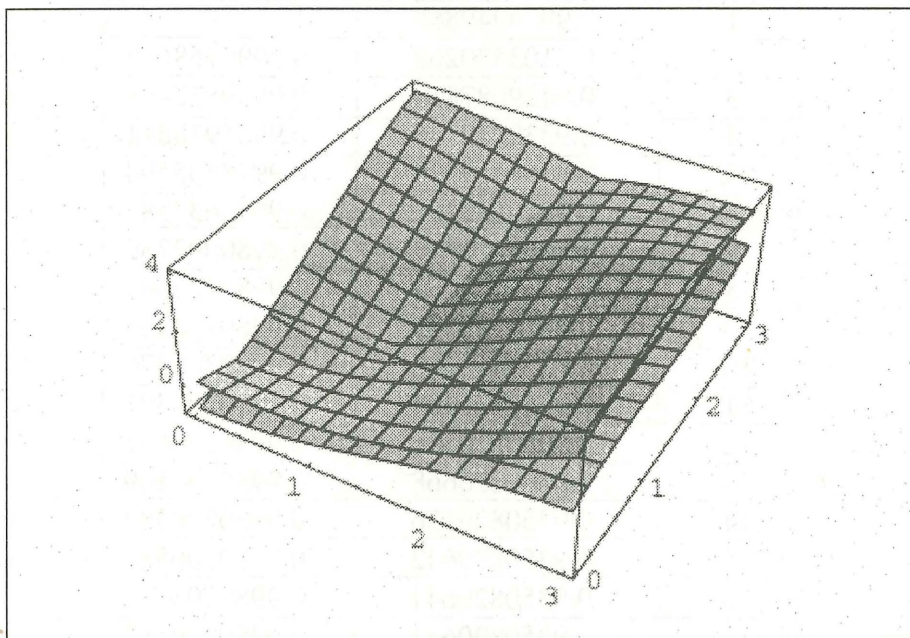


Gráfica 19

$$f(x,y) = x - \text{sen}(x + y)$$



Gráfica 20
 $g(x,y) = y - \cos(x-y)$



Gráfica 21 (f y g)

En la gráfica 21, se muestra la intersección de las dos superficies $f(x,y)$ y $g(x,y)$, entonces para un valor de $z = 0$ la intersección de f y g se ubicaría en la región $[0,1] \times [0,1]$ en el plano XY .

Si se define el algoritmo de Punto Fijo para dos variables como:

$$\begin{aligned}x_{n+1} &= F(x_n, y_n) = \text{sen}(x_n + y_n) \\y_{n+1} &= G(x_n, y_n) = \text{cos}(x_n - y_n);\end{aligned}$$

Observamos que:

$$\begin{aligned}\left| \frac{\partial F(x,y)}{\partial x} \right| + \left| \frac{\partial F(x,y)}{\partial y} \right| &= 2|\text{cos}(x + y)| \leq k_1 = 0.7247 < 1 \\ \left| \frac{\partial G(x,y)}{\partial x} \right| + \left| \frac{\partial G(x,y)}{\partial y} \right| &= 2|\text{sen}(x - y)| \leq k_2 = 0.7788 < 1\end{aligned}$$

en la región $D = [0.6, 1] \times [0.6, 1]$.

Aplicando el algoritmo, tenemos:

n	x_n	y_n
0	0.75	0.75
1	0.9974949866	1
2	0.9103370262	0.9999968625
3	0.9429087819	0.9959832488
4	0.9330142776	0.9985918811
5	0.9356112339	0.9978505594
6	0.934954525	0.9980637583
7	0.935111776	0.9980092732
8	0.9350753557	0.9980226044
9	0.9350835393	0.998019476
10	0.9350817476	0.9980201876
11	0.9350821304	0.9980200301
12	0.9350820506	0.9980200641
13	0.9350820668	0.9980200569
14	0.9350820636	0.9980200584
15	0.9350820642	0.9980200581
16	0.9350820641	0.9980200582
17	0.9350820641	0.9980200582

Tabla 18

Como podrán observar en este caso la convergencia ha sido lenta, esto se debe a que el $k_1 = 0.7247$ y $k_2 = 0.7788$, son próximos a 1 en cada caso; esto es a medida de que las cotas sean mayores y próximas a 1, la convergencia es más lenta.

Algoritmo de Punto Fijo para Sistemas de 2 Ecuaciones:

Objetivo: Hallar (α, β) punto fijo de $G = (F(x, y), G(x, y))$, con una tolerancia ε especificada (ver comentarios).

Paso 1. Hallar una aproximación inicial (x_0, y_0)

Paso 2. Calcular:

$$\begin{aligned} x_{n+1} &= F(x_n, y_n) & n &= 0, 1, 2, \dots \\ y_{n+1} &= G(x_n, y_n) & n &= 0, 1, 2, \dots \end{aligned}$$

Paso 3. Averiguar si se satisface la tolerancia especificada, si fuese así tomar:

$$\begin{aligned} \alpha &\approx x_{n+1} \\ \beta &\approx y_{n+1} \end{aligned}$$

Comentarios:

Los criterios que se emplean frecuentemente para detener el algoritmo de Punto Fijo para dos variables son:

1. Utilizar el concepto de distancia, esto es:

$$d((x_{n+1}, y_{n+1}), (x_n, y_n)) < \varepsilon \quad (\varepsilon > 0)$$

2. Emplear las diferencias para dos puntos sucesivos, en cada coordenada, esto es:

$$\Delta x = x_{n+1} - x_n \quad \text{y} \quad \Delta y = y_{n+1} - y_n;$$

$$\text{tomando luego: } |\Delta x| + |\Delta y| < \varepsilon \quad (\varepsilon > 0)$$

2.9 Método de Newton para Sistemas

El método de Newton para una ecuación no lineal con una variable, $f(x) = 0$, se puede interpretar de la siguiente forma: para un x_n cercano a α , cero de $f(x)$, se define

$$x_{n+1} = x_n + \Delta x_n, \text{ entonces } x_{n+1} \approx \alpha,$$

$$\text{y}$$

$$f(x_{n+1}) = f(x_n + \Delta x_n) \approx 0.$$

Usando la fórmula de Taylor, se obtiene:

$$f(x_n + \Delta x_n) \approx f(x_n) + f'(x_n) \Delta x_n \approx 0,$$

de donde, se puede tomar:

$$\Delta x_n = -f(x_n)/f'(x_n)$$

por lo tanto:

$$x_{n+1} = x_n + \Delta x_n = x_n - f(x_n)/f'(x_n)$$

Por analogía para el caso de dos ecuaciones, tendremos:

$$f(x, y) = 0$$

$$g(x, y) = 0,$$

con x_n cercano a α , e y_n cercano a β , donde (α, β) es solución del sistema dado, se puede tomar:

$$\alpha \approx x_{n+1} = x_n + \Delta x_n,$$

$$\beta \approx y_{n+1} = y_n + \Delta y_n$$

entonces,

$$f(x_n + \Delta x_n, y_n + \Delta y_n) \approx 0.$$

$$g(x_n + \Delta x_n, y_n + \Delta y_n) \approx 0.$$

Si $f, g \in C^2$, usando la Fórmula de Taylor, para funciones de dos variables (ver Apéndice A), se tiene:

$$\begin{aligned} f(x_n, y_n) + f_x(x_n, y_n)\Delta x_n + f_y(x_n, y_n)\Delta y_n &\approx 0 \\ g(x_n, y_n) + g_x(x_n, y_n)\Delta x_n + g_y(x_n, y_n)\Delta y_n &\approx 0 \end{aligned}$$

Si el Jacobiano del sistema:

$$J(f, g)_{(x_n, y_n)} = \begin{vmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{vmatrix} \neq 0$$

se puede resolver el sistema lineal, respecto a las incógnitas $\Delta x_n, \Delta y_n$ usando la regla de Crámer:

$$\Delta x_n = \frac{- \begin{vmatrix} f & f_y \\ g & g_y \end{vmatrix} (x_n, y_n)}{J(f, g) (x_n, y_n)}$$

$$\Delta y_n = \frac{- \begin{vmatrix} f_x & f \\ g_x & g \end{vmatrix} (x_n, y_n)}{J(f, g) (x_n, y_n)}$$

entonces:

$$x_{n+1} = x_n + \Delta x_n, \quad y_{n+1} = y_n + \Delta y_n.$$

Para que esta sucesión $\{(x_n, y_n)\}_{(n \geq 0)}$ generada con el método de Newton para dos variables, converja a la solución (α, β) del sistema dado, se deben cumplir las siguientes condiciones:

1. $f, g \in C^2[V(\alpha, \beta)]$.
2. $J(f, g) \neq 0$ en $V(\alpha, \beta)$.
3. (x_0, y_0) está suficientemente cerca de (α, β) .

Donde $V(\alpha, \beta)$, es una vecindad de (α, β) .

El Teorema de convergencia para el método de Newton y su demostración lo puede leer en Ortega [18].

Algoritmo de Newton para Sistemas de 2 Ecuaciones

Objetivo: Hallar (α, β) , cero del sistema:

$$\begin{aligned}f(x, y) &= 0 \\g(x, y) &= 0\end{aligned}$$

con una tolerancia ε .

Paso 1. Hallar una aproximación inicial (x_0, y_0)

Paso 2. Para $n = 0, 1, 2, \dots$

Calcular

$$\begin{aligned}x_{n+1} &= x_n + \Delta x_n \\y_{n+1} &= y_n + \Delta y_n\end{aligned}$$

$$\text{con } \Delta x_n = \frac{-f(x_n, y_n) g_y(x_n, y_n) + f_y(x_n, y_n) g(x_n, y_n)}{f_x(x_n, y_n) g_y(x_n, y_n) - f_y(x_n, y_n) g_x(x_n, y_n)}$$

$$\text{con } \Delta y_n = \frac{f(x_n, y_n) g_x(x_n, y_n) - f_x(x_n, y_n) g(x_n, y_n)}{f_x(x_n, y_n) g_y(x_n, y_n) - f_y(x_n, y_n) g_x(x_n, y_n)}$$

Paso 3. Verificación de la tolerancia.

Comentarios:

1. La convergencia del método de Newton para dos variables es cuadrática.

2. El método de Newton se puede extender a n ecuaciones con n incógnitas, en forma similar.

Aplicaciones:

Ejemplo 1:

Resolver el sistema:

$$\begin{aligned} f(x, y) &= x - 0.7\text{sen}x - 0.2\text{cos}y = 0 \\ g(x, y) &= y - 0.7\text{cos}x + 0.2\text{sen}y = 0 \end{aligned}$$

tomando $(x_0, y_0) = (0, 0)$, y empleando el método de Newton.

Solución:

Calculando las derivadas parciales:

$$\begin{aligned} f_x &= 1 - 0.7\text{cos}x, & f_y &= 0.2\text{sen}y \\ g_x &= 0.7\text{sen}x, & g_y &= 1 + 0.2\text{cos}y \end{aligned}$$

Se obtiene el Jacobiano de las funciones f y g :

$$J(f, g) = \begin{vmatrix} f_x & f_y \\ g_x & g_y \end{vmatrix} = (1 - 0.7\text{cos}x)(1 + 0.2\text{cos}y) - 0.7\text{sen}x \cdot 0.2\text{sen}y$$

que es distinto de cero en una vecindad de $(0, 0)$.

Entonces aplicando el algoritmo de Newton, para $n = 0, 1, 2, 3, \dots$:

$$x_{n+1} = x_n - \frac{[(x_n - 0.7\text{sen}x_n - 0.2\text{cos}y_n)(1 + 0.2\text{cos}y_n) - (y_n - 0.7\text{cos}x_n + 0.2\text{sen}y_n)(0.2\text{sen}y_n)]}{J(f, g)_{(x_n, y_n)}}$$

$$y_{n+1} = y_n - \frac{[(y_n - 0.7\text{cos}x_n + 0.2\text{sen}y_n)(1 - 0.7\text{cos}x_n) - (x_n - 0.7\text{sen}x_n - 0.2\text{cos}y_n)(0.7\text{sen}x_n)]}{J(f, g)_{(x_n, y_n)}}$$

A partir del cual, se obtienen los siguientes resultados:

n	x_n	y_n
0	0	0
1	0.6666666	0.5833333
2	0.5362400	0.5088490
3	0.5265620	0.5079319
4	0.5265226	0.5079197
5	0.5265226	0.5079197

Tabla 19

Ejemplo 2:

Resolver el sistema:

$$\begin{aligned} f(x, y) &= x + 3\log x - y^2 = 0 \\ g(x, y) &= 2x^2 - xy - 5x + 1 = 0 \end{aligned}$$

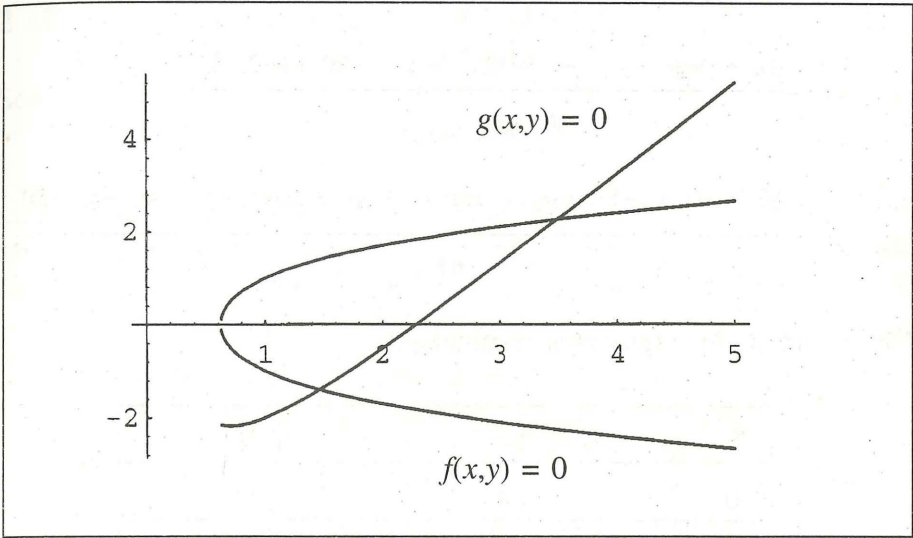
para la solución cercana a $(x_0, y_0) = (3.4, 2.2)$, empleando el método de Newton.

Solución:

La gráfica de las curvas

$$\begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0, \end{aligned}$$

se muestra a continuación:



Gráfica 22

Lo que nos permite observar que hay dos puntos de intersección, uno que está ubicado en la región $[3, 4] \times [2, 3]$ y el otro en la región $[0, 2] \times [-2, 0]$.

Entonces calculando las derivadas parciales:

$$f_x = 1 + (3 \log e)/x; \quad f_y = -2y; \quad g_x = 4x - y - 5; \quad g_y = -x$$

se obtiene el Jacobiano:

$$\begin{aligned} J(f, g) &= \begin{vmatrix} f_x & f_y \\ g_x & g_y \end{vmatrix} = (1 + 3(\log e)/x)(-x) - (4x - y - 5)(-2y) = \\ &= -x - 3 \log e + 8xy - 2y^2 - 10y \end{aligned}$$

que es diferente de cero en una vecindad de $(3.4, 2.2)$.

Se define el algoritmo, para $n = 0, 1, 2, \dots$

$$x_{n+1} = x_n - \frac{[(x_n + 3\log x_n - y_n^2)(-x_n) - (2x_n^2 - x_n y_n - 5x_n + 1)(-2y_n)]}{J(f, g)(x_n, y_n)}$$

$$y_{n+1} = y_n - \frac{[(2x_n^2 - x_n y_n - 5x_n + 1)((1 + (3\log e)/x_n) - (x_n + 3\log x_n - y_n^2)(x_n - y_n - 5)]}{J(f, g)(x_n, y_n)}$$

Obteniéndose los siguientes resultados:

n	x_n	y_n
0	3.4	2.2
1	3.4899123	2.2633644
2	3.4874446	2.2616299
3	3.4874428	2.2616286
4	3.4874428	2.2616286

Tabla 20

Si se utilizase otro punto inicial, que corresponde al cuarto cuadrante $(x_0, y_0) = (1, -2)$, la convergencia es rápida pero en este caso hacia la otra raíz, como lo muestran los resultados de la siguiente tabla:

n	x_n	y_n
0	1	-2
1	1.4759726	-1.5240274
2	1.4573549	-1.4011608
3	1.4588864	-1.3967715
4	1.4588902	-1.3967670
5	1.4588902	-1.3967670

Tabla 21

2.10 Método de Bairstow para Determinar el Factor Cuadrático de un Polinomio

Sea

$$P_N(x) = a_0 x^N + a_1 x^{N-1} + \dots + a_{N-1} x + a_N$$

un polinomio de grado N , con coeficientes reales.

Se sabe que si $P_N(x)$ tiene un cero complejo en $z_1 = a + ib$, también tiene otro cero complejo que es el conjugado de z_1 ; es decir, $z_2 = \bar{z}_1 = a - ib$. Por lo tanto $P_N(x)$ tiene un factor cuadrático, conocido con el nombre de factor cuadrático irreducible que es:

$$(x - z_1)(x - z_2) = x^2 - (z_1 + z_2)x + z_1 z_2$$

Para determinar los ceros complejos de $P_N(x)$ se puede construir el algoritmo que determina los factores cuadráticos de un polinomio.

Este algoritmo, que no usa aritmética compleja, se conoce con el nombre de algoritmo de Bairstow.

Todo polinomio de grado N puede presentarse como:

$$P_N(x) = (x^2 - ux - v)Q_{N-2}(x) + b_{N-1}(x - u) + b_N$$

con

$$Q_{N-2}(x) = b_0 x^{N-2} + b_1 x^{N-3} + \dots + b_{N-2},$$

polinomio de grado $N-2$.

Reemplazando $Q_{N-2}(x)$ en la ecuación e igualando los coeficientes se obtiene:

$$\begin{aligned} a_0 &= b_0 \\ a_1 &= b_1 - ub_0 \\ a_2 &= b_2 - ub_1 - vb_0 \\ &\dots \end{aligned}$$

en general:

$$a_n = b_n - ub_{n-1} - vb_{n-2} \text{ para } n = 2, 3, \dots, N.$$

de donde despejando b_n se tiene:

$$b_n = a_n + ub_{n-1} + vb_{n-2},$$

una relación de recurrencia para b_n , con $n = 0, 1, \dots, N$ y $b_{-1} = b_{-2} = 0$.

Si b_{N-1} y b_N se consideran funciones de u, v tales que:

$$\begin{aligned} b_{N-1} &= a_{N-1} + ub_{N-2} + vb_{N-3} = 0 \\ b_N &= a_N + ub_{N-1} + vb_{N-2} = 0 \end{aligned}$$

observamos que $(x^2 - ux - v)$ es un factor cuadrático de $P_N(x)$.

Entonces se plantea el problema de resolver el sistema:

$$b_{N-1}(u, v) = 0 \text{ y } b_N(u, v) = 0,$$

el cual se puede resolver aplicando el método de Newton.

Construiremos el algoritmo de Newton. Derivando con respecto a u :

$$\begin{aligned} \frac{\partial b_0}{\partial u} &= 0; \quad \frac{\partial b_1}{\partial u} = b_0 \\ \frac{\partial b_2}{\partial u} &= b_1 + u \frac{\partial b_1}{\partial u} \\ \frac{\partial b_3}{\partial u} &= b_2 + u \frac{\partial b_2}{\partial u} + v \frac{\partial b_1}{\partial u} \end{aligned}$$

En general

$$\frac{\partial b_{n+1}}{\partial u} = b_n + u \frac{\partial b_n}{\partial u} + v \frac{\partial b_{n-1}}{\partial u}; \quad n = 2, 3, \dots, (N-1).$$

Si $c_n = \frac{\partial b_{n+1}}{\partial u} = b_n + uc_{n-1} + vc_{n-2}$ y con $c_{-1} = c_{-2} = 0$

se tiene una relación de recurrencia para $n = 0, 1, 2, \dots, N-1$.

Luego para $n = N-2$,

$$c_{N-2} = \frac{\partial b_{N-1}}{\partial u}$$

y para $n = N-1$, tenemos:

$$c_{N-1} = \frac{\partial b_N}{\partial u}$$

que son las derivadas parciales que intervienen en el algoritmo de Newton.

Derivando con respecto a v :

$$\frac{\partial b_0}{\partial v} = 0; \quad \frac{\partial b_1}{\partial v} = 0; \quad \frac{\partial b_2}{\partial v} = b_0;$$

$$\frac{\partial b_3}{\partial v} = b_1 + u \frac{\partial b_2}{\partial v} + v \frac{\partial b_1}{\partial v}$$

$$\frac{\partial b_4}{\partial v} = b_2 + u \frac{\partial b_3}{\partial v} + v \frac{\partial b_2}{\partial v}, \dots$$

y en general:

$$\frac{\partial b_{n+2}}{\partial v} = b_n + u \frac{\partial b_{n+1}}{\partial v} + v \frac{\partial b_n}{\partial v}; \quad n = 2, 3, \dots, N-1.$$

Si se define

$$d_n = \frac{\partial b_{n+2}}{\partial v} = b_n + u d_{n-1} + v d_{n-2} \quad \text{y con} \quad d_{-1} = d_{-2} = 0$$

para $n = 0, 1, 2, \dots, N-2$, se obtiene la misma relación de recurrencia que se obtuvo para c_n . Luego para $n = N-3$,

$$c_{N-3} = \frac{\partial b_{N-1}}{\partial v},$$

y para $n = N-2$

$$c_{N-2} = \frac{\partial b_N}{\partial v},$$

que son las otras derivadas parciales que intervienen en el algoritmo de Newton.

Entonces en el algoritmo de Newton para determinar u y v , tenemos:

$$\begin{aligned} u_{n+1} &= u_n + \Delta u_n \\ v_{n+1} &= v_n + \Delta v_n \end{aligned}$$

en donde:

$$\Delta u_n = \frac{- \begin{vmatrix} b_{N-1} & c_{N-3} \\ b_N & c_{N-2} \end{vmatrix}}{\begin{vmatrix} c_{N-2} & c_{N-3} \\ c_{N-1} & c_{N-2} \end{vmatrix}}; \quad \Delta v_n = \frac{- \begin{vmatrix} c_{N-2} & b_{N-1} \\ c_{N-1} & b_N \end{vmatrix}}{\begin{vmatrix} c_{N-2} & c_{N-3} \\ c_{N-1} & c_{N-2} \end{vmatrix}}$$

Algoritmo de Bairstow

Objetivo: Hallar un factor cuadrático de la forma $x^2 - ux - v$ del polinomio $P_N(x)$ dados los valores iniciales aproximados u_0, v_0 , de u, v respectivamente.

Paso 1: Para $k = 0, 1, 2, \dots, N-1, N$ calcular:

$$b_k = a_k + u_0 b_{k-1} + v_0 b_{k-2} \text{ con } b_{-1} = b_{-2} = 0.$$

Paso 2: Para $k = 0, 1, 2, \dots, N-3, N-2, N-1$, calcular:

$$c_k = b_k + u_0 c_{k-1} + v_0 c_{k-2} \text{ con } c_{-1} = c_{-2} = 0$$

Paso 3: Tomar:

$$\Delta u = \frac{-(b_{N-1} c_{N-2} - b_N c_{N-3})}{(c_{N-2}^2 - c_{N-1} c_{N-3})}; \quad \Delta v = \frac{-(b_N c_{N-2} - b_{N-1} c_{N-1})}{(c_{N-2}^2 - c_{N-1} c_{N-3})}$$

Paso 4: Tomar $u = u + \Delta u$; $v = v + \Delta v$

Comentarios:

1. Como criterio para detener el algoritmo se puede usar:

$$\text{TOLERANCIA: } |\Delta u| + |\Delta v| < \varepsilon \quad (\varepsilon > 0)$$

2. En cada paso iterativo es útil presentar a los resultados en el siguiente formato:

n	a_n	b_n	c_n	$\Delta u, \Delta v$
0	a_0	b_0	c_0	
1	a_1	b_1	c_1	
...
$N-2$	a_{N-2}	b_{N-2}	c_{N-2}	
$N-1$	a_{N-1}	b_{N-1}	c_{N-1}	
N	a_N	b_N		

Tabla 22

3. Sólo los valores b_{N-1}, b_N de la columna b_n , y $c_{N-3}, c_{N-2}, c_{N-1}$ de la columna c_n se usan para construir Δu , y Δv .
4. Se puede observar que la columna de b_n nos da los coeficientes del polinomio $Q_{N-2}(x)$ de grado $(N-2)$, que puede utilizarse para determinar en forma aproximada los otros ceros de $P_N(x)$, ya que $b_{N-1} \approx 0$; $b_N \approx 0$.

Ejemplo:

Sea

$$P_3(x) = x^3 + 0.5x^2 - 2x + 7.5$$

Como es un polinomio de grado impar entonces $P_3(x)$ tiene por lo menos una raíz real, que se encuentra en el intervalo $] -3, -2[$, ya que $P_3(-2) > 0$ y $P_3(-3) < 0$.

Si dividimos $P_3(x)$ entre $(x + 2)$, se puede obtener un factor cuadrático aproximado $(x^2 - 1.5x + 1)$ lo que permite tomar los valores iniciales

$$u_0 = 1.5, v_0 = -1.$$

Aplicaremos ahora el algoritmo de Bairstow con los valores iniciales u_0 , y v_0 .

Paso 1: Consiste en determinar:

$$\begin{aligned}u_1 &= u_0 + \Delta u \\v_1 &= v_0 + \Delta v,\end{aligned}$$

donde Δu , Δv se obtienen a partir de los coeficientes b_k , c_k en función de a_k , coeficientes del polinomio dado, y los valores iniciales u_0 , v_0 obtenidos anteriormente.

Como

$$\begin{aligned}b_k &= a_k + u_0 b_{k-1} + v_0 b_{k-2} \\c_k &= b_k + u_0 c_{k-1} + v_0 c_{k-2} \\u_0 &= 1.5, v_0 = -1.0\end{aligned}$$

Los resultados de estos cálculos pueden presentarse a través de la tabla:

k	a_k	b_k	c_k
0	1.0	1.0	1.0
1	0.5	2.0	3.5
2	-2.0	0.0	4.25
3	7.5	5.5	

Tabla 23

entonces:

$$\begin{aligned}\Delta u &= 0.6875 \\ \Delta v &= -2.40625\end{aligned}$$

entonces:

$$\begin{aligned}u_1 &= u_0 + \Delta u = 1.5 + 0.6875 = 2.1875 \\v_1 &= v_0 + \Delta v = -1 - 2.40625 = -3.40625\end{aligned}$$

Paso 2: Consiste en determinar u_2 , v_2 . Para la aplicación del algoritmo tomamos $u_0 = 2.1875$ y $v_0 = -3.40625$. Entonces con las fórmulas dadas anteriormente se obtiene la tabla para b_k , c_k en la forma:

k	a_k	b_k	c_k
0	1.0	1.0	1.0
1	0.5	2.6875000	4.875
2	-2.0	0.472656250	7.73046875
3	7.5	-0.6203613285	

Tabla 24

entonces:

$$\Delta u = -0.1823842875$$

$$\Delta v = 0.4164671514$$

por lo tanto:

$$u_2 = 2.1875 - 0.1823842875 = 2.005115713$$

$$v_2 = -3.40625 + 0.4164671514 = -2.989782849$$

Paso 3: Tomando

$$u_0 = 2.005115713$$

$$v_0 = -2.989782849$$

se obtiene:

k	a_k	b_k	c_k
0	1.0	1.0	1.0
1	0.5	2.505115713	4.510231426
2	-2.0	0.03326403	6.087017083
3	7.5	0.07694623574	

Tabla 25

Entonces,

$$\Delta u = -0.005126717923$$

$$\Delta v = -0.01014134572$$

luego:

$$u_3 = 2.005115713 - 0.005126717923 = 1.999988995$$

$$v_3 = -2.989782849 - 0.01014134572 = -2.999924195$$

Para comprobar los resultados tomaremos:

$$\begin{aligned}u_0 &= 2 \\v_0 &= -3\end{aligned}$$

entonces se obtiene:

k	a_k	b_k
0	1.0	1.0
1	0.5	2.5
2	-2.0	0
3	7.5	0

Tabla 26

y como $b_2 = b_3 = 0$ entonces el polinomio se puede escribir como:

$$P_3(x) = (x + 2.5)(x^2 - 2x + 3)$$

y el factor cuadrático

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

nos proporciona las raíces complejas de $P_3(x)$:

$$x_1 = 1 + \sqrt{2}i \quad \text{y} \quad x_2 = 1 - \sqrt{2}i$$

Comentario:

En el ejemplo anterior si tomamos como aproximaciones iniciales $u_0 = v_0 = 0$, y aplicamos el Algoritmo de Bairstow, tendremos:

En el primer paso los valores b_k y c_k iguales a a_k , entonces la tabla es de la forma:

k	a_k	b_k	c_k
0	1.0	1.0	1.0
1	0.5	0.5	0.5
2	-2.0	-2.0	-2.0
3	7.5	7.5	

Tabla 27

De donde

$$\begin{aligned}\Delta u &= 3.777777778 \\ \Delta v &= 0.111111\end{aligned}$$

por lo tanto:

$$u_1 = u_0 + \Delta u = 3.777777778$$

$$v_1 = v_0 + \Delta v = 0.111111111$$

En el segundo paso tomamos como aproximaciones iniciales:

$$u_0 = 3.777777778$$

$$v_0 = 0.111111111$$

y se obtiene,

k	a_k	b_k	c_k
0	1.0	1.0	1.0
1	0.5	4.277777778	8.055555556
2	-2.0	14.27160494	44.81481484
3	7.5	61.89026064	

Tabla 28

Si observamos los valores obtenidos $b_{N-1} = b_2$ y $b_N = b_3$ y los comparamos con los valores b_{N-1} , b_N de la primera tabla, podemos afirmar que el proceso no es convergente ya que estos valores han crecido en vez de ir disminuyéndose. Lo que prueba que las aproximaciones para u y v nulas no son suficientemente "buenas".

2.11 Ejercicios

1. Dada la ecuación:

$$f(x) = x^2 - 4x \operatorname{sen} x + (2 \operatorname{sen} x)^2 = 0$$

- Analice donde se ubican sus raíces, por medio de una gráfica.
- Halle una raíz positiva empleando el método de Bisección, con una tolerancia de 10^{-3} .

2. Dada

$$f(x) = [(\operatorname{sen} x) / x^2] - 1/2 = 0,$$

encuentre todas sus raíces empleando:

- El método de Punto Fijo con una tolerancia = 10^{-5} , si $TOL = |f(x_n)|$
 - Si se quisiera una precisión de $\varepsilon = 10^{-10}$, ¿cuántas iteraciones se deberían realizar?
3. Dada la ecuación:

$$f(x) = 2x^2 + 6e^{-x} - 4 = 0$$

- Analice gráficamente donde se ubican las raíces.
 - Use el método de Punto Fijo (verificando las condiciones) para calcular las raíces con una precisión de 10^{-3} , empleando procedimientos diferentes.
 - Analizando la cota para el error absoluto diga, ¿Cuántas iteraciones se deberían realizar en cada caso para lograr una precisión de 10^{-8} ?
4. Dada la ecuación:

$$f(x) = \tan^{-1} x = 0$$

- Emplee el método de Newton para hallar la raíz de dicha ecuación, en los intervalos:
 - $[-2, 2]$, con $x_0 = 1.8$ como aproximación inicial; presentando gráficamente los resultados para 3 iteraciones.

(ii) $[-1.3, 1.3]$, con $x_0 = 1.1$; presentando gráficamente los resultados para 3 iteraciones.

b. Analice los resultados en (i) y en (ii).

5. Resolver la ecuación:

$$f(x) = (4x-7) / (x-2) = 0$$

a. Aplicando el método de Newton, para los siguientes valores iniciales:

i. $x_0 = 1.625$

ii. $x_0 = 1.5$

iii. $x_0 = 3$

b. Analice cada uno de los resultados hallados en la parte (a).

6. Dadas las ecuaciones:

i. $f(x) = 4x^4 - 9x^3 - 1 = 0$ iii. $f(x) = x^3 - e^x + 3 = 0$

ii. $f(x) = x^3 - 5e^x + 3 = 0$ iv. $f(x) = x^2 - e^{2x} - 1 = 0$

v. $f(x) = e^{-x} - \cos x = 0$

a. Ubicar las raíces.

b. Emplear el método de Bisección para hallar dichas raíces con una aproximación de $\varepsilon = 10^{-6}$.

c. Emplee el método de Punto Fijo para hallar dichas raíces con una aproximación de $\varepsilon = 10^{-6}$.

d. Repita el procedimiento del ítem (c), empleando el valor x_5 de la sucesión obtenida en (b), como aproximación inicial.

7. Sean las ecuaciones:

i. $f(x) = \ln x + 1 - 4(x-2)^2 = 0$;

ii. $f(x) = e^{-x^2} - 3 + (x-2)^2 = 0$

a. Haciendo uso de un gráfico, ubique sus raíces.

b. Aplique el método de Bisección para hallar cada una de las raíces, con $\varepsilon = 10^{-8}$, y haciendo uso del teorema, halle una cota superior para el error absoluto.

- c. Aplique el método de Punto Fijo, para la raíz de mayor magnitud con $\varepsilon = 10^{-8}$, verificando las condiciones del teorema.
8. Una determinada sustancia se desintegra según la ecuación: $A = Pe^{-0.0248t}$, donde P es la cantidad inicial en el tiempo $t = 0$ y A la cantidad resultante después de t años. Si inicialmente se depositan 500 miligramos de dicha sustancia, ¿Cuánto tiempo habrá de transcurrir para que quede el uno por ciento de ésta? Utilice el método de Newton, hallando previamente una “buena” aproximación con el método de Bisección, con una tolerancia de 10^{-5} , si

$$TOL = \left| \frac{t_{i+1} - t_i}{t_{i+1}} \right|.$$

9. Sea:

$$f(x) = e^x - 2x - 21 = 0, \quad x \in \mathfrak{R}$$

- Analice gráficamente las raíces.
 - Utilice el método de Newton para hallar las raíces, con $\varepsilon = 10^{-6}$
 - Emplee el método de Punto Fijo para hallar las raíces de $f(x)$, verificando las condiciones.
10. Sea:

$$f(x) = 2^{-x} + e^x + 2 \cos x - 6, \quad x \in [1, 3]$$

- Aplique Bisección para hallar una solución aproximada, con $TOL = 10^{-2}$.
 - Utilice la solución aproximada hallada en (a) como aproximación inicial para aplicar el método de Newton.
11. Dado:

$$P(x) = x^3 - 6.8x^2 + 15.41x - 11.638$$

- Haga la gráfica en el intervalo $[2, 3]$, y analice el comportamiento de $P(x)$ en el entorno de $x = 2.2$.
- Halle una aproximación inicial x_0 adecuada, de manera que el método de Newton converja a $x = 2.2$. Aplique el método, hasta obtener la mayor precisión.

12. Sea $P(x) = 5x^3 - 18.55x^2 + 22.9395x - 9.455625$
- $P(x)$ tiene una raíz en $x = 1.25$, use el método de Newton con una aproximación inicial:
 - $x_0 = 1.235$,
 - $x_0 = 1.239$,
 - $x_0 = 1.24$
 para tratar de encontrar esta raíz.
 - Haga un bosquejo de la gráfica de la ecuación.
 - ¿Qué pasa?. ¿Cómo podría obtener una primera aproximación lo suficientemente buena, para que el método converja a $x=1.25$? Halle la raíz con una precisión de $\varepsilon = 10^{-6}$.

13. Aplique el método modificado de Newton para hallar la raíz múltiple de:

$$f(x) = \operatorname{sen}^2 x - x \operatorname{sen} x + x^2 / 4 = 0$$

Comenzando con $x_0 = \pi/2$; con una tolerancia $= 10^{-6}$, si $TOL = |x_{i+1} - x_i|$

14. a. Pruebe que si z es una raíz de multiplicidad m de un polinomio $P(x)$, entonces

$$P(z) = P'(z) = \dots = P^{(m-1)}(z) = 0 \text{ y } P^{(m)}(z) \neq 0.$$

- b. Sea $P(x) = x^5 - 4x^4 + 14x^2 - 17x + 6$, un polinomio con raíz en $z = 1$ de multiplicidad 3,
- Halle dos iteraciones empleando el método de Newton, con $x_0 = 0$.
 - Halle las dos primeras iteraciones con el método modificado de Newton, con $x_0 = 0$.
 - Analice ambos resultados.

15. Un Método para encontrar una raíz de la ecuación $f(x) = 0$ que se encuentra en el intervalo $[a, b]$ es el llamado de Posición Falsa o Regula Falsi, que consiste en lo siguiente:

“Si el intervalo $[a_i, b_i]$ contiene una raíz de $f(x) = 0$, se calcula la intersección con el eje X de la recta que une los puntos $(a_i, f(a_i))$, $(b_i, f(b_i))$, denotando este punto por x_i . Si $f(a_i)f(x_i) < 0$, se define $a_{i+1} = a_i$ y $b_{i+1} = x_i$, de otra manera, se define $a_{i+1} = x_i$ y $b_{i+1} = b_i$.

Construya un algoritmo para este método, indicando los pasos a seguir. Luego aplíquelo, para hallar las tres primeras iteraciones, a fin de resolver la ecuación $f(x) = 3x - 2 + e^x - x^2 = 0$.

16. Resuelva la ecuación:

$$f(x) = 0.5 - x + 0.2\text{sen}x = 0$$

- Use el algoritmo de Aitken para calcular una sucesión acelerada \hat{x}_k .
- Use el algoritmo de Steffensen para calcular una sucesión acelerada y_k .

Sugerencia: Trabaje hasta con las $n = 9$ primeras iteraciones para Punto Fijo, y use $x_0 = 0.5$.

17. Utilice el Método de Steffensen para resolver las ecuaciones:

- $f(x) = x - 1/2 - \text{sen}x = 0$
- $e^{-x} = \text{cos}x$ en $[\pi/4, \pi/2]$

verifique las condiciones del Método de Punto Fijo.

18. Resolver la ecuación:

$$f(x) = 3x^2 - e^x = 0$$

en el intervalo $[0, 1]$.

Construya dos funciones $g(x)$ diferentes, y aplique el método de Steffensen a cada una con $\varepsilon = 10^{-6}$.

19. Resolver:

$$f(x) = x^3 + 5e^x + 3$$

- Usando el método de Punto Fijo con $TOL = 10^{-3}$, verifique las condiciones.
- Aplique el método de Steffensen, con un $TOL = 10^{-6}$.

20. La función de transferencia para un sistema está dada por:

$$F(s) = H(s)/(1+G(s)H(s))$$

donde

$$G(s) = \frac{1}{s} e^{-0.1s}, \quad H(s) = K$$

- Si $K = 3$, halle usando el método de Punto Fijo, la mayor de las raíces de la ecuación $1 + G(s)H(s) = 0$. Verifique las condiciones del método con una precisión de $\varepsilon = 10^{-4}$.
 - Aplique el método de Steffensen para acelerar la convergencia de la sucesión obtenida con Punto Fijo.
21. Si se compra una pieza de equipo en \$20,000 en abonos, pagando \$5,000 durante 5 años. ¿Qué tasa de interés I se está pagando? La fórmula que relaciona el costo actual (P), los pagos anuales (A), el número de años (n) y la tasa de interés es:

$$A = P \left(\frac{I(1+I)^n}{(1+I)^n - 1} \right)$$

- Haga una estimación gráfica de la raíz I , de

$$f(I) = \frac{I(1+I)^n}{(1+I)^n - 1} - \frac{A}{P} = 0$$

- Halle 5 iteraciones, empleando el método de Bisección y diga ¿cómo se comporta el error absoluto correspondiente a la aproximación lograda?

c. Como:

$$f'(x_n) = \lim_{x \rightarrow x_n} \left[\frac{f(x) - f(x_n)}{x - x_n} \right]$$

y tomando $x = x_{n-1}$

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

y utilizando este resultado en la fórmula de Newton:

$$x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}, \quad n \geq 1$$

que recibe el nombre del método de la secante.

Aplice este método para hallar la raíz de $f(I) = 0$, empleando 5 iteraciones.

Compare ambos resultados.

22. El valor acumulado en una cuenta de ahorros basada en pagos periódicos regulares puede determinarse de la *ecuación de vencimiento anual*:

$$A = \frac{P}{I} [(1+I)^n - 1]$$

en esta ecuación, A es la cantidad en la cuenta, P es la cantidad depositada regularmente, e I es la tasa de interés por período para los n períodos de depósito.

A una persona le gustaría tener una cantidad de \$2500 en una cuenta de ahorros al finalizar un año y para este fin puede depositar \$150 al mes durante el año.

- a. Hacer un bosquejo de la gráfica, en el intervalo $[0, 0.1]$

- b. Hallar la función g que permita calcular la tasa mínima de interés en $[0, 0.1]$ por medio del método de Punto Fijo para lograr esta cantidad de dinero. Verifique las condiciones.
 - c. Aplicar el método de Steffensen para determinar esta tasa de interés, con una precisión de $\varepsilon = 10^{-6}$.
 - d. Aplique el método de Newton para determinar esta tasa de interés.
 - e. Si en la ecuación se cambia $A = \$6000$ y $P = 400$. Haga un bosquejo de la gráfica y aplique el método de Newton para hallar la tasa mínima.
23. Un jugador A dejará en cero (ganará con una anotación de 21-0) a un jugador B en un juego de racketball con probabilidad:

$$P(p) = \left(\frac{1+p}{2} \right) \left(\frac{p}{1-p+p^2} \right)^{21}$$

donde p denota la probabilidad de que A gane un punto cualquiera. Se desea que A gane a B en la mitad de los juegos en los que ellos se enfrentan.

- a. Hacer un bosquejo de la gráfica, en el intervalo $[0, 1]$
 - b. Hallar una función g que permita calcular la probabilidad mínima en $[0, 1]$ por medio del método de Punto Fijo para lograr que A gane a B en la mitad de los juegos. Verifique las condiciones.
 - c. Aplique el método de Steffensen para determinar esta probabilidad p , con una precisión de $\varepsilon = 10^{-6}$.
 - d. Aplique el método de Newton para determinar esta probabilidad p .
 - e. Si se desea que A gane a B en un 60% de los juegos. Haga un bosquejo de la gráfica y aplique el método de Newton para hallar la probabilidad mínima.
24. El crecimiento de grandes poblaciones puede ser modelado en períodos cortos suponiendo que la población crece continuamente en el tiempo a una tasa proporcional al número de individuos presentes en ese lapso. Si denotamos por $N(t)$ al número de individuos en el tiempo t , y λ ($0 < \lambda < 1$) es la tasa de natalidad de la población, y si además se considera que existe inmigración a una tasa constante v , la ecuación diferencial que gobierna esta situación se convierte en,

$$\frac{dN(t)}{dt} = \lambda N(t) + v, \text{ cuya solución es } N(t) = N_0 e^{\lambda t} + \frac{v}{\lambda} (e^{\lambda t} - 1)$$

Suponga que:

$$N(t) = 1564000$$

$$N_0 = 1000000$$

$$v = 435000$$

$$t = 1 \text{ año}$$

- Realice 4 iteraciones usando el método de Bisección, para hallar una aproximación de λ .
- Realice 2 iteraciones usando el método de Newton, empleando como aproximación inicial para λ , el valor obtenido en (a).

25. Emplee el Método de Punto Fijo para resolver el sistema:

$$x^4 + y^4 - 67 = 0$$

$$x^3 - 3xy^2 + 35 = 0$$

cerca de $x = 2, y = 2$.

26. Resuelva el sistema:

$$x^2 + y^2 - x = 0$$

$$x^2 - y^2 - y = 0$$

haciendo uso del método de Punto Fijo en la región $R = [0, 1] \times [0, 1]$.

27. Dado el sistema:

$$f(x, y) = 16x^2 + y^2 - 16 = 0$$

$$g(x, y) = y - 3e^x - 1 = 0$$

- Hacer un bosquejo de la gráfica de $f(x, y)$ y $g(x, y)$. Explique. ¿Cuántas soluciones tiene el sistema.
- Hallar las funciones F y G , para aplicar el método del Punto Fijo para dos variables, a fin de encontrar la solución que está ubicada en el 2do. cuadrante del sistema de coordenadas cartesianas.

- c. Reduzca el sistema dado a una sola ecuación y halle la raíz correspondiente al ítem (b), aplicando el método de Newton.

28. Dado el sistema:

$$\begin{aligned}f(x, y) &= xy - 1 = 0 \\g(x, y) &= e^x - e^y - 2 = 0\end{aligned}$$

- Hacer un bosquejo de la gráfica de f y g . ¿Cuántas soluciones tiene el sistema.
- Hallar las funciones F y G , para aplicar el método del Punto Fijo para dos variables, y encontrar la solución que está ubicada en el 1er. cuadrante del sistema de coordenadas cartesianas.

29. Resolver el sistema:

$$\begin{aligned}f(x, y) &= 2x^2 - xy - 5x + 1 = 0 \\g(x, y) &= x + 3\ln x - y^x = 0\end{aligned}$$

en la región $R = [3, 4] \times [2, 3]$.

- Construya dos sistemas diferentes y equivalentes al dado, de la forma:

$$\begin{aligned}x &= F(x, y) \\y &= G(x, y)\end{aligned}$$

- Aplique el algoritmo de Punto Fijo para dos variables, a cada uno de los sistemas hallados anteriormente

30. Dado el sistema:

$$\begin{aligned}f(x, y) &= 4(x+1)^2 + (y+1)^2 - 16 = 0 \\g(x, y) &= y - e^x\end{aligned}$$

en la región $R = [0, 1] \times [1, 2]$.

Construya dos sistemas diferentes y equivalentes al dado, de la forma:

$$\begin{aligned}x &= F(x, y) \\y &= G(x, y)\end{aligned}$$

y aplique el algoritmo de Punto Fijo para dos variables, a cada uno de los sistemas hallados anteriormente y encuentre la solución en dicha región.

31. Dado el sistema:

$$\begin{aligned} f(x, y) &= 4x^2 - 9y^2 - 9 = 0 \\ g(x, y) &= (x - 1)^2 + (y - 1)^2 - 4 = 0 \end{aligned}$$

- Analice por medio de una gráfica, donde se ubican las soluciones del sistema.
 - Empleando el método de Punto Fijo halle la solución que se ubica en el cuarto cuadrante del sistema de coordenadas cartesianas.
 - Aplique el método de Steffensen a la sucesión obtenida en (b), para acelerar la convergencia.
32. Hallar la intersección ubicada en el 1er. Cuadrante del Sistema de Coordenadas Cartesianas, de las curvas:

- $$\begin{aligned} f(x, y) &= (x-1)^2 + (y-1)^2 - 9 = 0 \\ g(x, y) &= y^2 - x^2 - 1 = 0 \end{aligned}$$
- $$\begin{aligned} f(x, y) &= 0.5\text{sen}(xy) - 0.0795y - 0.5x = 0 \\ g(x, y) &= 0.9204(e^{2x} - 2.7182) + 0.8652y - 5.43565x = 0 \end{aligned}$$

cercana a (0.7623, 3.5985), empleando el Método de Newton.

33. Sea el sistema:

$$\begin{aligned} f(x, y) &= (x-2)^2 + (y-2)^2 - 4 = 0 \\ g(x, y) &= 2(x-4)^2 + 3(y-3)^2 - 12 = 0 \end{aligned}$$

- Analice gráficamente donde se ubican las raíces del sistema.
 - Aplicar el método de Newton, para hallar las dos soluciones del sistema.
34. Aplicar el Método de Bairstow a:

$$x^4 + 2x^3 + 7x^2 - 11 = 0$$

para hallar todas sus raíces con seis cifras exactas.

35. Aplicar el Método de Bairstow a:

a. $x^6 - x^4 - x^3 - 1 = 0$

para obtener un factor cuadrático cercano a $x^2 + 1.32x + 0.7424$.

b. $x^4 + 5x^3 - 3x^2 - 5x - 9 = 0$

para obtener un factor cuadrático cercano a $x^2 + 3x - 5$.

36. Hallar el factor cuadrático que contenga a las raíces complejas de los polinomios:

a. $P(x) = 2x^3 + x^2 - 3x + 2 = 0$

b. $P(x) = x^3 + x^2 - x - 6 = 0$

c. $P(x) = -x^3 + 6x^2 - 3x + 4 = 0$

37. Dado el polinomio:

$$P(x) = 50 - 70x + 39x^2 - 10x^3 + x^4$$

a. Hallar las raíces complejas que correspondan al factor cuadrático aproximado a:

$$x^2 - 3x + 4$$

empleando el método de Bairstow.

b. Halle las otras dos raíces complejas del polinomio $P(x)$, empleando el método de Bairstow.

38. Dado el polinomio:

$$P(x) = 10 - 14x + 11x^2 - 4x^3 + x^4$$

a. Hallar las raíces complejas que correspondan al factor cuadrático aproximado a:

$$x^2 - 1.5x + 2.3$$

empleando el método de Bairstow.

b. Halle las otras dos raíces complejas del polinomio $P(x)$, empleando el método de Bairstow.

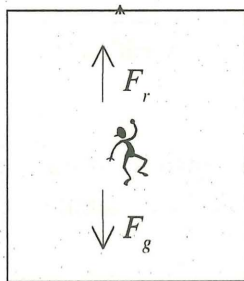
Solución de Sistemas de Ecuaciones Lineales

Muchas de las ecuaciones fundamentales de la Ingeniería se basan en las leyes de la conservación y forman parte de ellas, cantidades como la masa, la fuerza, la energía, el momento, ..., etc., leyes que en términos matemáticos nos conducen a sistemas de ecuaciones lineales. Así mismo en Química o en Física encontramos fenómenos de la naturaleza que matemáticamente quedan expresados en términos de sistemas lineales.

En Física, recordemos por ejemplo, la segunda ley del movimiento de Newton, que afirma que la velocidad de cambio de la cantidad de movimiento de un cuerpo con respecto al tiempo es igual a la fuerza resultante que actúa sobre él; por lo tanto la expresión matemática para esta segunda ley será:

$$F = ma$$

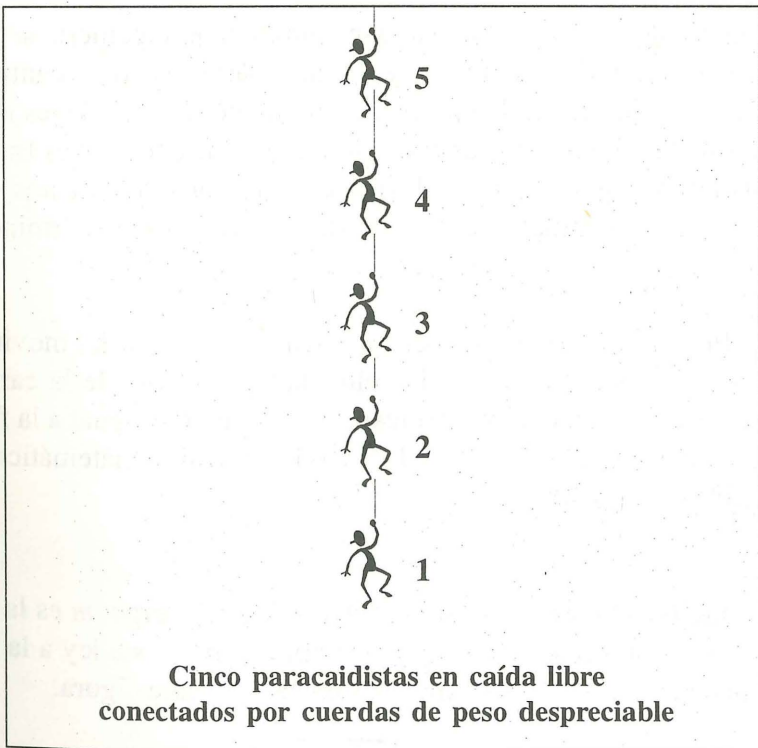
Donde F es la fuerza neta que actúa sobre el cuerpo, m es la masa del objeto y a su aceleración. Luego si aplicásemos esta ley a la caída de un paracaidista, como se muestra en la siguiente figura:



Gráfica 1

se tendrá la ecuación: $F = m \frac{dv}{dt}$, donde v es la velocidad de caída del paracaidista y $F = F_g + F_r$, en la cual $F_g = mg$ es la fuerza debida a la atracción de la gravedad y $F_r = -cv$ es la fuerza debida a la resistencia del aire (c coeficiente de rozamiento).

Supongamos que un grupo de cinco paracaidistas se conectan por medio de cuerdas muy ligeras mientras que caen a una velocidad de 5 m/s, como se muestra en la figura:

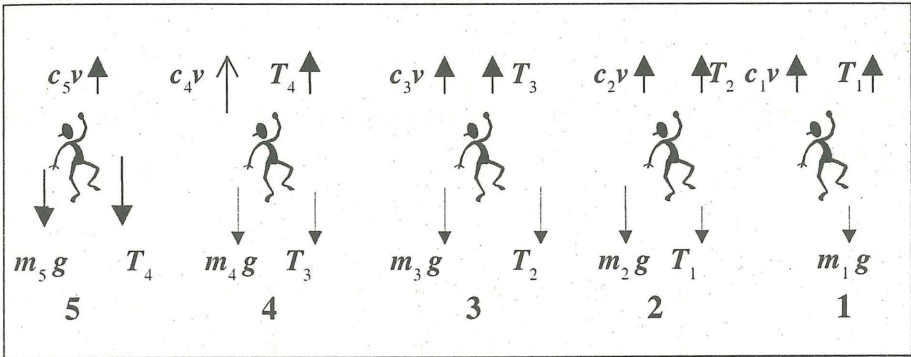


Gráfica 2

Se desea calcular la tensión en cada sección de la cuerda y la aceleración del grupo, dada la siguiente información:

Paracaidista	Masa, Kg.	Coefficientes de rozamiento
1	70	10
2	60	12
3	50	14
4	60	17
5	80	16

Los diagramas de cuerpo libre para cada uno de los tres paracaidistas se muestran en el siguiente gráfico:



Gráfica 3

lo que lleva a las siguientes ecuaciones basándonos en la segunda ley de Newton:

$$\begin{aligned}
 m_1 g - T_1 - c_1 v &= m_1 a \\
 m_2 g + T_1 - c_2 v - T_2 &= m_2 a \\
 m_3 g + T_2 - c_3 v - T_3 &= m_3 a \\
 m_4 g + T_3 - c_4 v - T_4 &= m_4 a \\
 m_5 g + T_4 - c_5 v &= m_5 a
 \end{aligned}$$

con $g = 9.8 \text{ m/s}^2$.

Basándonos en los datos dados se obtiene el siguiente sistema lineal:

$$686 - T_1 - 50 = 70a$$

$$588 + T_1 - 60 - T_2 = 60a$$

$$490 + T_2 - 70 - T_3 = 50a$$

$$588 + T_3 - 85 - T_4 = 60a$$

$$784 + T_4 - 80 = 80a$$

Como se ve, se reduce a un sistema lineal de 5 ecuaciones, donde las incógnitas serán T_1 , T_2 , T_3 , T_4 , y a . Para resolver este sistema, hay diferentes modos de hacerlo, en este capítulo trataremos algunos de ellos.

En general, consideremos un sistema lineal de n ecuaciones con n incógnitas de la forma:

$$\begin{aligned} E_1: & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ E_2: & a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ E_3: & a_{31}x_1 + a_{32}x_2 + \dots + a_{3n}x_n = b_3 \\ & \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ E_n: & a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{aligned}$$

que puede escribirse en forma matricial como:

$$Ax = b \quad \text{ó} \quad \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix}$$

con $A = [a_{ij}]_{(i,j=1,2,\dots,n)}$; una matriz cuadrada de orden n , en donde los elementos a_{ij} , reciben el nombre de elementos pivotes de A y los vectores

$$\text{columnas } n\text{- dimensionales } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix}$$

Para determinar la solución aproximada del sistema (si esta existe), existen métodos numéricos directos e iterativos. Estos métodos permiten hallar la solución numérica del sistema en forma aproximada, sin necesidad de hacer el análisis de existencia de la misma.

3.1 Métodos Directos de Solución de Sistemas Lineales

Los métodos numéricos directos son aquellos que usan un número finito de operaciones, a diferencia de los métodos iterativos, cuyo procedimiento infinito se corta en algún momento originando el error del método; es decir el error de truncamiento, además de los errores del cálculo. Los métodos numéricos directos están sujetos sólo a los errores del cálculo (errores del redondeo), por lo tanto el resultado siempre es una solución aproximada.

Entre ellos el más conocido es el método de Eliminación Gaussiana. Para definir este método vemos el caso particular del sistema lineal: $A\mathbf{x} = \mathbf{b}$, donde A matriz triangular superior, y todas las entradas diagonales no nulas.

Definición:

Se dice que $A = [a_{ij}]_{(i,j=1,2,\dots,n)}$, es una matriz triangular superior si:

$$a_{ij} = 0, \quad \forall i > j.$$

Algoritmo de Sustitución Reversa

Que puede expresarse en la forma:

$$x_k = \frac{b_k - \sum_{j=k+1}^n a_{kj} x_j}{a_{kk}} ; \text{ si } a_{kk} \neq 0 \text{ y para } k = n, n-1, \dots, 1.$$

En el caso de ser $k = n$, la sumatoria se toma como cero.

Para el caso de una matriz A no triangular superior, se usa el procedimiento de Eliminación de Gauss, que consiste en eliminar las incógnitas de las ecuaciones lineales, que están por debajo de la diagonal, transformando así el sistema: $Ax = b$ en un sistema $Ux = c$, sistemas equivalentes en el sentido de que tengan la misma solución, pero con la matriz U triangular superior.

Existen tres operaciones que pueden usarse con tal fin:

1. Multiplicar a una ecuación por una constante no nula:

$$\lambda E_i \rightarrow E_i$$

2. Intercambiar dos ecuaciones de lugar dentro del sistema:

$$E_i \leftrightarrow E_j$$

3. Sumar un múltiplo de una ecuación a otra:

$$E_i + \lambda E_j \rightarrow E_i$$

La ecuación E_j se conoce como ecuación pivotal.

Estas operaciones transforman al sistema, $Ax = b$ en el sistema equivalente $Ux = c$.

Para el procedimiento de Eliminación Gaussiana se usará el arreglo inicial $[A, b]$, y se realizarán los siguientes pasos:

Paso 1: consiste en determinar la ecuación pivotal, será aquella donde esté presente la 1^{ra} incógnita x_1 , esta ecuación se llamará E_1 y se usará para eliminar la incógnita x_1 de las restantes $(n-1)$ -ecuaciones.

Usando la operación (3):

$E_i - \lambda E_1 \rightarrow E_i$; para $i = 2, 3, \dots, n$
 con $\lambda = m_{i1} = a_{i1} / a_{11}$, para $i = 2, 3, \dots, n$, $a_{11} \neq 0$; el arreglo inicial se convierte en:

$$[A, \mathbf{b}] = \begin{matrix} E_1 \\ E_2 \\ E_3 \\ \vdots \\ E_n \end{matrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{bmatrix} \rightarrow \begin{matrix} E_1 \\ E_2 \\ E_3 \\ \vdots \\ E_n \end{matrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ 0 & a_{32}^{(1)} & \cdots & a_{3n}^{(1)} & b_3^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{bmatrix}$$

La primera columna de ceros se puede usar para guardar en la memoria de la computadora a los multiplicadores m_{i1} .

Paso 2: consiste primero, en determinar la ecuación pivotal, llamándola E_2 , (donde está presente la 2^{da} variable x_2) y se procede luego a eliminar la incógnita x_2 de las restantes $(n-2)$ -ecuaciones: E_3, E_4, \dots, E_n , usando la operación:

$$E_i - \lambda E_2 \rightarrow E_i \text{ con } i = 3, 4, \dots, n \text{ y } \lambda = m_{i2} = a_{i2} / a_{22}; i = 3, 4, \dots, n.$$

obteniendo el arreglo $[A, \mathbf{b}]$ en la siguiente forma:

$$\begin{array}{l}
 E_1 \\
 E_2 \\
 E_3 \\
 \vdots \\
 E_n
 \end{array}
 \left[
 \begin{array}{cccccc}
 a_{11} & a_{12} & \cdots & \cdots & a_{1n} & b_1 \\
 0 & a_{22}^{(1)} & \cdots & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\
 0 & 0 & a_{33}^{(2)} & & a_{3n}^{(2)} & b_3^{(2)} \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 0 & 0 & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)}
 \end{array}
 \right]$$

Al aplicar $(n-1)$ - veces el proceso de Eliminación Gaussiana al arreglo inicial $[A, b]$, se obtiene el arreglo final:

$$\begin{array}{l}
 E_1 \\
 E_2 \\
 E_3 \\
 \vdots \\
 E_{n-1} \\
 E_n
 \end{array}
 \left[
 \begin{array}{cccccc}
 a_{11} & a_{12} & \cdots & \cdots & \cdots & a_{1n} & b_1 \\
 0 & a_{22}^{(1)} & \cdots & & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\
 0 & 0 & a_{33}^{(2)} & \cdots & \cdots & a_{3n}^{(2)} & b_3^{(2)} \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 0 & 0 & 0 & \cdots & a_{n-1,n-1}^{(n-2)} & a_{n-1,n}^{(n-2)} & b_{n-1}^{(n-2)} \\
 0 & 0 & 0 & \cdots & 0 & a_{nn}^{(n-1)} & b_n^{(n-1)}
 \end{array}
 \right]$$

que es un sistema $Ux = c$ con la matriz U triangular superior, y que es equivalente al sistema $Ax = b$. Luego, aplicando el Algoritmo de Sustitución Reversiva, al sistema $Ux = c$ después de un número finito de operaciones, se obtiene una solución del sistema $Ax = b$.

Algoritmo de Eliminación Gaussiana con Sustitución hacia atrás

Objetivo: Para resolver el sistema lineal de $n \times n$ ecuaciones:

$$\begin{array}{l}
 E_1: a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = a_{1(n+1)} = b_1 \\
 E_2: a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = a_{2(n+1)} = b_2 \\
 E_3: a_{31}x_1 + a_{32}x_2 + \dots + a_{3n}x_n = a_{3(n+1)} = b_3 \\
 \vdots \\
 E_n: a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = a_{n(n+1)} = b_n
 \end{array}$$

con los pivotes $a_{ii} \neq 0, \forall i$.

Hacer para:

Paso 1 :

$$\begin{aligned} j &= 1, \dots, n-1 \\ i &= j + 1, \dots, n \end{aligned}$$

los pasos 2-3 (Proceso de Eliminación).

Paso 2: Tomar $m_{ij} = a_{ij} / a_{ii}$.

Paso 3: Efectuar $(E_i - m_{ij} E_j) \rightarrow (E_i)$.

Hacer para $k = n, n-1, \dots, 1$ (Algoritmo de Sustitución Reversa)
los pasos 4 y 5:

Paso 4: Tomar $x_n = a_{n(n+1)} / a_{nn}$.

Paso 5: Para $k = n - 1, n - 2, \dots, 1$ tomar $x_k = [a_{k(n+1)} - \sum_{j=k+1}^n a_{kj} x_j] / a_{kk}$

Ejemplo 1

Resolver el sistema lineal:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ -6 & 6 & 12 & 6 \\ 3 & 3 & -3 & 4 \\ 2 & -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 36 \\ -1 \\ 10 \end{bmatrix} \begin{matrix} \dots E_1 \\ \dots E_2 \\ \dots E_3 \\ \dots E_4 \end{matrix}$$

aplicando el método de Eliminación Gaussiana:

Solución:

Paso 1: Usaremos como ecuación pivotal a la fila E_1 , ya que $a_{11} \neq 0$.
Para eliminar los elementos de la primera columna debajo de la diagonal se realizarán las siguientes operaciones con las ecuaciones:

$$\begin{aligned} (6)E_1 + E_2 &\rightarrow E_2 \\ (-3)E_1 + E_3 &\rightarrow E_3 \end{aligned}$$

$$(-2)E_1 + E_4 \rightarrow E_4$$

obteniéndose el siguiente sistema:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 18 & 18 & 18 \\ 0 & -3 & -6 & -2 \\ 0 & -5 & -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 72 \\ -19 \\ -2 \end{bmatrix} \begin{matrix} \dots E_1 \\ \dots E_2 \\ \dots E_3 \\ \dots E_4 \end{matrix}$$

Paso 2: Usaremos como ecuación pivotal a la fila E_2 . Para eliminar las entradas de la segunda columna debajo de la diagonal, se realizarán las siguientes operaciones:

$$\begin{aligned} (1/6)E_2 + E_3 &\rightarrow E_3 \\ (5/18)E_2 + E_4 &\rightarrow E_4 \end{aligned}$$

obteniéndose el siguiente sistema:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 18 & 18 & 18 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 72 \\ -7 \\ 18 \end{bmatrix} \begin{matrix} \dots E_1 \\ \dots E_2 \\ \dots E_3 \\ \dots E_4 \end{matrix}$$

Paso 3: Usaremos como ecuación pivotal a la fila E_3 . Para anular las entradas de la tercera columna debajo de la diagonal, se realizará la siguiente operación:

$$(4/3)E_3 + E_4 \rightarrow E_4$$

obteniéndose el arreglo triangular superior final:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 18 & 18 & 18 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 13/3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 72 \\ -7 \\ 26/3 \end{bmatrix} \begin{matrix} \dots E_1 \\ \dots E_2 \\ \dots E_3 \\ \dots E_4 \end{matrix}$$

y aplicando el Algoritmo de Sustitución Reversiva se obtiene:

$$\begin{aligned} x_4 &= 2 \\ -3x_3 + x_4 &= -7 \Rightarrow x_3 = 3 \\ 18x_2 + 18x_3 + 18x_4 &= 72 \Rightarrow x_2 = -1 \\ x_1 + 2x_2 + x_3 + 2x_4 &= 6 \Rightarrow x_1 = 1 \end{aligned}$$

entonces la solución del sistema será:

$$x = [1 \ -1 \ 3 \ 2]^t$$

Comentarios sobre el método de Eliminación Gaussiana:

1. El procedimiento de Eliminación Gaussiana se puede aplicar a sistemas que no tienen solución ($\det A = 0$), o a aquellos que tienen más de una solución. Sólo que en estos casos, el algoritmo se detiene en el momento de tener una fila de ceros, sin posibilidad de avanzar. Se recomienda en este caso, detenerse y analizar la situación.

Ejemplo 2:

Podemos observar que en el siguiente sistema lineal:

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

el algoritmo de Eliminación Gaussiana falla, ya que realizando el primer paso del algoritmo, con ecuación pivotal la fila E_1 , se obtendrá el arreglo de la forma $[A, b]$:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 4 \\ 1 & 1 & -1 & -1 & 0 \\ 1 & 1 & 2 & 0 & 3 \\ 1 & 1 & 1 & -1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & 4 \\ 0 & 0 & -2 & -2 & -4 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & -2 & -2 \end{bmatrix}$$

donde vemos que los elementos a_{22} , a_{32} , a_{42} son iguales a cero, lo que no permitiría realizar el siguiente paso del algoritmo.

Analizando esta situación, notamos que de la última ecuación se tiene:

$$-2x_4 = -2 \Rightarrow x_4 = 1,$$

de la penúltima ecuación se tiene

$$x_3 - x_4 = -1 \Rightarrow x_3 = -1 + 1 = 0$$

y de la segunda ecuación se tiene:

$$-2x_3 - 2x_4 = -4$$

lo que es incompatible con los resultados obtenidos. En este caso el sistema no tiene solución.

Ejemplo 3:

Podemos observar que en el siguiente sistema lineal:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 4 \\ 2 \end{bmatrix}$$

el algoritmo de Eliminación Gaussiana falla, ya que realizando el primer paso del algoritmo, con ecuación pivotal la fila E_1 , se obtendrá el arreglo de la forma $[A, \mathbf{b}]$:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 4 \\ 1 & 1 & -1 & -1 & 0 \\ 1 & 1 & 2 & 0 & 4 \\ 1 & 1 & 1 & -1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & 4 \\ 0 & 0 & -2 & -2 & -4 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -2 & -2 \end{bmatrix}$$

Resolviendo, de la última ecuación se tiene

$$-2x_4 = -2 \Rightarrow x_4 = 1,$$

de la penúltima ecuación se tiene

$$x_3 - x_4 = 0 \Rightarrow x_3 = x_4 = 1$$

y de la segunda ecuación se tiene

$$-2x_3 - 2x_4 = -4$$

lo que prueba los resultados anteriores para x_3 y x_4 . Y de la primera ecuación se tiene

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 4 \Rightarrow \\ x_1 &= 4 - x_2 - x_3 - x_4 = 4 - x_2 - 2 = 2 - x_2 \end{aligned}$$

lo que implica que el sistema tiene infinitas soluciones.

2. Las operaciones elementales realizadas con las ecuaciones del sistema lineal, afectan al determinante de la matriz A en la siguiente forma:
 - a. Al sumar a una fila un múltiplo de otra no altera el valor de $\det(A)$.
 - b. Un intercambio de filas introduce un signo $(-)$ en el valor del $\det(A)$.
 - c. La multiplicación de una fila por una constante también multiplica por esta constante al $\det(A)$.

Tomando en cuenta estas observaciones y como $\det(U) = u_{11}u_{22}u_{33}\dots u_{mm}$; ésto es, el producto de los elementos de la diagonal del último arreglo triangular superior $[A, \mathbf{b}]$, se puede obtener $\det(A)$, haciendo uso de este resultado.

Ejemplo 4:

Con relación al **Ejemplo 1** para la aplicación del método de Eliminación Gaussiana, el determinante de A estará dado por el determinante de U ; esto es: $\det(A) = \det(U) = 18(-3)(13/3) = -234$.

3. Si guardamos los multiplicadores m_{ij} en los espacios con ceros del arreglo $[A, \mathbf{b}]$, se obtiene el arreglo final en la forma:

$$\begin{array}{l}
 E_1 \\
 E_2 \\
 E_3 \\
 \vdots \\
 \vdots \\
 E_n
 \end{array}
 \left[\begin{array}{cccccc}
 a_{11} & a_{12} & \cdots & \cdots & a_{1n} & b_1 \\
 m_{21} & a_{22}^{(1)} & \cdots & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\
 m_{31} & m_{32} & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & b_3^{(2)} \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
 m_{n1} & m_{n2} & \cdots & m_{n,n-1} & a_{nn}^{(n-1)} & b_n^{(n-1)}
 \end{array} \right]$$

lo que nos permite resolver el sistema $A\mathbf{x} = \mathbf{d}$, con $\mathbf{d} \neq \mathbf{b}$, aplicando las operaciones realizadas sólo para la última columna del arreglo.

Esta idea de resolver el sistema $A\mathbf{x} = \mathbf{b}$ para distintos vectores \mathbf{b} , se puede utilizar para construir la inversa de la matriz A .

3.2 Cálculo de la Inversa de una Matriz

Sea A una matriz cuadrada de orden n , entonces A^{-1} es inversa de A , sí y sólo si

$$AA^{-1} = A^{-1}A = I,$$

con $I = [e_{ij}]_{(i,j=1, 2, \dots, n)}$ y $e_{ij} = \begin{cases} 1, & i = j; \\ 0, & i \neq j \end{cases}$;

es decir,

$$I = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}_{n \times n}$$

Si designamos a los vectores columna de la matriz I , como:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_k = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \rightarrow k\text{-ésima componente, y } e_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

entonces

$$I = [e_1, e_2, \dots, e_n].$$

Resolviendo los n sistemas, se obtienen las columnas de A^{-1} ; es decir:

1. Del sistema $Ax = e_1 \Rightarrow x = A^{-1}e_1$, se obtiene la 1^{ra} Columna de A^{-1} .
2. Del sistema $Ax = e_2 \Rightarrow x = A^{-1}e_2$, se obtiene la 2^{da} Columna de A^{-1}
-
- n . Del sistema $Ax = e_n \Rightarrow x = A^{-1}e_n$, se obtiene la n -ésima Columna de A^{-1}

De esta forma el proceso de construcción de la inversa de A se reduce al proceso de resolver n sistemas lineales:

$$Ax = e_k, \quad k = 1, 2, \dots, n$$

con la misma matriz dada A y con n vectores distintos e_k .

Para realizar este procedimiento se puede construir el arreglo aumentado $[A, I] = [A, e_1, e_2, \dots, e_n]$ y aplicar la Eliminación Gaussiana, obteniendo el arreglo final en la forma

$$[A, I] \rightarrow [U, b_1, b_2, \dots, b_n]$$

esto es, con la matriz U triangular superior, lo que permite determinar las soluciones para n vectores distintos b_k , usando n veces el Algoritmo de Sustitución Reversiva.

Ejemplo 1:

Calcular la inversa de la matriz:

$$A = \begin{bmatrix} 6 & -2 & 2 & 4 \\ 12 & -8 & 6 & 10 \\ 3 & -13 & 9 & 3 \\ -6 & 4 & 1 & -18 \end{bmatrix}$$

Solución:

Sea $B = [b_1, b_2, b_3, b_4] = A^{-1}$, hallemos cada uno de los vectores columna de la inversa de A , resolviendo cada uno de los sistemas lineales siguientes:

(i) Resolviendo el sistema lineal:

$$Ax = e_1$$

$$\begin{bmatrix} 6 & -2 & 2 & 4 \\ 12 & -8 & 6 & 10 \\ 3 & -13 & 9 & 3 \\ -6 & 4 & 1 & -18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

se tiene

$$b_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3.486111 \\ 8.2916666 \\ 11.916666 \\ 3.6666666 \end{bmatrix}$$

(ii) Resolviendo el sistema lineal:

$$Ax = e_2$$

$$\begin{bmatrix} 6 & -2 & 2 & 4 \\ 12 & -8 & 6 & 10 \\ 3 & -13 & 9 & 3 \\ -6 & 4 & 1 & -18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

se tiene:

$$b_2 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2.1527777 \\ -4.7916666 \\ -6.9166666 \\ -2.1666666 \end{bmatrix}$$

(iii) Resolviendo el sistema lineal:

$$Ax = e_3$$

$$\begin{bmatrix} 6 & -2 & 2 & 4 \\ 12 & -8 & 6 & 10 \\ 3 & -13 & 9 & 3 \\ -6 & 4 & 1 & -18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

se tiene:

$$b_3 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -0.694444 \\ 1.416666 \\ 2.166666 \\ 0.666666 \end{bmatrix}$$

(iv) Resolviendo el sistema lineal:

$$Ax = e_4$$

$$\begin{bmatrix} 6 & -2 & 2 & 4 \\ 12 & -8 & 6 & 10 \\ 3 & -13 & 9 & 3 \\ -6 & 4 & 1 & -18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

se tiene

$$b_4 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0.305555 \\ -0.583333 \\ -0.833333 \\ -0.333333 \end{bmatrix}$$

Por lo tanto la inversa de A , es:

$$A^{-1} = \begin{bmatrix} -3.48611 & 2.152777 & -0.69444 & 0.30555 \\ 8.29166 & -4.791666 & 1.41666 & -0.58333 \\ 11.9166 & -6.916666 & 2.16666 & -0.83333 \\ 3.66666 & -2.16666 & 0.66666 & -0.33333 \end{bmatrix}$$

Comentario:

Una variación del método anterior, que se conoce con el nombre de procedimiento de Gauss-Jordan, consiste en llevar el arreglo inicial $[A, I]$ a la forma $[I, B]$ con $B = A^{-1}$. Lo que se logra, primero anulando las entradas debajo de la diagonal y luego anulando las entradas sobre la diagonal.

Entonces el procedimiento consiste en:

$$[A, I] \rightarrow \dots \rightarrow [I, A^{-1}]$$

Ejemplo 2:

Hallar la inversa de la matriz A del **Ejemplo 1**, empleando el método de Gauss-Jordan.

Solución:

El arreglo aumentado será:

$$\left[\begin{array}{cccc|cccc} 6 & -2 & 2 & 4 & 1 & 0 & 0 & 0 & \dots E_1 \\ 12 & -8 & 6 & 10 & 0 & 1 & 0 & 0 & \dots E_2 \\ 3 & -13 & 9 & 3 & 0 & 0 & 1 & 0 & \dots E_3 \\ -6 & 4 & 1 & -18 & 0 & 0 & 0 & 1 & \dots E_4 \end{array} \right]$$

- (i) Para anular los elementos debajo del pivote a_{11} , realizaremos las operaciones siguientes:

$$\begin{aligned} (-2)E_1 + E_2 &\rightarrow E_2 \\ (-1/2)E_1 + E_3 &\rightarrow E_3, \\ E_1 + E_4 &\rightarrow E_4 \end{aligned}$$

lográndose el siguiente arreglo:

$$\left[\begin{array}{cccc|cccc} 6 & -2 & 2 & 4 & 1 & 0 & 0 & 0 & \dots E_1 \\ 0 & -4 & 2 & 2 & -2 & 1 & 0 & 0 & \dots E_2 \\ 0 & -12 & 8 & 1 & -1/2 & 0 & 1 & 0 & \dots E_3 \\ 0 & 2 & 3 & -14 & 1 & 0 & 0 & 1 & \dots E_4 \end{array} \right]$$

- (ii) Para anular los elementos debajo del pivote a_{22} , realizaremos las operaciones siguientes:

$$\begin{aligned} (-3)E_2 + E_3 &\rightarrow E_3 \\ (1/2)E_2 + E_4 &\rightarrow E_4 \end{aligned}$$

lográndose el siguiente arreglo:

$$\left[\begin{array}{cccc|cccc} 6 & -2 & 2 & 4 & 1 & 0 & 0 & 0 & \dots E_1 \\ 0 & -4 & 2 & 2 & -2 & 1 & 0 & 0 & \dots E_2 \\ 0 & 0 & 2 & -5 & 11/2 & -3 & 1 & 0 & \dots E_3 \\ 0 & 0 & 4 & -13 & 0 & 1/2 & 0 & 1 & \dots E_4 \end{array} \right]$$

- (iii) Para anular los elementos debajo del pivote a_{33} , realizaremos las operaciones siguientes:

$$(-2)E_3 + E_4 \rightarrow E_4$$

lográndose el siguiente arreglo:

$$\left[\begin{array}{cccc|cccc} 6 & -2 & 2 & 4 & 1 & 0 & 0 & 0 & \dots E_1 \\ 0 & -4 & 2 & 2 & -2 & 1 & 0 & 0 & \dots E_2 \\ 0 & 0 & 2 & -5 & 11/2 & -3 & 1 & 0 & \dots E_3 \\ 0 & 0 & 0 & -3 & -11 & 13/2 & -2 & 1 & \dots E_4 \end{array} \right]$$

- (iv) Para anular los elementos sobre el pivote a_{44} , realizaremos las operaciones siguientes:

$$\begin{aligned} (-5/3)E_4 + E_3 &\rightarrow E_3 \\ (2/3)E_4 + E_2 &\rightarrow E_2 \\ (4/3)E_4 + E_1 &\rightarrow E_1 \end{aligned}$$

lográndose el siguiente arreglo:

$$\left[\begin{array}{cccc|cccc} 6 & -2 & 2 & 0 & -41/3 & 26/3 & -8/3 & 4/3 & \dots E_1 \\ 0 & -4 & 2 & 0 & -28/3 & 16/3 & -4/3 & 2/3 & \dots E_2 \\ 0 & 0 & 2 & 0 & 143/6 & -83/6 & 13/3 & -5/3 & \dots E_3 \\ 0 & 0 & 0 & -3 & -11 & 13/2 & -2 & 1 & \dots E_4 \end{array} \right]$$

- (v) Para anular los elementos sobre el pivote a_{33} , realizaremos las operaciones siguientes:

$$\begin{aligned} (-1)E_3 + E_2 &\rightarrow E_2 \\ (-1)E_3 + E_1 &\rightarrow E_1 \end{aligned}$$

lográndose el siguiente arreglo:

$$\left[\begin{array}{cccc|cccc} 6 & -2 & 0 & 0 & -225/6 & 135/6 & -21/3 & 9/3 & \dots E_1 \\ 0 & -4 & 0 & 0 & -199/6 & 115/6 & -17/3 & 7/3 & \dots E_2 \\ 0 & 0 & 2 & 0 & 143/6 & -83/6 & 13/3 & -5/3 & \dots E_3 \\ 0 & 0 & 0 & -3 & -11 & 13/2 & -2 & 1 & \dots E_4 \end{array} \right]$$

(vi) Para anular los elementos sobre el pivote a_{22} realizaremos las operaciones siguientes:

$$(-1/2)E_2 + E_1 \rightarrow E_1$$

lográndose el siguiente arreglo:

$$\begin{array}{cccc|cccc} 6 & 0 & 0 & 0 & -251/12 & 155/12 & -25/6 & 11/6 & \dots E_1 \\ 0 & -4 & 0 & 0 & -199/6 & 115/6 & -17/3 & 7/3 & \dots E_2 \\ 0 & 0 & 2 & 0 & 143/6 & -83/6 & 13/3 & -5/3 & \dots E_3 \\ 0 & 0 & 0 & -3 & -11 & 13/2 & -2 & 1 & \dots E_4 \end{array}$$

(vii) Para obtener en la primera mitad del arreglo a la matriz identidad dividiremos cada fila entre su pivote; esto es:

$$\begin{array}{l} (1/6)E_1 \rightarrow E_1 \\ (-1/4)E_2 \rightarrow E_2 \\ (1/2)E_3 \rightarrow E_3 \\ (-1/3)E_4 \rightarrow E_4 \end{array}$$

lográndose el siguiente arreglo:

$$\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -251/72 & 155/72 & -25/36 & 11/36 & \dots E_1 \\ 0 & 1 & 0 & 0 & 199/24 & -115/24 & 17/12 & -7/12 & \dots E_2 \\ 0 & 0 & 1 & 0 & 143/12 & -83/12 & 13/6 & -5/6 & \dots E_3 \\ 0 & 0 & 0 & 1 & 11/3 & -13/6 & 2/3 & -1/3 & \dots E_4 \end{array}$$

hallándose finalmente la inversa de A:

$$A^{-1} = \begin{bmatrix} -251/72 & 155/72 & -25/36 & 11/36 \\ 199/24 & -115/24 & 17/12 & -7/12 \\ 143/12 & -83/12 & 13/6 & -5/6 \\ 11/3 & -13/6 & 2/3 & -1/3 \end{bmatrix}$$

3.3 Factorización Matricial Triangular

El proceso de factorización de una matriz cuadrada A consiste en determinar dos matrices triangulares L y U , tales que $A = LU$.

Si estas matrices L y U poseen características especiales, su conocimiento puede ser muy útil en la práctica, cuando se quiere trabajar con la matriz A .

Veamos el caso, cuando la matriz $U = [u_{ij}]_{(i,j=1,2,\dots,n)}$ es una matriz triangular superior, esto es:

$$u_{ij} = 0 \quad \text{para todo } i > j, \quad i, j = 1, 2, \dots, n$$

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & u_{nn} \end{bmatrix};$$

y la matriz $L = [l_{ij}]_{(i,j=1,2,\dots,n)}$ como matriz triangular inferior, donde

$$l_{ij} = 0 \quad \text{para todo } i < j, \quad i, j = 1, 2, \dots, n$$

$$L = \begin{bmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ l_{21} & l_{22} & 0 & \cdots & 0 \\ l_{31} & l_{32} & l_{33} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ l_{n1} & l_{n2} & \cdots & \cdots & l_{nn} \end{bmatrix}.$$

Se puede ver, que si el método de Eliminación Gaussiana se ha realizado sin intercambio de filas, genera dos matrices de este tipo: la matriz U que es el arreglo triangular superior final del método de Eliminación Gaussiana, y la matriz L que se construye con los multiplicadores m_{ij} , usados para la Eliminación Gaussiana en la forma:

$$L = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ m_{21} & 1 & 0 & \dots & 0 \\ m_{31} & m_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & m_{n3} & \dots & 1 \end{bmatrix}; \quad \text{esto es: } l_{ij} = \begin{cases} m_{ij}, & i > j \\ 1, & i = j \\ 0, & i < j \end{cases}$$

Formando el producto LU , podemos comprobar que sus filas son las filas de A :

$$LU = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ m_{21} & 1 & 0 & \dots & 0 \\ m_{31} & m_{32} & 1 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ m_{n1} & m_{n2} & m_{n3} & \dots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & u_{22} & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{33} & \dots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & u_{nn} \end{bmatrix}$$

Multiplicando la 1^{ra} fila de L por las columnas de U se obtiene la 1^{ra} fila de la matriz producto: $u_{11}, u_{12}, \dots, u_{1n}$, que es también 1^{ra} fila de la matriz A , ya que la Eliminación Gaussiana se realizó sin intercambios de las filas y por lo tanto la 1^{ra} ecuación es la ecuación pivotal en el paso 1 de la eliminación, es decir: $u_{1j} = a_{1j}$, para $j = 1, 2, \dots, n$.

Para formar la 2^{da} fila del producto, tomaremos la 2^{da} fila de L y la multiplicamos por las columnas de U ; entonces los elementos de la 2^{da} fila del producto son:

$$[(m_{21}a_{11}), (m_{21}a_{12} + u_{22}), (m_{21}a_{13} + u_{23}), \dots, (m_{21}a_{1n} + u_{2n})];$$

Pero:

$$\begin{array}{lll} m_{21}a_{11} = a_{21}, & \text{ya que} & m_{21} = a_{21}/a_{11} \\ m_{21}a_{12} + u_{22} = a_{22} & \text{ya que} & u_{22} = a_{22} - m_{21}a_{12} \\ m_{21}a_{13} + u_{23} = a_{23} & \text{ya que} & u_{23} = a_{23} - m_{21}a_{13} \\ \dots\dots\dots & & \dots\dots\dots \\ m_{21}a_{1n} + u_{2n} = a_{2n} & \text{ya que} & u_{2n} = a_{2n} - m_{21}a_{1n} \end{array}$$

lo que prueba que la 2^{da} fila de LU es igual a la 2^{da} fila de A . Avanzando con el proceso y reconstruyendo u_{ij} a través de los multiplicadores m_{ij} y los elementos de A , se puede probar que todas las filas de la matriz LU coinciden con las filas de la matriz A . Por lo tanto: $A = LU$.

Ejemplo 1:

Con relación al **Ejemplo 1** para el método de Eliminación Gaussiana, tendremos:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -6 & 1 & 0 & 0 \\ 3 & -1/6 & 1 & 0 \\ 2 & -5/18 & -4/3 & 1 \end{bmatrix}; \quad U = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 18 & 18 & 18 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 13/3 \end{bmatrix}$$

y al multiplicar L por U , se comprueba que $LU = A$.

Comentario:

Si la Eliminación Gaussiana se realiza con intercambios de filas, entonces la matriz triangular superior U , coincide con el último arreglo obtenido en la Eliminación Gaussiana, pero la matriz L , en general ya no es el arreglo triangular inferior formado por los elementos m_{ij} .

Factorización Directa de una matriz A

Sin necesidad de realizar el método de Eliminación Gaussiana, se puede plantear el problema de Factorización Directa de una matriz cuadrada A , descomponiéndola en el producto de dos matrices triangulares L y U .

Se conoce tres casos de Factorización Directa:

- (1) El método de Doolittle, al tomar $l_{ii} = 1, \forall i=1, 2, \dots, n$,
- (2) El método de Crout, al tomar $u_{ii} = 1, \forall i=1, 2, \dots, n$, y
- (3) El método de Choleski, al tomar $u_{ii} = l_{ii}, \forall i=1, 2, \dots, n$

Veremos uno de ellos digamos el método de Crout, tomando $u_{ii} = 1$, para el caso de la matriz A de orden 4. Los otros dos métodos usan las mismas ideas, que trabajaremos en este método.

Suponiendo que $A = LU$, buscaremos los elementos de L y U , realizando la multiplicación de las dos matrices en la forma más eficiente posible.

Tenemos: $A = LU$ ó

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & u_{13} & u_{14} \\ 0 & 1 & u_{23} & u_{24} \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Paso 1: Para determinar la 1^{ra} columna de L , tomaremos las filas de L y las multiplicaremos por la 1^{ra} columna de U , y obtendremos la 1^{ra} columna de A :

$$\begin{aligned} l_{11} &= a_{11} \\ l_{21} &= a_{21} \\ l_{31} &= a_{31} \\ l_{41} &= a_{41} \end{aligned}$$

Paso 2: Para determinar la 1^{ra} fila de U , tomaremos la 1^{ra} fila de L y la multiplicaremos por la columna de U , obteniéndose la 1^{ra} fila de A :

$$\begin{aligned} l_{11} &= a_{11} \\ l_{11}u_{12} &= a_{12} & \Rightarrow & u_{12} = a_{12} / l_{11} \\ l_{11}u_{13} &= a_{13} & \Rightarrow & u_{13} = a_{13} / l_{11} \\ l_{11}u_{14} &= a_{14} & \Rightarrow & u_{14} = a_{14} / l_{11} \end{aligned}$$

Paso 3: Para determinar la 2^{da} columna de L , es decir, l_{22} , l_{32} , l_{42} , tomaremos las filas (2^{da}, 3^{ra}, 4^{ta}) de L y las multiplicaremos por la 2^{da} columna de U , obteniéndose la 2^{da} columna de A :

$$\begin{aligned}
 l_{21}u_{12} + l_{22} &= a_{22} & \Rightarrow & & l_{22} &= a_{22} - l_{21}u_{12} \\
 l_{31}u_{12} + l_{32} &= a_{32} & \Rightarrow & & l_{32} &= a_{32} - l_{31}u_{12} \\
 l_{41}u_{12} + l_{42} &= a_{42} & \Rightarrow & & l_{42} &= a_{42} - l_{41}u_{12}.
 \end{aligned}$$

Procediendo de la misma forma, es decir alternando la obtención de las columnas de L , y de las filas de U , se tiene:

Paso 4: La 2^{da} fila de U ,

$$\begin{aligned}
 l_{21}u_{13} + l_{22}u_{23} &= a_{23} & \Rightarrow & & u_{23} &= (a_{23} - l_{21}u_{13}) / l_{22} \\
 l_{21}u_{14} + l_{22}u_{24} &= a_{24} & \Rightarrow & & u_{24} &= (a_{24} - l_{21}u_{14}) / l_{22}
 \end{aligned}$$

Paso 5: La 3^{ra} columna de L ,

$$\begin{aligned}
 l_{31}u_{13} + l_{32}u_{23} + l_{33} &= a_{33} & \Rightarrow & & l_{33} &= a_{33} - l_{31}u_{13} - l_{32}u_{23} \\
 l_{41}u_{13} + l_{42}u_{23} + l_{43} &= a_{43} & \Rightarrow & & l_{43} &= a_{43} - l_{41}u_{13} - l_{42}u_{23}
 \end{aligned}$$

Paso 6: La 3^{ra} fila de U ,

$$l_{31}u_{14} + l_{32}u_{24} + l_{33}u_{34} = a_{34} \Rightarrow u_{34} = (a_{34} - l_{31}u_{14} - l_{32}u_{24}) / l_{33}$$

Paso 7: La 4^{ta} columna de A , l_{44} :

$$a_{44} = l_{41}u_{14} + l_{42}u_{24} + l_{43}u_{34} + l_{44} \Rightarrow l_{44} = a_{44} - l_{41}u_{14} - l_{42}u_{24} - l_{43}u_{34}.$$

Este procedimiento se puede generalizar para una matriz de orden n , a través del siguiente algoritmo:

Algoritmo de Factorización Triangular de Crout

Objetivo: Descomponer a la matriz A , en el producto de dos matrices triangulares L y U ; esto es,

$$A = LU$$

Paso 1: Para determinar la 1^{ra} columna de la matriz L , se tiene:

$$l_{i1} = a_{i1}$$

Paso 2: Para determinar la 1^{ra} fila de la matriz U , se tiene:

$$u_{1j} = a_{1j} / l_{11} = a_{1j} / a_{11}$$

Paso 3: Para determinar la j -ésima columna de la matriz L , calcular:

$$l_{ij} = a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \quad \text{para } j \leq i, \quad i = 1, 2, \dots, n$$

Paso 4: Para determinar la i -ésima fila de la matriz U , calcular:

$$u_{ij} = \frac{\left(a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj} \right)}{l_{ii}}, \quad \text{para } i < j, \quad j = 1, 2, \dots, n.$$

Ejemplo 2:

Sea la matriz de orden 3:

$$A = \begin{bmatrix} 4 & 0 & 2 \\ -1 & 3 & 1 \\ 1 & -1 & 5 \end{bmatrix}$$

supongamos que se desea aplicar la factorización triangular de Crout. Luego basándose en los resultados obtenidos anteriormente:

a. Primera columna de L :

$$l_{i1} = a_{i1}; \text{ ésto es: } l_{11} = 4; l_{21} = -1; l_{31} = 1$$

b. Primera fila de U :

$$u_{1j} = a_{1j}/l_{11}; \text{ ésto es } u_{12} = 0 \text{ y } u_{13} = 1/2$$

c. Segunda columna de L :

$$l_{i2} = a_{i2} - l_{i1}u_{12}, \text{ con } i = 2, 3; \text{ ésto es:}$$

$$l_{22} = 3 \quad \text{y} \quad l_{32} = -1$$

d. Segunda fila de U :

$$u_{23} = [a_{23} - l_{21}u_{13}]/l_{22} = 1/2$$

e. Tercera columna de L :

$$l_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23} = 5$$

Concluimos que:

$$L = \begin{bmatrix} 4 & 0 & 0 \\ -1 & 3 & 0 \\ 1 & -1 & 5 \end{bmatrix}; \quad U = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 1 \end{bmatrix}$$

de manera que, $A = LU$.

Comentarios:

1. El método descrito se conoce como el esquema compacto, ya que economiza el espacio de almacenamiento. Conociendo la forma de las matrices L y U , no hay necesidad de almacenar ceros o unos sino otros

elementos esenciales de LU , es decir, el arreglo A se transforma en LU de la forma:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \rightarrow \dots \rightarrow \begin{bmatrix} l_{11} & u_{12} & u_{13} & u_{14} \\ l_{21} & l_{22} & u_{23} & u_{24} \\ l_{31} & l_{32} & l_{33} & u_{34} \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix}$$

2. La factorización triangular es útil para resolver un sistema lineal y sobretodo es conveniente cuando se tiene varios sistemas lineales con diferentes vectores \mathbf{b} .

Conociendo las matrices L y U tales que $A = LU$, se puede presentar al sistema $A\mathbf{x} = \mathbf{b}$ como el sistema $LU\mathbf{x} = \mathbf{b}$.

Si definimos $U\mathbf{x} = \mathbf{y}$, podemos con facilidad resolver el sistema $L\mathbf{y} = \mathbf{b}$, usando el Algoritmo de Sustitución hacia delante; puesto que la matriz L es triangular inferior. Luego, conociendo el vector \mathbf{y} , se puede resolver el sistema $U\mathbf{x} = \mathbf{y}$, usando el Algoritmo de Sustitución hacia atrás, ya que la matriz U es triangular superior.

Ejemplo 3:

Apliquemos la factorización triangular de Crout, para resolver el sistema lineal:

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

a. Resolviendo el sistema lineal $Ly = b$, esto es:

$$\begin{bmatrix} 4 & 0 & 0 \\ -1 & 3 & 0 \\ 1 & -1 & 5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 5 \end{bmatrix}$$

$$\Rightarrow 4y_1 = 6$$

$$-y_1 + 3y_2 = 3$$

$$y_1 - y_2 + 5y_3 = 5$$

$$\therefore y_1 = 3/2; y_2 = 3/2; y_3 = 1.$$

b. Resolviendo el sistema lineal $Ux = y$:

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

$$\Rightarrow x_1 + (1/2)x_3 = 3/2$$

$$x_2 + (1/2)x_3 = 3/2$$

$$x_3 = 1$$

$$\therefore x_3 = 1; x_2 = 1; x_1 = 1.$$

Luego se ha realizado en la parte (a), un Proceso de Sustitución Hacia delante, y después en la parte (b) se realizó un Proceso de Sustitución reversiva, obteniéndose la solución del sistema original $x = [1, 1, 1]^t$.

3. La factorización triangular $A = LU$, también es útil para determinar la inversa de A .

Como:

$$A^{-1} = U^{-1}L^{-1} \quad \text{y} \quad AA^{-1} = I$$

entonces tomando, $A^{-1} = X$ se tiene que:

$$LUX = L(UX) = I$$

Y si definimos $UX = Y \Rightarrow LY = I$, lo que permite hallar la matriz Y , para luego determinar X a partir de la ecuación $UX = Y$, la matriz inversa de A .

Ejemplo 4:

Determinar la inversa de A , usando la factorización matricial.

Sea

$$A = \begin{bmatrix} -3 & 6 & -11 \\ 3 & -4 & 6 \\ 4 & -6 & 13 \end{bmatrix}$$

Se puede ver que $A = LU$ donde

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -4/3 & 1 & 1 \end{bmatrix}; \quad U = \begin{bmatrix} -3 & 6 & -11 \\ 0 & 2 & -5 \\ 0 & 0 & 10/3 \end{bmatrix}$$

Sea

$$X = A^{-1} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$$

Como

$$I = A A^{-1} \quad \text{y} \quad A = LU$$

se tiene que

$$(LU)X = I \quad \text{ó} \quad L(UX) = I$$

Si tomamos

$$Y = UX \quad \text{entonces} \quad LY = I \quad \text{ó}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -4/3 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

De donde, multiplicando la primera fila de L por las columnas de Y se obtiene

$$\begin{aligned} y_{11} &= 1 \\ y_{12} &= 0 \\ y_{13} &= 0 \end{aligned}$$

Luego multiplicando la segunda fila de L por las columnas de Y se tiene

$$\begin{aligned} -y_{11} + y_{21} &= 0 &\Rightarrow & y_{21} = y_{11} = 1 \\ -y_{12} + y_{22} &= 1 &\Rightarrow & y_{22} = 1 + y_{12} = 1 \\ -y_{13} + y_{23} &= 0 &\Rightarrow & y_{23} = y_{13} = 0 \end{aligned}$$

Multiplicando la última fila de L por las columnas de Y se tiene:

$$\begin{aligned} -\frac{4}{3}y_{11} + y_{21} + y_{31} &= 0 &\Rightarrow & y_{31} = \frac{1}{3} \\ -\frac{4}{3}y_{12} + y_{22} + y_{32} &= 0 &\Rightarrow & y_{32} = -1 \\ -\frac{4}{3}y_{13} + y_{23} + y_{33} &= 1 &\Rightarrow & y_{33} = 1 \end{aligned}$$

De donde la matriz Y es de la forma

$$Y = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1/3 & -1 & 1 \end{bmatrix}$$

Resolviendo $U X = Y$:

$$\begin{bmatrix} -3 & 6 & -11 \\ 0 & 2 & -5 \\ 0 & 0 & 10/3 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1/3 & -1 & 1 \end{bmatrix}$$

Nos conviene comenzar con determinar los elementos de la 3^{ra} fila de X . Multiplicando la 3^{ra} fila de U por las columnas de X se obtiene

$$\begin{aligned} \frac{10}{3} x_{31} &= \frac{1}{3} \Rightarrow x_{31} = \frac{1}{10} \\ \frac{10}{3} x_{32} &= -1 \Rightarrow x_{32} = -\frac{3}{10} \\ \frac{10}{3} x_{33} &= 1 \Rightarrow x_{33} = \frac{3}{10} \end{aligned}$$

Multiplicando la 2^{da} fila de U por las columnas de X se tiene:

$$\begin{aligned} 2x_{21} - 5x_{31} &= 1 \Rightarrow x_{21} = 3/4 \\ 2x_{22} - 5x_{32} &= 1 \Rightarrow x_{22} = -1/4 \\ 2x_{23} - 5x_{33} &= 0 \Rightarrow x_{23} = 3/4 \end{aligned}$$

Multiplicando la 1^{ra}. fila de U por las columnas de X se obtienen las siguientes ecuaciones:

(i)

$$-3x_{11} + 6x_{21} - 11x_{31} = 1$$

de donde

$$x_{11} = 0.8$$

(ii)

$$-3x_{12} + 6x_{22} - 11x_{32} = 0$$

de donde

$$x_{12} = 0.6$$

(iii)

$$-3x_{13} + 6x_{23} - 11x_{33} = 0$$

de donde

$$x_{13} = 0.4$$

Entonces

$$A^{-1} = \begin{bmatrix} 4/5 & 3/5 & 2/5 \\ 3/4 & -1/4 & 3/4 \\ 1/10 & -3/10 & 3/10 \end{bmatrix}$$

Comprobando

$$AA^{-1} = \begin{bmatrix} -3 & 6 & -11 \\ 3 & -4 & 6 \\ 4 & -6 & 13 \end{bmatrix} \begin{bmatrix} 4/5 & 3/5 & 2/5 \\ 3/4 & -1/4 & 3/4 \\ 1/10 & -3/10 & 3/10 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

No hay errores del cálculo, ya que trabajamos con aritmética finita.

Algoritmo de Factorización Directa

Objetivo: Para factorizar una matriz $A = [a_{ij}]_{(n \times n)}$ de orden n , en el producto de la matriz triangular inferior $L = [l_{ij}]_{(n \times n)}$ con la matriz triangular superior $U = [u_{ij}]_{(n \times n)}$; esto es, $A = LU$, donde está dada la diagonal principal de cualquiera de L ó U :

Paso 1: Para $j = 1, 2, \dots, n$ tomar: $u_{1j} = a_{1j} / l_{11}$. (Primer renglón de U).

Para $i = 2, 3, \dots, n$ tomar: $l_{i1} = a_{i1} / u_{11}$. (Primera columna de L)

Paso 2: Hacer para:

$$i = 2, 3, 4, \dots, n$$

$$j = i + 1, \dots, n$$

Los siguientes cálculos:

$$u_{ij} = \frac{1}{l_{ii}} \left[a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj} \right] \quad (i\text{-ésimo renglón de } U).$$

$$l_{ji} = \frac{1}{u_{ii}} \left[a_{ji} - \sum_{k=1}^{i-1} l_{jk} u_{ki} \right] \quad (i\text{-ésima columna de } L).$$

3.4 Estrategias de Pivoteo

El método de Eliminación Gaussiana se basa en la idea de que en cada paso al eliminar la i -ésima incógnita de cada una de las ecuaciones, debe estar presente esta incógnita en la i -ésima fila; es decir, el coeficiente pivotal de esta incógnita, el elemento a_{ii} , es diferente de cero. Sin embargo, en muchas situaciones, esto no es suficiente. En la práctica con frecuencia se presentan sistemas que involucran cantidades de magnitudes muy distintas y algunos coeficientes por lo tanto son mucho más pequeños que otros.

Estrategia de Pivoteo Parcial

Si el coeficiente pivotal es muy pequeño, las operaciones que lo involucren, aumentan considerablemente el error del cálculo debido al error del redondeo cometido en cada operación realizada; como una división entre un número pequeño o una multiplicación por un número grande.

Ejemplo:

Sea un sistema:

$$\begin{aligned} E_1: 0.0006x_1 + 2.655x_2 &= 2.659 \\ E_2: 0.4354x_1 - 1.436x_2 &= 0.8970 \end{aligned}$$

La solución exacta del sistema es: $x_1 = 5.359265681$, $x_2 = 1.000295458$.

Empleando una aritmética de cuatro dígitos y tomando E_1 como ecuación pivotal, ya que $a_{11} \neq 0$, se tiene:

$$m_{21} = a_{21}/a_{11} = 725.7$$

Multiplicando $m_{21} \times E_1$, y restando, $E_2 - m_{21}E_1 \rightarrow E_2$, el sistema toma la forma:

$$\begin{aligned} E_1: 0.0006x_1 + 2.655x_2 &= 2.659 \\ E_2: -1928x_2 &= -1929 \end{aligned}$$

lo que implica que:

$$\begin{aligned} x_2 &= 1.001 \\ x_1 &= 1.667 \end{aligned}$$

y si la respuesta para x_2 es aceptable, para x_1 no, ya que ningún dígito es exacto.

Los errores del cálculo, originados por la división entre un número pequeño, originaron una pérdida total de dígitos exactos.

Para enfrentar este tipo de problemas se puede desarrollar una estrategia que consista en evitar las operaciones que aumenten el error del cálculo, como la división entre un número pequeño.

Una de las estrategias se conoce como estrategia de pivoteo parcial y consiste en elegir por ecuación pivotal a aquella, que tiene por coeficiente pivotal el valor más grande en su valor absoluto. Este proceso se realiza a través de intercambios de filas, sin intercambio de columnas.

Describiendo:

- a. En el 1^{er} paso de Eliminación Gaussiana, escogemos por ecuación pivotal a aquella donde se cumple:

$$|a_{p1}^{(1)}| = \max_{i=1, 2, \dots, n} |a_{i1}|, \text{ y efectuaremos el intercambio de filas } E_p \leftrightarrow E_1,$$

usamos luego E_1 para eliminar la incógnita x_1 de las restantes $(n-1)$ -ecuaciones, obteniendo por debajo de la diagonal ceros, en la 1^{ra} columna de A .

- b. En el 2^{do} paso de eliminación, revisamos los elementos de la 2^{da} columna de A , a partir del segundo elemento:

$$|a_{p2}^{(2)}| = \max_{i=2, 3, \dots, n} |a_{i2}^{(1)}|, \text{ luego efectuaremos el intercambio de filas}$$

$E_p \leftrightarrow E_2$, y se usará E_2 como ecuación pivotal para eliminar la 2^{da} incógnita x_2 , en la 2^{da} columna de A .

En general, en el k -ésimo paso de la Eliminación Gaussiana, se tomará por ecuación pivotal, a aquella donde se logra:

$$a_{pk}^{(k)} = \max_{i=k, \dots, n} |a_{ik}^{(k-1)}|, \text{ para luego efectuar el intercambio de filas } E_p \leftrightarrow E_k.$$

Regresando a nuestro ejemplo:

Podemos observar que,

$$\max_{i=1, 2} |a_{ik}| = \max\{|a_{11}|, |a_{21}|\} = \max\{0.0006, 0.4354\} = a_{21} = 0.4354 \Rightarrow E_2 \leftrightarrow E_1$$

por lo que tomamos a E_2 como ecuación pivotal.

Entonces el sistema toma la forma:

$$\begin{aligned} E_1: 0.4354x_1 - 1.436x_2 &= 0.8970 \\ E_2: 0.0006x_1 + 2.655x_2 &= 2.659 \end{aligned}$$

Y realizando los cálculos:

$$m_{21} = 0.1378 \times 10^{-2}$$

entonces calculando $E_2 - m_{21} E_1$ se obtiene el sistema:

$$E_1: 0.4354x_1 - 1.436x_2 = 0.897$$

$$E_2: 2.657x_2 = 2.658$$

lo que implica que

$$x_2 = 1.000$$

$$x_1 = 5.358$$

y la respuesta tiene tres dígitos exactos.

Algoritmo de Eliminación Gaussiana con Pivoteo Parcial

Objetivo : Para resolver el sistema lineal de $n \times n$ ecuaciones:

$$\begin{array}{rcccccccc} E_1: & a_{11}x_1 & + & a_{12}x_2 & + \dots & + & a_{1n}x_n & = & a_{1(n+1)} \\ E_2: & a_{21}x_1 & + & a_{22}x_2 & + \dots & + & a_{2n}x_n & = & a_{2(n+1)} \\ E_3: & a_{31}x_1 & + & a_{32}x_2 & + \dots & + & a_{3n}x_n & = & a_{3(n+1)} \\ & \vdots & & \vdots & & & \vdots & & \vdots \\ E_n: & a_{n1}x_1 & + & a_{n2}x_2 & + \dots & + & a_{nn}x_n & = & a_{n(n+1)} \end{array}$$

con los pivotes $a_{ii} \neq 0$, para $i = 1, 2, 3, \dots, n$.

Hacer para $j = 1, 2, \dots, n-1$ los pasos 1 y 2.

Paso 1: Para $i = j, \dots, n$ tomar p el menor número entero con $i \leq p \leq n$, tal que

$$|a(p, j)| = \max_i |a(i, j)|,$$

Luego intercambiar las filas $E_p \leftrightarrow E_i$.

Paso 2: Aplicar el proceso de Eliminación Gaussiana.

Estrategia de Pivoteo Escalonado o Escalado

Si el sistema tiene coeficientes de distinto tamaño, se puede diseñar otro tipo de estrategia, conocida como estrategia de pivoteo escalonado o escalado de columna. Tal estrategia consiste en uniformizar los coeficientes a través del siguiente procedimiento:

En 1^{er} lugar podemos establecer una “medida” o “tamaño” de la i -ésima fila como:

$$S_i = \max_{j=1, \dots, n} |a_{ij}|.$$

Al dividir luego cada fila por su medida, todos los coeficientes serán no mayores que 1. Tal división se realiza sólo con el propósito de determinar la ecuación pivotal, no así para la aplicación de la Eliminación Gaussiana.

Luego la matriz con los elementos uniformizados, es:

$$\begin{bmatrix} a_{11} / S_1 & a_{12} / S_1 & \cdots & a_{1n} / S_1 \\ a_{21} / S_2 & a_{22} / S_2 & \cdots & a_{2n} / S_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} / S_n & a_{n2} / S_n & \cdots & a_{nn} / S_n \end{bmatrix}$$

Esta matriz sólo se utilizará para determinar la ecuación pivotal E_p , en cada paso de la eliminación. Una vez determinada E_p , la Eliminación Gaussiana se realiza con los elementos de la matriz A , para no introducir los errores que originan las divisiones de los elementos de A entre S_i .

Entonces el procedimiento

consiste en determinar $\max_{j=k, \dots, n} \frac{|a_{jk}|}{S_j}$, para cada k -ésimo paso de Eliminación Gaussiana.

Paso 1: si $k=1$, se toma por la ecuación pivotal E_p , a aquella donde se logra:

$$\max_{j=1, \dots, n} \frac{|a_{j1}|}{S_j} = \max \left\{ \frac{|a_{11}|}{S_1}, \frac{|a_{21}|}{S_2}, \dots, \frac{|a_{n1}|}{S_n} \right\} = \frac{|a_{p1}|}{S_p}$$

Si este máximo se logra para $j=p$, se toma $E_p \leftrightarrow E_1$ y se procede a trabajar con los elementos de la matriz A , con el fin de obtener ceros en la 1^{ra} columna de A , bajo la diagonal.

Paso 2: Si $k=2$, se comparan los elementos $[a_{j2}/S_j]$, para $j=2, 3, \dots, n$; y se toma por E_p a aquella fila que satisface:

$$\max_{j=2, 3, \dots, n} \left\{ \frac{|a_{j2}|}{S_j} \right\} = \max \left\{ \frac{|a_{22}|}{S_2}, \frac{|a_{32}|}{S_3}, \dots, \frac{|a_{n2}|}{S_n} \right\} = \frac{|a_{p2}|}{S_p},$$

entonces se intercambian las filas, \dots

$$E_p \leftrightarrow E_2$$

y así sucesivamente para $k=3, 4, \dots, n$.

Veamos un ejemplo:

$$E_1: 0.0004x_1 + 1.402x_2 = 1.406$$

$$E_2: 0.4003x_1 - 1.502x_2 = 2.501$$

Empleando una aritmética de cuatro dígitos, tenemos:

$$S_1 = \max \{0.0004, 1.402\} = 1.402$$

$$S_2 = \max \{0.4003, 1.502\} = 1.502,$$

Paso 1: Para $i=1, 2, \dots, n$; tomar $S_i = \max_{1 \leq j \leq n} |a_{ij}|$

Hacer para $j = 1, \dots, n-1$ los pasos 2 y 3.

Paso 2: Para $i = j, j+1, \dots, n$; tomar p el menor entero con $i \leq p \leq n$, tal que satisfaga:

$$\frac{|a(p, j)|}{a(p, j)} = \max_j \left\{ \frac{|a(i, j)|}{S_i} \right\},$$

entonces se intercambian las filas $E_p \leftrightarrow E_i$.

Paso 3: Realizar el proceso de Eliminación Gaussiana.

3.5 Métodos Iterativos para la Solución de Sistemas Lineales

Los métodos directos son aptos para las matrices densas (con pocos elementos nulos) y de orden "pequeño".

Si la matriz del sistema es de orden grande, el número de operaciones a realizarse crece de tal forma que el error del cálculo afecta a la solución aproximada.

En muchas aplicaciones, se presentan modelos matemáticos como ecuaciones diferenciales ordinarias o con derivadas parciales. Al aproximar tales ecuaciones por diferencias finitas se obtienen sistemas lineales con la matriz de orden grande, pero con muchas entradas nulas. Para este tipo de sistemas se utilizan los métodos iterativos.

A diferencia de los métodos directos, los métodos iterativos se basan en el concepto de límite de una sucesión de vectores en \mathbb{R}^n .

Sea $Ax = b$, con x la solución exacta.

El método iterativo consiste en construir una sucesión vectorial $\{x^{(k)}\}_{(k \geq 0)}$ que converge a x , es decir:

$$\lim_{k \rightarrow \infty} x^{(k)} = x$$

y por lo tanto para obtener una aproximación

$$\mathbf{x}^{(k)} \approx \mathbf{x}$$

hay que truncar el proceso infinito de límite, lo que significa que los métodos iterativos están sujetos a dos tipos de errores: al error de truncamiento, y al error de cálculo.

La aproximación $\mathbf{x}^{(k)} \approx \mathbf{x}$ para algún k suficientemente grande, se logra a través de sus componentes

$$x_i^{(k)} \approx x_i \text{ para } i = 1, \dots, n$$

con una medida que utiliza el concepto de la norma vectorial.

Normas de vectores y de matrices

Definición:

Sea $\mathbf{x} = (x_1, x_2, \dots, x_n)$ un vector en \mathfrak{R}^n , entonces se puede definir la norma de este vector, de las siguientes formas:

$$(1) \quad \|\mathbf{x}\|_u = \|\mathbf{x}\|_\infty = \max_i \{ |x_i| \}, \quad (\text{norma uniforme})$$

$$(2) \quad \|\mathbf{x}\|_s = \|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|, \quad (\text{norma suma})$$

$$(3) \quad \|\mathbf{x}\|_e = \|\mathbf{x}\|_2 = \left\{ \sum_{i=1}^n x_i^2 \right\}^{1/2}, \quad (\text{norma euclidiana})$$

y cada una de estas normas satisface las siguientes propiedades:

- (i) $\|\mathbf{x}\| \geq 0, \forall \mathbf{x} \in \mathfrak{R}^n,$
- (ii) $\|\mathbf{x}\| = 0, \Leftrightarrow \mathbf{x} = \mathbf{0}$
- (iii) $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|, \alpha \text{ es escalar.}$
- (iv) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \forall \mathbf{x}, \mathbf{y} \in \mathfrak{R}^n.$
- (v) $\|-\mathbf{x}\| = \|\mathbf{x}\|, \text{ o equivalente } \forall \mathbf{x}, \mathbf{y} \in \mathfrak{R}^n \text{ se tiene } \|\mathbf{x} - \mathbf{y}\| = \|\mathbf{y} - \mathbf{x}\|$
- (vi) $\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

Definición:

Si $\mathbf{x} \in \mathfrak{R}^n$ y es tal que $\|\mathbf{x}\| = 1$, entonces se dice que es un vector unitario.

Definición:

Sea \mathbf{x} un vector en \mathfrak{R}^n , se define al vector normalizado de \mathbf{x} , como al vector $\frac{\mathbf{x}}{\|\mathbf{x}\|}$ en \mathfrak{R}^n , es decir al vector cuya norma es 1.

Definición:

Sean $\mathbf{x}, \mathbf{y} \in \mathfrak{R}^n$, se define la distancia entre \mathbf{x} e \mathbf{y} , como:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$

Definición:

Se dice que una sucesión de vectores $\{\mathbf{x}^{(k)}\}_{(k \geq 0)}$ en \mathfrak{R}^n , donde

$$\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}),$$

converge a \mathbf{x} , con respecto a cualesquiera de las normas, si dado cualquier $\varepsilon > 0$, existe un entero $N(\varepsilon)$ tal que:

$$\|\mathbf{x}^{(k)} - \mathbf{x}\| < \varepsilon, \forall k \geq N(\varepsilon)$$

o lo que es equivalente a decir: $\lim_{k \rightarrow \infty} \|\mathbf{x}^{(k)} - \mathbf{x}\| = 0$.

El concepto de la norma de un vector se extiende a una matriz, como norma matricial.

Definición:

Sea $A = [a_{ij}]_{n \times n}$ una matriz de orden n se puede definir la norma de A , como

$$\|A\| = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}, \mathbf{x} \in \mathfrak{R}^n,$$

y se puede probar que esta definición es equivalente a las dos condiciones siguientes:

- (1) $\forall x \in \mathfrak{R}^n, \|Ax\| \leq \|A\| \|x\|$
 (2) $\exists x \neq 0, x \in \mathfrak{R}^n / \|x\| = 1$ y $\|Ax\| = \|A\| \|x\|$

Demostración:

Como

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

Se tiene:

$$\begin{aligned} \frac{\|Ax\|}{\|x\|} &\leq \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \|A\| \\ \Rightarrow \|Ax\| &\leq \|A\| \|x\| \end{aligned}$$

lo que prueba la condición (1).

Para probar la condición (2) tomemos:

$$x \neq 0, x \in \mathfrak{R}^n / \|x\| = 1$$

entonces.

$$\begin{aligned} \|Ax\| &\leq \|A\| \|x\| = \|A\| = \\ &= \max_{\|x\|=1} \|Ax\| \geq \|Ax\| \\ \Rightarrow \|Ax\| &\leq \|A\| \geq \|Ax\| \Rightarrow \|Ax\| = \|A\| = \|A\| \|x\| \square \end{aligned}$$

Propiedades de la norma matricial

De la definición de norma de matriz se deducen las siguientes propiedades:

- (i) $\|A\| \geq 0$, \forall matriz $A_{n \times n}$ y $\|A\| = 0 \Leftrightarrow A = O_{n \times n}$
- (ii) $\|\alpha A\| = |\alpha| \|A\|$, $\alpha \in \mathfrak{R}^n$
- (iii) Para dos matrices cualesquiera A y B
- $$\|A + B\| \leq \|A\| + \|B\|$$
- (iv) $\|AB\| \leq \|A\| \|B\|$

Como la norma matricial se define sobre la base de la norma vectorial, ésta podría darse para cualesquiera de ellas: uniforme, suma, euclidiana:

- (1) $\|A\|_u = \max_i \left\{ \sum_{j=1}^n |a_{ij}| \right\}$, correspondiente a la norma vectorial uniforme $\|x\|_u$.
- (2) $\|A\|_s = \max_j \left\{ \sum_{i=1}^n |a_{ij}| \right\}$, correspondiente a la norma vectorial suma $\|x\|_s$.
- (3) $\|A\|_e = \max_{x \neq 0} \frac{\|Ax\|_e}{\|x\|_e}$, correspondiente a la norma vectorial euclidiana $\|x\|_e$.

Demostración:

Probaremos la afirmación para el caso (1).

Según el resultado anterior, para probar que

$$\|x\|_u = \max_i \{ |x_i| \}$$

da lugar a la norma matricial:

$$\|A\|_u = \max_i \left\{ \sum_{j=1}^n |a_{ij}| \right\}$$

debemos probar que:

- a. $\forall x \in \mathfrak{R}^n$, $\|Ax\|_u \leq \|A\|_u \|x\|_u$.
- b. $\exists x \neq 0$, $x \in \mathfrak{R}^n$, $\|x\|_u = 1$ tal que $\|Ax\|_u = \|A\|_u \|x\|_u$.
- a. Para todos los vectores $x \in \mathfrak{R}^n$, $x \neq 0$ se tiene:

$$\begin{aligned}\|Ax\|_u &= \max_i \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \max_i \sum_{j=1}^n |a_{ij}| |x_j| \leq \max_i \left[\max_j |x_j| \sum_{j=1}^n |a_{ij}| \right] = \\ &= \max_i \sum_{j=1}^n |a_{ij}| \cdot \|x\|_u = \|A\|_u \|x\|_u.\end{aligned}$$

b. Sea $\max_i \sum_{j=1}^n |a_{ij}|$ se logra para algún $i = k$. Tomaremos el vector x ,

con las componentes $x_j = \frac{|a_{kj}|}{|a_{kj}|}$, si $a_{kj} \neq 0$ y $x_j = 1$ si $a_{kj} = 0$, entonces

$$\begin{aligned}\|x\|_u &= 1 \text{ y } \|Ax\|_u = \max_i \left| \sum_{j=1}^n a_{ij} x_j \right| = \left| \sum_{j=1}^n a_{kj} x_j \right| = \sum_{j=1}^n |a_{kj}| = \max_i \sum_{j=1}^n |a_{ij}| = \\ &= \|A\|_u = \|A\|_u \|x\|_u \quad \square\end{aligned}$$

Métodos Iterativos

Una técnica iterativa para resolver un sistema lineal,

$$Ax = b$$

de n ecuaciones consiste en tomar una aproximación inicial de x , $x^{(0)}$, a partir de la cual se genera una sucesión de vectores $\{x^{(k)}\}_{(k \geq 0)}$ que converge a x .

La mayoría de estas técnicas iterativas involucran un proceso que convierte al sistema $Ax = b$ en un sistema equivalente de la forma:

$$x = Bx + c$$

para alguna matriz $B_{n \times n}$ y un vector c . Ya seleccionado el vector inicial, $x^{(0)}$, la sucesión de vectores que se aproxima a la solución buscada, se genera calculando:

$$x^{(k+1)} = B x^{(k)} + c, \quad k = 0, 1, 2, 3, \dots$$

Se espera que la sucesión resultante converja a x , esto es

$$B = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \cdots & -\frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{nn}} & \cdots & -\frac{a_{n,n-1}}{a_{nn}} & 0 \end{bmatrix}; \quad c = \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{bmatrix}$$

y se define el algoritmo de Jacobi, como:

$$\text{Algoritmo: } \mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c} \quad \forall k \geq 0.$$

con $\mathbf{x}^{(0)}$ una aproximación inicial.

Comentario:

Para detener el algoritmo, dada una tolerancia especificada ε se emplean cualesquiera de los siguientes criterios:

$$(i) \quad \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_u < \varepsilon$$

$$(ii) \quad \frac{\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_u}{\|\mathbf{x}^{(k+1)}\|_u} < \varepsilon$$

Ejemplo 1:

Resolver el sistema lineal:

$$\begin{bmatrix} 3 & 1 & 0 \\ -1 & 3 & 1 \\ 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \text{con } \mathbf{x}^{(0)} = \mathbf{0}.$$

Solución:

$$B = \begin{bmatrix} 0 & -1/3 & 0 \\ 1/3 & 0 & -1/3 \\ 0 & -2/3 & 0 \end{bmatrix}; \quad \mathbf{c} = \begin{bmatrix} 1/3 \\ -1/3 \\ 0 \end{bmatrix}$$

entonces

(i) Para $k = 0$: $\mathbf{x}^{(1)} = B\mathbf{x}^{(0)} + \mathbf{c}$

$$\mathbf{x}^{(1)} = \mathbf{c} = \begin{bmatrix} 1/3 \\ -1/3 \\ 0 \end{bmatrix}$$

(ii) Para $k = 1$:

$$\mathbf{x}^{(2)} = B\mathbf{c} + \mathbf{c} = \begin{bmatrix} 0.4 \\ -0.2 \\ 0.2 \end{bmatrix};$$

Las iteraciones se presentan en la siguiente tabla:

i	$x_1^{(i)}$	$x_2^{(i)}$	$x_3^{(i)}$
2	0.44444444	-0.22222222	0.22222222
3	0.40740740	-0.25925925	0.14814814
4	0.41975308	-0.24691358	0.17283950
5	0.41563786	-0.25102806	0.16460905
6	0.41700960	-0.24965706	0.16735253
7	0.41655235	-0.25011431	0.16643804
8	0.41670477	-0.24996189	0.16674287
9	0.41665396	-0.25001270	0.16664126
10	0.4166709	-0.24999576	0.16667513
11	0.41666525	-0.25000141	0.16666384
12	0.41666713	-0.24999529	0.16666707
13	0.41666509	-0.25000156	0.16666353
14	0.41666718	-0.24999947	0.16666771
15	0.41666649	-0.25000017	0.16666631
16	0.41666672	-0.24999994	0.16666678
17	0.41666664	-0.25000001	0.16666662
18	0.41666667	-0.24999999	0.16666666
19	0.41666666	-0.25	0.16666666
20	0.41666666	-0.25	0.16666666
21	0.41666666	-0.24999999	0.16666666
22	0.41666666	-0.25	0.16666666

Tabla 1

notamos que la sucesión de vectores que se ha generado converge a la solución del sistema lineal dado:

$$\mathbf{x} = \begin{bmatrix} 0.41\widehat{6} \\ -0.25 \\ 0.1\widehat{6} \end{bmatrix}$$

De la tabla, se puede observar que en las últimas líneas las cantidades obtenidas van a coincidir para cada una de las componentes.

Ejemplo 2:

Resolver el sistema lineal:

$$\begin{bmatrix} -8 & 3 & 1 \\ 2 & -6 & -1 \\ 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -4 \\ -3 \\ 5 \end{bmatrix}, \quad \text{con } \mathbf{x}^{(0)} = \mathbf{0}.$$

empleando el método de Jacobi.

Solución:

$$B = \begin{bmatrix} 0 & 3/8 & 1/8 \\ 2/6 & 0 & -1/6 \\ 1/3 & 1/3 & 0 \end{bmatrix}; \quad \mathbf{c} = \begin{bmatrix} 1/2 \\ 1/2 \\ -5/3 \end{bmatrix}$$

entonces

(i) Para $k = 0$: $\mathbf{x}^{(1)} = B \mathbf{x}^{(0)} + \mathbf{c}$

$$\mathbf{x}^{(1)} = \mathbf{c} = \begin{bmatrix} 1/2 \\ 1/2 \\ -5/3 \end{bmatrix}$$

(ii) Para $k = 1$:

$$\mathbf{x}^{(2)} = B\mathbf{c} + \mathbf{c} = \begin{bmatrix} 0.4791\widehat{6} \\ 0.9\widehat{4} \\ -1.\widehat{3} \end{bmatrix};$$

Las siguientes iteraciones se presentan a continuación:

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	0	0	0
1	0.5	0.5	-1.6666666
2	0.47916670	0.9444444	-1.3333333
3	0.6875	0.8819444	-1.19212963
4	0.681712963	0.9278549383	-1.143518518
5	0.7050057871	0.9178240741	-1.130144033
10	0.7053948703	0.9224989129	-1.124043407
15	0.7054264722	0.9224804846	-1.124031045
20	0.7054263557	0.922480621	-1.124031008
23	0.7054263566	0.9224806202	-1.124031008
24	0.7054263566	0.9224806202	-1.124031008

Tabla 2

Observamos que converge hacia la solución aproximada del sistema lineal dado:

$$\mathbf{x} = \begin{bmatrix} 0.7054263566 \\ 0.9224806202 \\ -1.124031008 \end{bmatrix}$$

Comentarios:

- Este método consiste en resolver la i -ésima ecuación de $A\mathbf{x} = \mathbf{b}$, para x_i (siempre que $a_{ii} \neq 0$), si este no es el caso las ecuaciones pueden ser reordenadas de tal manera que a_{ii} sea lo más grande posible para acelerar la convergencia y obtener que:

$$x_i = \sum_{\substack{j=1 \\ j \neq i}}^n \left(-\frac{a_{ij} x_j}{a_{ii}} \right) + \frac{b_i}{a_{ii}}; \quad \forall i = 1, 2, \dots, n.$$

Para luego generar cada $x_i^{(k+1)}$ de las componentes $\mathbf{x}^{(k)}$, con:

$$x_i^{(k+1)} = \frac{\left(\sum_{\substack{j=1 \\ j \neq i}}^n -a_{ij} x_j^{(k)} + b_i \right)}{a_{ii}}; \quad \text{para } i = 1, 2, \dots, n.$$

Hacer para $k = 1, 2, \dots$, hasta donde se satisfaga la tolerancia especificada ε , los siguientes cálculos:

Para $i = 1, 2, \dots, n$

$$x_i^{(k+1)} = \frac{- \sum_{\substack{j=1 \\ j \neq i}}^n (a_{ij} x_j^{(k)}) + b_i}{a_{ii}}$$

Método de Gauss-Seidel

Analicemos la ecuación del Algoritmo de Jacobi:

$$x_i^{(k+1)} = \frac{\sum_{\substack{j=1 \\ j \neq i}}^n (-a_{ij} x_j^{(k)}) + b_i}{a_{ii}} \quad \text{con } a_{ii} \neq 0.$$

Se ve que para hallar la i -ésima componente del vector $\mathbf{x}^{(k+1)}$, se requiere de las componentes del vector k -ésimo, $x_i^{(k)}$.

Supongamos que para $i > 1$, se han calculado las $(i-1)$ -primeras componentes del vector $(k+1)$ -ésimo:

$$x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)},$$

y éstas serán mejores aproximaciones a las componentes correspondientes de la solución exacta \mathbf{x} :

$$x_1, x_2, \dots, x_{i-1}$$

que las $(i-1)$ -primeras componentes del último vector $\mathbf{x}^{(k)}$:

$$x_1^{(k)}, x_2^{(k)}, \dots, x_{i-1}^{(k)}.$$

Entonces sería razonable calcular $x_i^{(k+1)}$ usando los últimos valores obtenidos; es decir:

$$x_i^{(k+1)} = \frac{\sum_{j=1}^{i-1} (-a_{ij}x_j^{(k+1)}) - \sum_{j=i+1}^n a_{ij}x_j^{(k)} + b_i}{a_{ii}} \dots \quad (1)$$

con $a_{ii} \neq 0$ para todo $i = 1, 2, \dots, n$.

Y este método recibe el nombre de Gauss-Seidel.

Desarrollando la última ecuación:

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

y dejando en un miembro a las componentes del $(k + 1)$ -ésimo vector y en el otro a las componentes del k -ésimo vector, se obtendrá:

$$\sum_{j=1}^i a_{ij}x_j^{(k+1)} = -\sum_{j=i+1}^n a_{ij}x_j^{(k)} + b_i, \text{ para todo } i \geq 1$$

Desarrollando estas ecuaciones para $i = 1, 2, 3, \dots, n$:

$$\begin{aligned} i=1, a_{11}x_1^{(k+1)} &= -a_{12}x_2^{(k)} - a_{13}x_3^{(k)} \dots - a_{1n}x_n^{(k)} + b_1 \\ i=2, a_{21}x_1^{(k+1)} + a_{22}x_2^{(k+1)} &= -a_{23}x_3^{(k)} - a_{24}x_4^{(k)} \dots - a_{2n}x_n^{(k)} + b_2 \\ \dots & \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ i=n, a_{n1}x_1^{(k+1)} + a_{n2}x_2^{(k+1)} \dots + a_{nn}x_n^{(k+1)} &= b_n \end{aligned}$$

y escribiéndolas matricialmente, ya que $A = L + D + U$:

$$(L+D)x^{(k+1)} = -Ux^{(k)} + b, \text{ entonces } x^{(k+1)} = -(L+D)^{-1}Ux^{(k)} + (L+D)^{-1}b$$

por lo tanto:

$$B = -(L+D)^{-1} U, \quad c = (L+D)^{-1} b$$

Lo que lleva al método de Gauss-Seidel en forma matricial:

$$x^{(k+1)} = Bx^{(k)} + c, \quad \forall k \geq 0 \dots (2)$$

con $x^{(0)}$ aproximación inicial.

Algoritmo iterativo de Gauss-Seidel

Objetivo: Para resolver $Ax = b$, con la matriz A con pivotes $a_{ii} \neq 0$, para una tolerancia especificada ε . Con una aproximación inicial dada $x^{(0)}$.

Hacer para $k = 1, 2, \dots$, hasta donde se satisfaga la tolerancia especificada ε , los siguientes cálculos:

Para $i = 1, 2, \dots, n$

$$x_i^{(k+1)} = \frac{-\sum_{j=1}^{i-1} (a_{ij} x_j^{(k+1)}) - \sum_{j=i+1}^n a_{ij} x_j^{(k)} + b_i}{a_{ii}}.$$

Ejemplo 3:

Resolver el sistema lineal:

$$\begin{bmatrix} 3 & 1 & 0 \\ -1 & 3 & 1 \\ 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \text{ con } x^{(0)} = \mathbf{0}.$$

Empleando el método de Gauss-Seidel en forma matricial.

Solución:

Se calcula:

$$B = -(L+D)^{-1} U; \quad c = (L+D)^{-1} b$$

$$B = -\begin{bmatrix} 3 & 0 & 0 \\ -1 & 3 & 0 \\ 0 & 2 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = -\begin{bmatrix} 0 & 1/3 & 0 \\ 0 & 1/9 & 1/3 \\ 0 & -2/27 & -2/9 \end{bmatrix}; \quad c = \begin{bmatrix} 1/3 \\ -2/9 \\ 4/27 \end{bmatrix}$$

Como:

$$\mathbf{x}^{(k+1)} = B \mathbf{x}^{(k)} + \mathbf{c}; \quad k \geq 0.$$

las iteraciones se resumen en el siguiente cuadro:

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	0	0	0
1	1/3	-2/9	4/27
2	11/27	-20/81	40/243
3	101/243	-182/729	364/2187
4	911/2187	-1640/6561	3280/19683
5	8201/19683	-0.2499957662	0.1666638442
6	0.4166652554	-0.2499995296	0.1666663531
7	0.4166665099	-0.2499999477	0.1666666318
8	0.4166666492	-0.2499999942	0.1666666628
9	0.4166666647	-0.2499999994	0.1666666662
10	0.4166666665	-0.2499999999	0.1666666666
11	0.4166666666	-0.25	0.1666666667
12	0.4166666666	-0.25	0.1666666667

Tabla 3

Observar que en la aplicación de Jacobi a este sistema lineal fue necesario realizar un número mayor de iteraciones que en la aplicación de este método.

Al aplicarse este método (2) de Gauss-Seidel empleando el cálculo de la inversa de $(L+D)^{-1}$, se puede introducir errores del cálculo; sobretodo si se trata de sistemas lineales con órdenes mayores, por lo que es conveniente trabajar en estos casos con el método (1). Veamos el siguiente ejemplo:

Ejemplo 4:

Resolver el sistema lineal:

$$\begin{bmatrix} -8 & 3 & 1 \\ 2 & -6 & -1 \\ 1 & 3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -4 \\ -3 \\ 5 \end{bmatrix}, \quad \text{con } \mathbf{x}^{(0)} = \mathbf{0}.$$

empleando el método de Gauss-Seidel.

Solución:

Trabajaremos con el método (1); es decir:

(i) Para $k = 0$, se tiene:

$$\begin{aligned} a_{11}x_1^{(1)} &= 0 + b_1 = -4 \\ a_{21}x_1^{(1)} + a_{22}x_2^{(1)} &= 0 + b_2 = -3 \\ a_{31}x_1^{(1)} + a_{32}x_2^{(1)} + a_{33}x_3^{(1)} &= 0 + b_3 = 5 \end{aligned}$$

por lo tanto, obtenemos:

$$\mathbf{x}^{(1)} = \begin{bmatrix} 1/2 \\ 2/3 \\ -23/18 \end{bmatrix}$$

(ii) Para $k = 1$, se tiene:

$$\begin{aligned} a_{11}x_1^{(2)} &= -a_{12}x_2^{(1)} - a_{13}x_3^{(1)} + b_1 \\ a_{21}x_1^{(2)} + a_{22}x_2^{(2)} &= -a_{23}x_3^{(1)} + b_2 \\ a_{31}x_1^{(2)} + a_{32}x_2^{(2)} + a_{33}x_3^{(2)} &= 0 + b_3 \end{aligned}$$

por lo tanto, obtenemos:

$$\mathbf{x}^{(2)} = \begin{bmatrix} 0.5902\widehat{7} \\ 0.9097\widehat{2} \\ -1.1\widehat{6} \end{bmatrix}$$

Resumimos las siguientes iteraciones en el cuadro a continuación:

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
3	0.695313	0.926215	-1.126157
4	0.7065610532	0.9232132523	-1.123409
5	0.705778899	0.9224943938	-1.123908902
10	0.7054263744	0.9224806251	-1.124031
15	0.7054263566	0.9224806202	-1.124031008
16	0.7054263566	0.9224806202	-1.124031008

Tabla 4

De la tabla se puede apreciar que el método de Gauss-Seidel converge hacia la solución aproximada del sistema lineal dado, en un

número menor de iteraciones, con relación al resultado obtenido al aplicar el método de Jacobi a este mismo sistema lineal (**Ejemplo 2**).

Comentario:

En la mayoría de los casos el método de Gauss-Seidel converge más rápido que el método de Jacobi, sin embargo hay sistemas lineales para los cuales el método de Jacobi converge y el método de Gauss-Seidel no y otros para los cuales el método de Gauss-Seidel converge y el de Jacobi no.

3.6 Convergencia de los Métodos Iterativos

Estudiaremos la convergencia de las técnicas generales de iteración:

$$\mathbf{x}^{(k+1)} = B \mathbf{x}^{(k)} + \mathbf{c}, \quad k \geq 0.$$

Para lo cual necesitamos revisar algunos conceptos teóricos:

Valores propios de una matriz

Definición:

Se dice que λ real o complejo es el valor propio o característico de una matriz A , si para algún $\mathbf{x} \neq \mathbf{0}$, se tiene $A\mathbf{x} = \lambda\mathbf{x}$. El vector $\mathbf{x} \neq \mathbf{0}$ se conoce como vector propio o característico de la matriz A , correspondiente al valor propio λ .

De esta definición se puede obtener lo siguiente: como $\mathbf{x} \neq \mathbf{0}$ es solución del sistema $[A - \lambda I]\mathbf{x} = \mathbf{0}$, entonces $\det[A - \lambda I] = 0$.

Pero,

$$\det[A - \lambda I] = \det \begin{bmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{bmatrix} = (-\lambda)^n + \dots = P_n(\lambda)$$

es un polinomio de grado n , llamado polinomio característico de la matriz A .

Entonces los valores propios de una matriz son ceros del polinomio característico:

$$P_n(\lambda) = 0.$$

Ejemplo:

$$A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & -3 & 0 \\ 1 & 0 & 3 \end{bmatrix},$$

$$\begin{aligned} P_3(\lambda) &= \det [A - \lambda I] = \det \begin{bmatrix} 3-\lambda & 0 & 1 \\ 0 & -3-\lambda & 0 \\ 1 & 0 & 3-\lambda \end{bmatrix} = \\ &= (3-\lambda) \begin{vmatrix} -3-\lambda & 0 \\ 0 & 3-\lambda \end{vmatrix} + \begin{vmatrix} 0 & -3-\lambda \\ 1 & 0 \end{vmatrix} = \\ &= (3-\lambda)(-3-\lambda)(3-\lambda) - (-3-\lambda) = (-3-\lambda)(\lambda^2 - 6\lambda + 8) = 0 \\ &\Rightarrow \lambda_1 = 4; \quad \lambda_2 = 2; \quad \lambda_3 = -3 \end{aligned}$$

Todas las raíces son reales.

Hallando los vectores propios u_1, u_2, u_3 ; correspondientes a los valores propios hallados:

(i) Para $\lambda_1 = 4$, se tiene que:

$$[A - \lambda_1 I]u_1 = 0$$

o lo que es lo mismo:

$$\begin{bmatrix} -1 & 0 & 1 \\ 0 & -7 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow -u_1 + u_3 = 0$$

$$-7u_2 = 0$$

$$u_1 - u_3 = 0$$

$$\therefore u_2 = 0 \quad u_3 = u_1$$

luego el vector propio correspondiente al valor propio λ_1 es:

$$\mathbf{u}_1 = \begin{bmatrix} u_1 \\ 0 \\ u_1 \end{bmatrix}, \text{ para } u_1 \in \mathfrak{R}$$

(ii) Para $\lambda_2 = 2$, se tiene que:

$$[A - \lambda_2 I] \mathbf{u}_2 = \mathbf{0}$$

o lo que es lo mismo:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & -5 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow u_1 + u_3 = 0$$

$$-5u_2 = 0$$

$$u_1 + u_3 = 0$$

$$\therefore u_2 = 0 \quad u_3 = -u_1$$

luego el vector propio correspondiente al valor propio λ_2 es:

$$\mathbf{u}_2 = \begin{bmatrix} u_1 \\ 0 \\ -u_1 \end{bmatrix}, \text{ para } u_1 \in \mathfrak{R}$$

(iii) Para $\lambda_3 = -3$:

$$[A - \lambda_3 I] u_3 = \mathbf{0}$$

$$\begin{bmatrix} 6 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 6 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} \Rightarrow \quad 6u_1 + u_3 &= 0 \\ u_1 + 6u_3 &= 0 \\ \therefore u_3 = u_1 = 0, u_2 &\in \mathfrak{R} \end{aligned}$$

luego el vector propio es: $u_3 = \begin{bmatrix} 0 \\ u_2 \\ 0 \end{bmatrix}$, para $u_2 \in \mathfrak{R}$.

Definición:

El radio espectral o valor propio dominante $\rho(A)$ de una matriz A se define como:

$\rho(A) = \max_i |\lambda_i|$ donde los $\{\lambda_i\}_{(1 \leq i \leq n)}$ son los valores característicos de A .
(recordar que si $\lambda = a + bi$, es un número complejo entonces $|\lambda| = (a^2 + b^2)^{1/2}$)

Teorema:

Si A es una matriz real de orden n , entonces:

(i) $[\rho(A^t A)]^{1/2} = \|A\|_e$ (norma euclidiana)

(ii) $\rho(A) \leq \|A\|$ para cualquier norma.

La demostración puede encontrarse en Ortega [18].

Definición:

Llamamos a una matriz A de orden n una matriz convergente si:

$$\lim_{k \rightarrow \infty} (A^k) = \mathbf{0},$$

Teorema:

Las siguientes afirmaciones son equivalentes:

- (i) A es una matriz convergente.
- (ii) $\lim_{n \rightarrow \infty} \|A^n\| = 0$, para alguna norma $\| \cdot \|$
- (iii) $\rho(A) < 1$
- (iv) $\lim_{n \rightarrow \infty} A^n x = \mathbf{0}$, $\forall x$.

La demostración puede encontrarse en Isaacson y Keller [11].

Lema:

Si el radio espectral $\rho(B) < 1 \Rightarrow \exists (I - B)^{-1} = I + B + B^2 + \dots$

Demostración:

Sea λ un valor propio de la matriz B , entonces

$$Bx = \lambda x \quad (x \neq \mathbf{0})$$

Y como $Ix = x$, se tiene que

$$(I - B)x = (1 - \lambda)x,$$

lo que implica que

$$\mu = 1 - \lambda$$

es valor propio de la matriz $(I - B)$.

Como para todo λ , valor propio de B , se tiene que:

$$|\lambda| \leq \rho(B) < 1,$$

se puede afirmar, que todos los valores propios de la matriz $(I - B)$ son diferentes de cero, y por lo tanto la matriz $(I - B)$ es no singular, es decir existe la matriz inversa $(I - B)^{-1}$.

Ahora, sea

$$S_k = I + B + B^2 + B^3 + \dots + B^k.$$

Luego $(I - B)S_k = I - B^{k+1}$ y tomando el límite:

$$\lim_{k \rightarrow \infty} (I - B)S_k = \lim_{k \rightarrow \infty} (I - B^{k+1}) = I,$$

ya que

$$\lim_{k \rightarrow \infty} B^{k+1} = 0 \quad (\text{ver teorema anterior}).$$

De donde,

$$S_k = (I - B)^{-1}$$

Lo que prueba la afirmación del Lema. \square

Teorema de convergencia:

Para cualquier $\mathbf{x}^{(0)} \in \mathfrak{R}^n$, la sucesión $\{\mathbf{x}^{(k)}\}_{(k \geq 0)}$ definida por

$$\mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c}, \quad \forall k \geq 1 \text{ y } \mathbf{c} \neq \mathbf{0},$$

converge a la solución única de la ecuación $\mathbf{x} = B\mathbf{x} + \mathbf{c} \Leftrightarrow \rho(B) < 1$.

Demostración:

El proceso iterativo nos lleva a considerar la sucesión $\{\mathbf{x}^{(k)}\}_{(k \geq 0)}$:

$$\begin{aligned} \mathbf{x}^{(k)} &= B\mathbf{x}^{(k-1)} + \mathbf{c} = \\ &= B(B\mathbf{x}^{(k-2)} + \mathbf{c}) + \mathbf{c} = \\ &= B^2\mathbf{x}^{(k-2)} + (B + I)\mathbf{c} = \dots = \\ &= B^k\mathbf{x}^{(0)} + (B^{k-1} + \dots + B + I)\mathbf{c} \end{aligned}$$

y tomando el límite:

$$\lim_{k \rightarrow \infty} \mathbf{x}^k = \lim_{k \rightarrow \infty} B^k \mathbf{x}^{(0)} + \lim_{k \rightarrow \infty} (I + B + B^2 + \dots + B^{k-1}) \mathbf{c}$$

Pero, como $\rho(B) < 1$, se tiene que:

$$\lim_{k \rightarrow \infty} B^k \mathbf{x}^{(0)} = \mathbf{0}$$

por el teorema anterior. Además, por el Lema se sabe que:

$$(I + B + B^2 + \dots + B^{k-1}) = (I - B)^{-1},$$

por lo tanto:

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \lim_{k \rightarrow \infty} (I + B + B^2 + \dots + B^{k-1})\mathbf{c} = (I - B)^{-1}\mathbf{c}$$

Como:

$$\mathbf{x} = B\mathbf{x} + \mathbf{c} \quad \text{ó} \quad (I - B)\mathbf{x} = \mathbf{c}$$

Entonces la sucesión $\{\mathbf{x}^{(k)}\}_{(k \geq 0)}$ construida como:

$$\mathbf{x}^{(k)} = B\mathbf{x}^{(k-1)} + \mathbf{c}$$

converge a \mathbf{x} , la única solución de ésta última ecuación:

$$\mathbf{x} = (I - B)^{-1}\mathbf{c}.$$

Supongamos ahora que,

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x},$$

para probar la segunda parte del teorema.

Como se sabe que:

$$\begin{aligned} \mathbf{x} &= B\mathbf{x} + \mathbf{c} \\ \mathbf{x}^{(k)} &= B\mathbf{x}^{(k-1)} + \mathbf{c} \end{aligned}$$

Entonces

$$\mathbf{x} - \mathbf{x}^{(k)} = B(\mathbf{x} - \mathbf{x}^{(k-1)}) = \dots = B^k(\mathbf{x} - \mathbf{x}^{(0)}),$$

con $\mathbf{x}^{(0)}$ vector inicial arbitrario.

Sea $\mathbf{z} \in \mathfrak{R}^n$ un vector cualquiera, tomando $\mathbf{x}^{(0)} = \mathbf{x} - \mathbf{z}$, se tiene que:

$$\lim_{k \rightarrow \infty} B^k \mathbf{z} = \lim_{k \rightarrow \infty} B^k (\mathbf{x} - \mathbf{x}^{(0)}) = \lim_{k \rightarrow \infty} (\mathbf{x} - \mathbf{x}^{(k)}) = \mathbf{x} - \mathbf{x} = \mathbf{0}$$

y por lo tanto por el teorema anterior,

$$\rho(B) < 1. \square$$

Corolario:

Si $\|B\| < 1$ para cualquier norma matricial, entonces la sucesión $\{\mathbf{x}^{(k)}\}_{(k \geq 0)}$ obtenida de $\mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c}$, converge $\forall \mathbf{x}^{(0)} \in \mathfrak{R}^n$, a la solución $\mathbf{x} \in \mathfrak{R}^n$, de la ecuación $\mathbf{x} = B\mathbf{x} + \mathbf{c}$ y se satisfacen las siguientes cotas de error:

$$(a) \quad \|x - x^{(k)}\| \leq \|B\|^k \|x^{(0)} - x\| ;$$

$$(b) \quad \|x - x^{(k)}\| \leq \frac{\|B\|^k}{1 - \|B\|} \|x^{(1)} - x^{(0)}\|$$

Demostración:

Como

$$\begin{aligned} x &= Bx + c \\ x^{(k)} &= Bx^{(k-1)} + c \end{aligned}$$

se tiene que:

$$\begin{aligned} \|x - x^{(k)}\| &= \|B(x - x^{(k-1)})\| \leq \|B\| \|x - x^{(k-1)}\| = \|B\| \|B(x - x^{(k-2)})\| \leq \\ &\leq \|B\|^2 \|x - x^{(k-2)}\| \leq \dots \leq \|B\|^k \|x - x^{(0)}\| \end{aligned}$$

lo que prueba (a).

Para probar (b), hacemos:

$$x - x^{(k-1)} = x - x^{(k-1)} + x^{(k)} - x^{(k)} = (x - x^{(k)}) + (x^{(k)} - x^{(k-1)})$$

de donde:

$$\|x - x^{(k-1)}\| \leq \|x - x^{(k)}\| + \|x^{(k)} - x^{(k-1)}\|$$

de (a), tenemos:

$$\|x - x^{(k-1)}\| \leq \|B\| \|x - x^{(k-1)}\|$$

Entonces:

$$(1 - \|B\|) \|x - x^{(k-1)}\| \leq \|x^{(k)} - x^{(k-1)}\|$$

$$\|x - x^{(k)}\| \leq \frac{\|B\|}{1 - \|B\|} \|x^{(k)} - x^{(k-1)}\|$$

La fórmula que nos permite estimar la k -ésima aproximación del vector x , usando las dos aproximaciones sucesivas $x^{(k)}$, y $x^{(k-1)}$, además de $\|B\|$.

Procediendo de la misma forma se obtiene que:

$$\|x - x^{(k)}\| \leq \frac{\|B\|^k}{1 - \|B\|} \|x^{(1)} - x^{(0)}\|$$

La fórmula que usa la primera aproximación y la aproximación inicial, además de $\|B\|^k$. El término $\|B\|^k$, indica que la convergencia depende de $\|B\|$, en el sentido de que si $\|B\|$ es cercana a 1, la convergencia puede ser muy lenta; problema que se presenta con frecuencia en los métodos de Punto Fijo.

Para aplicar los resultados anteriores a los métodos iterativos dados tenemos la matriz B para el método de Jacobi y el de Gauss-Seidel respectivamente:

$$B_J = -D^{-1}(L+U), \quad B_G = -(L+D)^{-1}U$$

Si $\rho(B_J) < 1$ ó $\rho(B_G) < 1 \Rightarrow \{x^{(k)}\}_{(k \geq 0)}$ converge a la solución x de que también es solución de $x = Bx + c$

$$Ax = b.$$

Definición:

Si A es de orden n , $A = [a_{ij}]_{n \times n}$, se dice que es estrictamente diagonalmente dominante (por filas), si y sólo si:

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \text{ para } i = 1, 2, \dots, n.$$

Teorema:

Si A es estrictamente diagonalmente dominante entonces, para cualquier elección de $x^{(0)}$ ambos métodos, el de Jacobi o el de Gauss-Seidel, dan lugar a sucesiones $\{x^{(k)}\}_{(k \geq 0)}$ que convergen a la solución de $Ax = b$.

La demostración puede encontrarse en Ortega [18].

Comentario:

Con relación a la matriz B_J hallada para la aplicación del método de Jacobi del **Ejemplo 1**, observamos de que:

$$\|B_J\|_u = \frac{2}{3} < 1,$$

luego cumple con la condición de convergencia del corolario anterior, y en relación con la matriz B_G hallada para la aplicación del método de Gauss-Seidel, se puede notar de que:

$$\|B_G\|_u = \frac{4}{9} < 1,$$

vemos que también satisface la condición de convergencia del corolario. Notar que:

$$\|B_G\|_u < \|B_J\|_u,$$

lo que explica porque la convergencia del método de Gauss-Seidel es más rápida que la convergencia del método de Jacobi para el sistema lineal (ver corolario)

Sobre la base de la condición de convergencia del último teorema se puede apreciar que para la matriz A :

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

del sistema utilizado en los ejemplos (1) y (3), se cumple que la matriz A es estrictamente diagonalmente dominante.

La relación de la rapidez de convergencia con el radio espectral de la matriz de iteración B se puede ver de la desigualdad:

$$\begin{aligned} \|x - x^{(k)}\| &\leq \|B\|^k \|x^{(0)} - x\| \text{ y como } \rho(B) \leq \|B\| \\ &\Rightarrow \|x^{(k)} - x\| \approx \rho^k(B) \|x^{(0)} - x\| \end{aligned}$$

Supongamos que $\rho(B) < 1$ y que $x^{(0)} = \mathbf{0}$, luego el error relativo después de k iteraciones será aproximadamente igual a $\rho^k(B)$. Si se

trabaja con una aritmética de t dígitos, y se espera una precisión de 10^{-t} , se tiene:

$$\rho^k(B) \leq 10^{-t}$$

Entonces $k \geq -t / \log(\rho(B))$, es deseable escoger la técnica iterativa con $\min\{\rho(B)\}$ para resolver el sistema lineal $Ax = b$.

3.7 Error y Residual de una Solución Aproximada

Definición:

Si $x \in \mathfrak{R}^n$ es una aproximación a la solución del sistema lineal definido $Ax = b$, el vector residual de x con respecto a este sistema se define como:

$$r \text{ (vector residual)} = b - A\hat{x} = Ax - A\hat{x} = A(x - \hat{x}) = Ae,$$

donde

$$e \text{ (vector error)} = x - \hat{x} \Rightarrow e = A^{-1}r, \text{ si existe } A^{-1}.$$

Observar que el vector error es la solución del sistema lineal, $r = Ae$, con la matriz A del sistema dado y el vector no homogéneo, el vector residual.

De las ecuaciones:

$$Ae = r \quad \text{y} \quad e = A^{-1}r$$

y por las propiedades de normas, se obtiene (1) y (2):

$$(1)... \quad \|r\| \leq \|A\| \|e\| \Rightarrow \frac{\|r\|}{\|A\|} \leq \|e\|$$

$$(2)... \quad \|e\| \leq \|A^{-1}\| \|r\|$$

$$\Rightarrow \frac{\|r\|}{\|A\|} \leq \|e\| \leq \|A^{-1}\| \|r\| \quad \dots(i)$$

y de manera similar de las ecuaciones:

$$Ax = b, \text{ y } x = A^{-1}b,$$

se obtienen (3) y (4):

$$(3)... \quad \|b\| \leq \|A\| \|x\|$$

$$(4)... \quad \|x\| \leq \|A^{-1}\| \|b\|,$$

$$\text{por lo tanto } \frac{\|r\|}{\|b\|} \leq \|x\| \leq \|A^{-1}\| \|b\|, \quad \dots \text{ (ii)}$$

De (i) y (ii), se tiene:

$$\frac{\|r\|}{\|b\|} \frac{1}{\|A\| \|A^{-1}\|} \leq \frac{\|e\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|r\|}{\|b\|}$$

Lo que nos muestra las cotas para el error relativo: $\frac{\|e\|}{\|x\|}$ en

términos del residual relativo: $\frac{\|r\|}{\|b\|}$

Notar que en ambas cotas interviene el número $\|A\| \|A^{-1}\|$, que se define como el número de condición de A , $K(A) = \|A\| \|A^{-1}\|$, y se observa que el error relativo de una solución aproximada puede ser tan grande como $K(A)$ veces su residual relativo, y tan pequeño como $1/K(A)$ veces su residual relativo.

Si $K(A) \approx 1$, entonces el error relativo es aproximadamente igual al residual relativo, y éste último se puede usar como un estimativo del error. $K(A)$ depende de la norma matricial usada y puede para algunas matrices variar considerablemente cuando la norma matricial se cambia.

Como:

$$\|I\| = \max_{x \neq 0} \frac{\|x\|}{\|x\|} = 1,$$

Se tiene que:

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\| \|A^{-1}\| = K(A)$$

esto es:

$$K(A) \geq 1.$$

El número de condición es por consiguiente una medida de cuán bien podemos esperar distinguir una solución (aproximada) “buena” de una “mala” observando el error residual.

Generalmente, se dice que la matriz A tiene un buen comportamiento y que está bien condicionada si $K(A) \approx 1$, es un caso ideal, ya que en la práctica $K(A) \gg 1$.

Aplicaciones:

Ejemplo 1:

Sea la matriz:

$$A = \begin{bmatrix} 1 & 0.67 & 0.33 \\ 0.45 & 1 & 0.55 \\ 0.67 & 0.33 & 1 \end{bmatrix},$$

entonces $\|A\|_u = 2$ y $A^{-1} = \begin{bmatrix} 1.383068 & -0.948124 & 0.065056 \\ -0.137715 & 1.316154 & -0.678439 \\ -0.881210 & 0.200912 & 1.1802974 \end{bmatrix}$

y

$$\|A^{-1}\|_u = 2.39625, \text{ por lo tanto } K(A) = 4.7925$$

luego es una matriz bien condicionada.

Ejemplo 2:

Sea la matriz:

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}, \quad \text{y} \quad \|A\|_u = 11/6,$$

$$\text{y} \quad A^{-1} = \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}, \quad \text{luego} \quad \|A^{-1}\|_u = 408$$

$$\Rightarrow K(A) = 748.$$

Por lo que es una matriz mal condicionada. Esta matriz cae dentro de las matrices de la familia de Hilbert, de la forma:

$$A = [a_{ij}]_{(i,j=1,2,3)}, \quad \text{donde} \quad a_{ij} = \frac{1}{i+j-1}; \quad \text{para } i, j = 1, 2, 3.$$

3.8 Método de Potencia

El problema de valores propios de una matriz se presenta con frecuencia en los problemas de Física e Ingeniería, donde intervienen los movimientos oscilatorios o vibratorios y los valores propios se relacionan con las frecuencias de estos movimientos.

Resulta que si el orden de la matriz A es mayor que tres y hay pocos elementos nulos, el trabajo de construir el polinomio característico se complica mucho. Además en la práctica, con frecuencia sólo se necesita el valor propio de mayor magnitud o de menor magnitud de A .

Construyamos un método para determinar el valor propio dominante de A ; es decir el valor propio de mayor magnitud:

Sean $\lambda_1, \lambda_2, \dots, \lambda_n$ los valores propios de una matriz A y u_1, u_2, \dots, u_n los vectores propios correspondientes, de manera que λ_1 sea el único de mayor magnitud, es decir:

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|,$$

Si $u^{(0)}$ es un vector arbitrario, se puede presentar este vector, como una combinación lineal de los vectores propios correspondientes a λ :

$$u^{(0)} = \sum_{i=1}^n \alpha_i u_i$$

Ahora, multiplicando esta ecuación por la matriz A , se tiene:

$$Au^{(0)} = A \sum_{i=1}^n \alpha_i u_i = \sum_{i=1}^n \alpha_i Au_i = \sum_{i=1}^n \alpha_i \lambda_i u_i, \text{ ya que } Au_i = \lambda_i u_i.$$

Luego multiplicando nuevamente por A obtenemos:

$$A^2 u^{(0)} = A(Au^{(0)}) = \sum_{i=1}^n \alpha_i \lambda_i^2 u_i,$$

y así sucesivamente:

$$A^m u^{(0)} = \sum_{i=1}^n \alpha_i \lambda_i^m u_i = \lambda_1^m \sum_{i=1}^n \left[\alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^m u_i \right]$$

entonces si tomamos límite,

$$\lim_{m \rightarrow \infty} A^m u^{(0)} = \lim_{m \rightarrow \infty} (\lambda_1^m \alpha_1 u_1),$$

ya que

$$\lim_{m \rightarrow \infty} \left(\frac{\lambda_2}{\lambda_1} \right)^m = \dots = \lim_{m \rightarrow \infty} \left(\frac{\lambda_n}{\lambda_1} \right)^m = 0$$

por ser λ_1 el valor propio dominante.

A medida que se construyen los vectores:

$$A^m \mathbf{u}^{(0)}; \text{ para } m = 1, 2, 3, \dots,$$

a partir de $\mathbf{u}^{(0)}$ vector arbitrario, es conveniente normalizarlos con el fin de trabajar sólo con vectores unitarios. Designemos a $\mathbf{u}^{(0)}$, un vector unitario.

Sea el vector $A\mathbf{u}^{(0)}$, entonces supongamos que la componente que proporciona la norma de $\mathbf{u}^{(0)}$ está en el lugar i_0 , donde i_0 es el menor entero tal que $1 \leq i_0 \leq n$, y se define al vector unitario $\mathbf{u}^{(1)}$ como:

$$\mathbf{u}^{(1)} = \frac{A\mathbf{u}^{(0)}}{\|A\mathbf{u}^{(0)}\|_u}$$

y al valor $\lambda^{(1)}$, como:

$$\lambda^{(1)} = \frac{(A\mathbf{u}^{(0)})_{i_0}}{(\mathbf{u}^{(0)})_{i_0}}$$

donde $(\mathbf{x})_{i_0}$ indica la i_0 -ésima componente del vector \mathbf{x} , en general. Luego, construiremos al vector unitario:

$$\mathbf{u}^{(2)} = \frac{A^2 \mathbf{u}^{(0)}}{\|A^2 \mathbf{u}^{(0)}\|_u}$$

Y al valor asociado:

$$\lambda^{(2)} = \frac{(A^2 \mathbf{u}^{(0)})_{i_1}}{(A \mathbf{u}^{(0)})_{i_1}}$$

Donde i_1 es el menor entero tal que $1 \leq i_1 \leq n$, de manera que:

$$\|A \mathbf{u}^{(0)}\|_u = |(A \mathbf{u}^{(0)})_{i_1}|.$$

De esta forma se define al:

Algoritmo del Método de Potencia

Objetivo: Estimar al valor propio dominante y al vector propio correspondiente de la matriz A de orden $n \times n$ dado un vector no cero $\mathbf{u}^{(0)}$ de norma unitaria:

Hacer para $m = 1, 2, \dots$, hasta donde se satisfaga la tolerancia pedida:

$$\mathbf{u}^{(m)} = \frac{A^m \mathbf{u}^{(0)}}{\|A^m \mathbf{u}^{(0)}\|_u}$$

$$\lambda^{(m)} = \frac{(A^m \mathbf{u}^{(0)})_{i_{m-1}}}{(A^{(m-1)} \mathbf{u}^{(0)})_{i_{m-1}}}$$

con i_{m-1} el menor entero tal que $1 \leq i_{m-1} \leq n$, de manera que:

$$\|A^{(m-1)} \mathbf{u}^{(0)}\|_u = |(A^{(m-1)} \mathbf{u}^{(0)})_{i_{m-1}}|$$

Se puede probar que:

$\lim_{m \rightarrow \infty} \lambda^{(m)} = \lambda_1$, es el valor propio dominante,

y que $\lim_{m \rightarrow \infty} \mathbf{u}^{(m)} = \mathbf{u}_1$, es el vector propio correspondiente al valor propio dominante λ_1 .

Ejemplo:

Hallar el valor propio de mayor magnitud así como el vector propio correspondiente de la matriz:

$$A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & -3 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

Solución:

Sea $\mathbf{u}^{(0)} = [1, 1, 1]'$ tal que su norma es 1, $\|\mathbf{u}^{(0)}\|_u = 1$, realizando las iteraciones para el método de Potencia, tenemos:

(i) Para $m = 1$,

$$\lambda^{(1)} = \frac{(A\mathbf{u}^{(0)})_{i_0}}{(\mathbf{u}^{(0)})_{i_0}} = \frac{[4 \quad -3 \quad 4]_{i_0}'}{[1 \quad 1 \quad 1]_{i_0}'} = 4$$

$$\mathbf{u}^{(1)} = \frac{A\mathbf{u}^{(0)}}{\|A\mathbf{u}^{(0)}\|_u} = \frac{[4 \quad -3 \quad 4]_{i_0}'}{4} = [1 \quad -3/4 \quad 1]_{i_0}'$$

El valor de i_0 , se determina como la primera componente del vector $\mathbf{u}^{(0)}$, que nos proporciona su norma, y en este caso su norma es 1, y todas las componentes son iguales, luego se elige a $i_0 = 1$.

(ii) Para $m = 2$,

$$\lambda^{(2)} = \frac{(A^2\mathbf{u}^{(0)})_{i_1}}{(\mathbf{u}^{(0)})_{i_1}} = \frac{[16 \quad 9 \quad 16]_{i_1}'}{[4 \quad -3 \quad 4]_{i_1}'} = 4$$

$$u^{(2)} = \frac{A^2 u^{(0)}}{\|A^2 u^{(0)}\|_u} = \frac{[16 \quad 9 \quad 16]^t}{16} = [1 \quad 9/16 \quad 1]^t$$

El valor de i_1 , se determina como la primera componente del vector $Au^{(0)}$, que nos proporciona su norma, y en este caso su norma es 4, y dos de las componentes son iguales a 4, pero se elige a $i_1 = 1$, por ser la primera que nos da su norma.

(iii) Para $m = 3$,

$$\dot{e}^{(3)} = \frac{(A^3 u^{(0)})_{i_2}}{(A^2 u^{(0)})_{i_2}} = \frac{[64 \quad -27 \quad 64]_{i_2=1}^t}{[16 \quad 9 \quad 16]_{i_2=1}^t} = 4$$

$$u^{(3)} = \frac{A^3 u^{(0)}}{\|A^3 u^{(0)}\|_u} = \frac{[64 \quad -27 \quad 64]^t}{64} = [1 \quad -27/64 \quad 1]^t$$

En forma similar a los dos casos anteriores, se determina a $i_2 = 1$.

Resumiendo en la siguiente tabla se muestran las iteraciones:

m	$\lambda^{(m)}$	$u_1^{(m)}$	$u_2^{(m)}$	$u_3^{(m)}$
1	4	1	-3/4	1
2	4	1	9/16	1
3	4	1	-27/64	1
4	4	1	81/128	1
5	4	1	-243/1024	1
6	4	1	729/4096	1
10	4	1	0.0563135	1
20	4	1	3.17×10^{-3}	1
30	4	1	1.786×10^{-4}	1
40	4	1	1.006×10^{-5}	1
50	4	1	5.66×10^{-7}	1
60	4	1	3.1892×10^{-8}	1
70	4	1	1.79×10^{-9}	1
∞	4	1	0	1

Obteniéndose como valor propio de mayor magnitud a $\lambda = 4$, y al vector propio correspondiente $[1, 0, 1]^t$. Observamos que este

resultado coincide con la respuesta obtenida con el ejemplo que realizamos para los valores y vectores propios de la matriz A .

3.9 Ejercicios

1. Dado el siguiente sistema:

$$\begin{bmatrix} 0.0005 & 235789 \\ 9.237 & -8.954 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 235923 \\ 8.304 \end{bmatrix}$$

resolverlo, usando Eliminación Gaussiana, con Pivoteo Parcial, y Escalonado con una aritmética de cuatro dígitos.

2. Resolver el sistema:

$$\begin{aligned} -1.41x + 2y &= 1 \\ x - 1.41y + z &= 1 \\ 2y - 1.41z &= 1 \end{aligned}$$

Haciendo uso del método de Eliminación Gaussiana, con las estrategias de:

- Pivoteo Parcial
 - Pivoteo Escalonado
3. Resuelva:

$$\begin{bmatrix} 0 & 4 & 5 & 7 & 8 & 6 & 7 \\ 9 & 0.0078 & 3 & 1 & 7 & 9 & 6 \\ 5 & 3 & 2 & -1 & -7 & 8 & 10 \\ -1 & 2 & 3 & 0 & 4 & 5 & 3 \\ 11 & 10 & 9 & 4 & 2 & 7 & 6 \\ 15 & 11 & 8 & 7 & 6 & 9 & 3 \\ 12 & 1 & 3 & 6 & 5 & 7 & 10 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \\ 1 \\ 6 \\ 3 \end{bmatrix}$$

haciendo uso de la Eliminación Gaussiana con la estrategia de Pivoteo Escalonado.

4. Hallar el vector x , si:

$$\begin{bmatrix} 0 & -1 & 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 2 & 1 & 3 \\ -1 & 0 & 1 & -2 & 0 & 4 \\ 2 & 3 & -1 & 0 & 1 & 2 \\ 3 & -1 & 0 & 2 & 3 & 1 \\ -2 & 0 & 1 & -2 & 3 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 2 \\ 3 \\ -2 \end{bmatrix}$$

y si:

$$\begin{bmatrix} -1 & 2 & 3 & 4 & 1 & -1 \\ -2 & 0 & -1 & 1 & 2 & 3 \\ -1 & 2 & 3 & -2 & 1 & -3 \\ 4 & -1 & -2 & 3 & 2 & 1 \\ 0 & 1 & -1 & 2 & -2 & 0 \\ -1 & 0 & 1 & 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{bmatrix}$$

aplicando Eliminación Gaussiana con Pivoteo Escalado.

5. a. Haciendo uso del método de Eliminación Gaussiana con la estrategia de Pivoteo Escalonado, resolver:

$$(i) \begin{bmatrix} 1 & -2 & 1 & 1 \\ 2 & 0 & -1 & 2 \\ -1 & -3 & 2 & 1 \\ 3 & 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix} \quad (ii) \begin{bmatrix} -1 & 2 & 0 & 3 & 4 \\ 2 & 1 & -2 & 0 & 3 \\ 3 & -1 & 1 & 1 & -1 \\ 2 & 4 & -1 & 0 & 1 \\ 3 & -2 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \\ u \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$(iii) \begin{bmatrix} 1 & 1 & 3 & 2 \\ 5 & 0 & -2 & 7 \\ 9 & 2 & 2 & 5 \\ 2 & -1 & 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix} \quad (iv) \begin{bmatrix} 3 & 5 & 5 & 7 & 3 \\ 7 & 4 & 2 & 3 & 2 \\ 11 & 1 & 0 & 1 & -1 \\ 2 & -2 & 2 & 4 & 3 \\ 5 & 3 & 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \\ u \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \\ 7 \\ 9 \end{bmatrix}$$

b. Hallar la inversa de las matrices dadas en el ítem (a).

6. a. Descomponer la matriz:

$$A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 5 & 1 \\ 2 & 1 & 6 \end{bmatrix}$$

por medio del método de Doolittle.

b. Utilice la descomposición hallada en (a) para resolver los sistemas:

$$Ax = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} ; \quad Ax = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

7. Dada la matriz:

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & -3 \end{bmatrix}$$

- Descomponerla por el método de Doolittle.
- Hallar A^{-1}
- Resolver los sistemas:

$$(1) \quad Ax = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}; \quad (2) \quad Ax = \begin{bmatrix} 6 \\ 4 \\ 5 \end{bmatrix}, \quad (3) \quad Ax = \begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix} \quad (4) \quad Ax = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$$

8. Una compañía minera trabaja en 3 minas, cada una de las cuales produce minerales de tres clases. La primera mina puede producir 4 Tm. del mineral A, 3 Tm. del mineral B, y 5 Tm. del mineral C; la segunda mina puede producir 1 Tm. de cada uno de los minerales, y la tercera mina, 2 Tm. del A, 4 Tm del B y 3 Tm del C, por cada hora de funcionamiento.

Haciendo uso del método de factorización triangular (Doolittle) diga: ¿cuántas horas se debe trabajar en cada mina para satisfacer los tres pedidos siguientes?

Pedidos	Minerales		
	A	B	C
P ₁	19	25	25
P ₂	13	16	16
P ₃	8	12	10

9. El estudio de cadenas alimenticias es un tema importante en la determinación del esparcimiento y contaminación de contaminantes del medio ambiente en la materia viviente. Suponga que una cadena alimenticia tiene tres eslabones. El primer eslabón consiste de vegetación de los tipos: v_1, v_2, \dots, v_n , la cual provee todos los requerimientos alimenticios a los herbívoros de las especies h_1, h_2, \dots, h_n , en el segundo eslabón. El tercer eslabón consiste de animales carnívoros del tipo c , los cuales dependen totalmente para el suministro de su alimento de los herbívoros del segundo eslabón.

La componente a_{ij} de la matriz $A = [a_{ij}]_{n \times n}$, representa el número total de plantas del tipo v_j , que son comidas por los herbívoros de la especie h_i , mientras que la componente x_i del vector $x = [x_i]_{n \times 1}$, describe el número de herbívoros de la especie h_i , que son devorados por los animales del tipo c .

En la ecuación: $Ax = b$,

la componente b_i , del vector b , representa el número de plantas del tipo v_i que eventualmente terminarán en los animales de la especie c .

a. Sea:

$$\begin{bmatrix} 1 & 3 & 0 \\ 3 & 1 & 4 \\ 4 & 6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = b = \begin{bmatrix} 1 \\ 8 \\ 6 \end{bmatrix}$$

Hallar el número de herbívoros de la especie h_i , que son devorados por los animales del tipo c .

- i. Usando Eliminación Gaussiana, con una aritmética de tres dígitos.
- ii. Usando Eliminación Gaussiana con Pivoteo Parcial y empleando una aritmética de tres dígitos.
- iii. Hallar la factorización triangular de Doolittle de A y utilizar este resultado para resolver los sistemas:

$$Ax = \begin{bmatrix} 4 \\ 7 \\ 8 \end{bmatrix} \quad \text{y} \quad Ax = \begin{bmatrix} 5 \\ 4 \\ 6 \end{bmatrix}$$

iv. Hallar la inversa de A .

10. Dado el siguiente sistema:

$$\begin{bmatrix} 6 & 1 & 0 & 1 & 1 & 2 \\ 1 & 6 & 1 & 1 & 0 & 2 \\ 0 & 2 & 6 & 1 & 0 & 2 \\ 2 & 1 & 0 & 5 & 0 & 1 \\ 0 & 1 & 2 & 1 & 6 & 1 \\ 1 & 1 & 0 & 1 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \\ 5 \\ 8 \\ 7 \\ 6 \end{bmatrix}$$

- i. ¿Cuántas iteraciones se necesitarían realizar para lograr con el método de Jacobi una precisión de $\varepsilon = 10^{-3}$?
- ii. Resuelva el sistema dado, empleando el método de Gauss-Seidel, con una precisión de $\varepsilon = 10^{-6}$.
11. a. Considere la siguiente ecuación diferencial de segundo orden:

$$y''(x) - p(x)y'(x) - q(x)y(x) = r(x)$$

y los puntos igualmente espaciados:

$$x_j = a + jh \text{ con } x_0 = a \text{ y } x_N = b \text{ donde } j = 0, 1, \dots, N;$$

y empleando las aproximaciones:

$$y''(x_j) \approx [y(x_{j+1}) - 2y(x_j) + y(x_{j-1}))] / h^2$$

$$y'(x_j) \approx [y(x_{j+1}) - y(x_{j-1}))] / 2h$$

demuestre que el problema de resolver la ecuación diferencial de segundo orden aproximadamente para determinar los valores $y_j = y(x_j)$, se reduce a hallar la solución de un sistema de ecuaciones lineales.

- b. Haciendo uso de la parte (a), resuelva la ecuación:

$$y'' + y' + xy = 0 \quad \text{con } y(0) = 1 \text{ e } y(1) = 0$$

Emplee un método iterativo para:

- (i) $N = 5$, $TOL = 10^{-4}$
 (ii) $N = 10$, $TOL = 10^{-8}$

- c. Representemos por u al potencial electrostático entre dos esferas metálicas concéntricas de radios R_1 y R_2 ($R_1 < R_2$), tales que el potencial de la esfera interior se mantiene constante a V_1 voltios y el potencial de la esfera exterior a 0 voltios. El potencial en la región entre las dos esferas está gobernado por la ecuación de Laplace, la cual en esta aplicación se reduce a:

$$\begin{aligned} d^2u/dr^2 + (2/r)(du/dr) &= 0, \quad R_1 \leq r \leq R_2, \\ u(R_1) &= V_1, \quad u(R_2) = 0 \end{aligned}$$

Supóngase que $R_1 = 2$ pulg., $R_2 = 4$ pulg., y $V_1 = 110$ voltios.

Haciendo uso de la parte (a), genere el sistema lineal correspondiente para $N = 15$, y resuelva dicho sistema empleando el método de Jacobi, con $TOL = 10^{-6}$

12. Suponga que Juan es un caminante distraído que se desplaza de tal manera que da un paso hacia la izquierda con una frecuencia triple que hacia la derecha sobre una línea con posiciones numeradas del 0 al 4. Cuando llega a un extremo se para. Cada valor x_k , representa su probabilidad de llegar al extremo izquierdo a partir de su posición k . Se puede definir $x_0 = 1$ y $x_4 = 0$, en cuyo caso la ecuación tiene la forma:

$$x_k = (3/4) x_{k-1} + (1/4) x_{k+1} \quad ; \quad k = 1, 2, 3.$$

Aplique el método de Gauss-Seidel, para resolver el sistema dado, con una $TOL = 10^{-3}$.

13. La Universidad "AAA" tiene estudiantes de:

- Ingeniería Civil (C)
- Otras Ingenierías (I)
- Ciencias y Economía (E)

Los presupuestos respectivos por pagos de profesores son:

Ingeniería Civil	:	8'000,000 u.m.
Otras Ingenierías	:	16'000,000 u.m.
Ciencias y Economía	:	5'000,000 u.m.

El número de estudiantes en:

Ingeniería Civil	:	2,000
Otras Ingenierías	:	4,000
Ciencias y Economía	:	1,000

En el sistema de Departamentos la distribución de cursos fue:

Estudiantes	Profesores		
	C	I	E
C	70%	10%	15%
I	20%	90%	10%
E	10%	0%	75%

Hallar el costo por estudiante de cada grupo, considerando únicamente el costo ocasionado por los profesores de su facultad.

Haga uso de los métodos iterativos:

- Jacobi
 - Gauss-Seidel
- con una tolerancia de 10^{-4} .

Sugerencia: defina:

- x_1 = pago total a todos los profesores de Ingeniería Civil.
 x_2 = pago total a todos los profesores de las otras ingenierías.
 x_3 = pago total a todos los profesores de Ciencias y Economía.

14. a. Sea el sistema:

$$\begin{bmatrix} 8 & 0 & -5 \\ 1 & 3 & 1 \\ 3 & 1 & -6 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

- emplee el método de potencia para hallar el radio espectral de las matrices, B , correspondientes a cada uno de los métodos iterativos, presentando en cada caso las 6 primeras iteraciones.
- b. Resuelva el sistema dado, presentando las 10 primeras iteraciones, empleando el método más conveniente.
- c. Calcule la cota del error absoluto para la solución obtenida en (b)
15. Sea el sistema lineal:

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

- a. ¿Qué método iterativo: Jacobi o Gauss-Seidel le convendría usar? Justifique
- b. Realice tres iteraciones usando el método que eligió en (a) e indique una cota superior para el error absoluto de la solución aproximada que ha hallado.
16. a. Utilice el método de potencia para hallar el valor propio de mayor magnitud con una $TOL = 10^{-3}$ y el vector propio correspondiente, de la matriz:

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

- b. Si λ es un valor propio de A y u el vector propio correspondiente entonces $A^{-1}u = \lambda^{-1}u$; lo que significa que el valor propio de menor magnitud de A se puede hallar como el recíproco del valor propio de mayor magnitud de A^{-1} .

Halle el valor propio de menor magnitud de A , con una $TOL=10^{-3}$.

17. Una matriz persimétrica es una matriz que es simétrica alrededor de ambas diagonales; es decir una matriz $A = [a_{ij}]_{n \times n}$ es persimétrica si $a_{ij} = a_{ji} = a_{N+1-i, N+1-j}$ para $i, j = 1, 2, \dots, N$. Muchos problemas en la teoría de la comunicación tienen soluciones que involucran a los valores y vectores característicos de matrices que están en forma persimétrica. Por ejemplo, el vector característico correspondiente al menor valor característico de la matriz persimétrica:

$$A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

da la respuesta impulsiva de un canal de energía-unidad para una secuencia de error dada, de longitud 2 y subsecuentemente, el peso mínimo de cualquier secuencia de error posible.

Se sabe que si λ satisface la igualdad:

$$|\lambda - 4| = \rho(A - 4I)$$

donde ρ denota el radio espectral, entonces λ es el valor característico mínimo de A . Emplee el método de Potencia para hallar al valor característico mínimo de A , y su vector característico correspondiente con $\varepsilon = 10^{-4}$.

18. Suponga que una especie de escarabajo tiene un lapso de vida de 4 años y que una hembra en el primer año tiene una tasa de sobrevivencia de $1/2$; en el segundo, una de $1/4$; y en el tercer año, una de $1/8$. Suponga además que una hembra procrea, en promedio, dos nuevas hembras en el tercer año de vida, y cuatro nuevas hembras en el cuarto. La matriz que describe la contribución de una hembra en un año a la población de sexo femenino, en los siguientes años es:

$$A = \begin{bmatrix} 0 & 0 & 2 & 4 \\ 1/2 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1/8 & 0 \end{bmatrix}$$

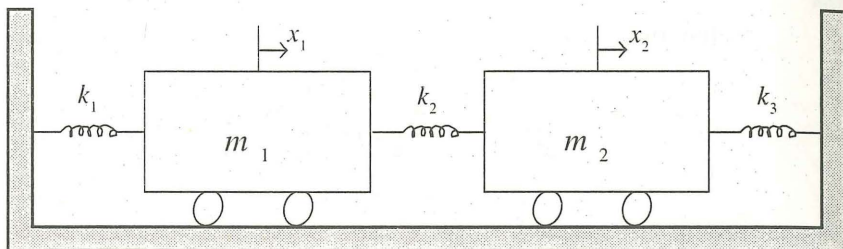
donde a_{ij} = representa la contribución probabilística que hace una hembra de edad j a la población de hembras de edad i en el año siguiente.

- La aportación que un escarabajo de sexo femenino hará a la población en dos años se puede determinar a partir de A^2 , de 3 años a partir de A^3 y así sucesivamente. ¿Cuál es su predicción a largo plazo para la población de estos escarabajos?
 - Emplee el método de Potencia para hallar el valor propio dominante y su correspondiente vector propio, con $TOL = 10^{-2}$,
 - Si inicialmente la población de escarabajos, $x^{(0)}$, coincide con el vector propio hallado en la parte (b), ¿Qué podría afirmar acerca de la población después de 25 años?
19. Para el siguiente sistema:

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

- a. Determine las matrices correspondientes al método iterativo: $\mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c}$ para $k \geq 0$, tanto para el método de Jacobi, como para el de Gauss-Seidel.
- b. ¿Cuántas iteraciones serían necesarias realizarse, para asegurar una precisión de $TOL = 10^{-5}$?, para el método de Jacobi, y para el método de Gauss-Seidel.

20. Considerar el movimiento horizontal del siguiente sistema de masa-resorte, que se muestra en la figura:



Las desviaciones horizontales x_1 y x_2 son medidas relativas a la posición del equilibrio estático. Las constantes de los resortes: k_1 , k_2 , y k_3 , son las fuerzas requeridas para extender o comprimir cada resorte por unidad de longitud.

Si las ecuaciones de movimiento son:

$$m_1 \ddot{x}_1 = -k_1 x_1 + k_2 (x_2 - x_1)$$

$$m_2 \ddot{x}_2 = k_2 (x_1 - x_2) - k_3 x_2$$

- a. Reescribir las ecuaciones en la forma matricial:

$$\ddot{\mathbf{x}} = A\mathbf{x}, \text{ donde } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

- b. Probar que la sustitución $\mathbf{x} = \mathbf{b} e^{j\omega t}$, donde $j = \sqrt{-1}$, induce al problema de valores propios:

$$A\mathbf{b} = \lambda \mathbf{b}, \text{ donde } \lambda = -\omega^2$$

Los posibles valores que ω puede asumir son las frecuencias naturales circulares de vibración del sistema.

- c. Si $k_1 = k_2 = k_3 = 1 \text{ lb/seg}^2$, y $m_1 = m_2 = 1 \text{ lb}$.
Encontrar los valores y vectores propios de A .
- d. Si las condiciones iniciales en $t = 0$, son :

$$x_1(0) = 1, \quad x_2(0) = \dot{x}_1(0) = \dot{x}_2(0) = 0$$

probar que:

$$x(t) = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \cos t + \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix} \cos(t\sqrt{3})$$

sugerencia: Se sabe que:

$$\text{si } t = 0, \quad x(0) = \sum_{i=1}^2 \alpha_i v_i \quad \text{con } \alpha_i \in \mathfrak{R},$$

donde v_i son los vectores propios de A .

Por lo tanto la solución del sistema planteado será:

$$\sum_{i=1}^2 \alpha_i v_i e^{jt\omega_i}, \quad \text{donde } \omega_i = (-\lambda_i)^{1/2},$$

λ_i valor propio de A

- e. Generalizar el problema anterior (a - b) a la situación que incluye n masas: m_1, m_2, \dots, m_n , conectadas con $n + 1$ resortes con constantes: k_1, k_2, \dots, k_{n+1} .

f. Aplique la parte (e) al caso particular:

$$n = 3, m_1 = m_2 = m_3 = 1; k_1 = k_2 = k_3 = k_4 = 1 \text{ lb/seg}^2$$

con las condiciones iniciales:

si $t = 0$,

$$x_1(0) = x_2(0) = 1$$

$$x_3(0) = \dot{x}_1(0) = \dot{x}_2(0) = \dot{x}_3(0) = 0$$

y hallar la solución.

Aproximación de Funciones

En el área de Ciencias e Ingeniería se presenta con frecuencia la necesidad de hacer cálculos, basándose en un número limitado de puntos tabulados que proceden muchas veces de resultados experimentales, y que se desconoce la función matemática que los define. En estos casos es necesario buscar procedimientos que permitan construir modelos matemáticos, que describan el comportamiento de los puntos tabulados, con el objeto de hacer estimaciones de otros valores que no se encuentran tabulados. Dentro de estas técnicas de *Aproximación de Funciones* se encuentran las de: *Interpolación Polinómica* y *Ajuste de Datos*; las cuales trataremos en este capítulo.

La palabra interpolación significa pasar una curva por un conjunto dado de puntos; es decir, matemáticamente entenderíamos que si se tiene un conjunto de puntos:

$$(x_i, y_i) \text{ para } i = 0, 1, 2, \dots, N.$$

deberemos hallar una *función interpolante*, cuya gráfica pase por el conjunto de puntos dado. Así por ejemplo, si tenemos al conjunto de puntos:

i	0	1	2	3
x_i	0.1	0.5	0.9	1.3
$y_i = f(x_i)$	0.099833	0.479426	0.783327	0.963558

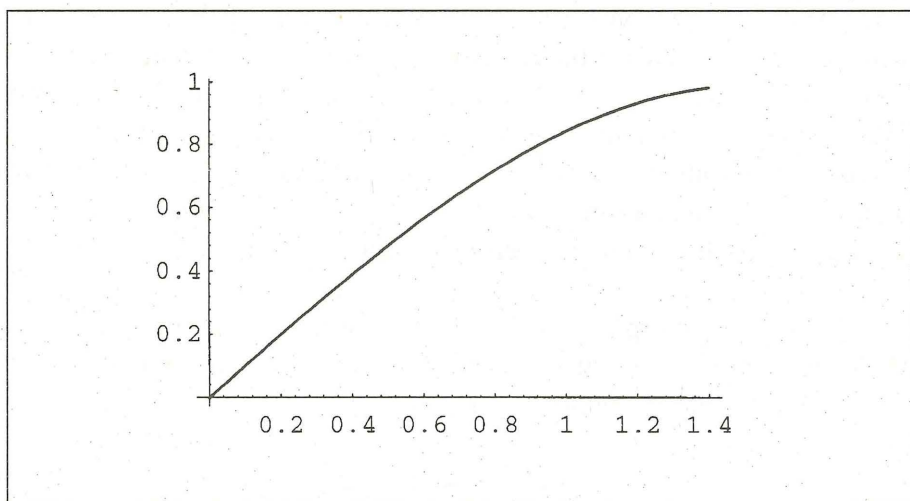
Tabla 1

A través de la *Interpolación Polinómica* construiremos al polinomio de grado tres que pasa por estos puntos:

$$P_3(x) = -0.0012697 + 1.01719x - 0.0491234x^2 - 0.124943x^3$$

que se aproxima a la función $f(x)$ y describe el comportamiento de dichos puntos.

A continuación se muestra en la gráfica como el polinomio $P_3(x)$ pasa por el conjunto de puntos dado:

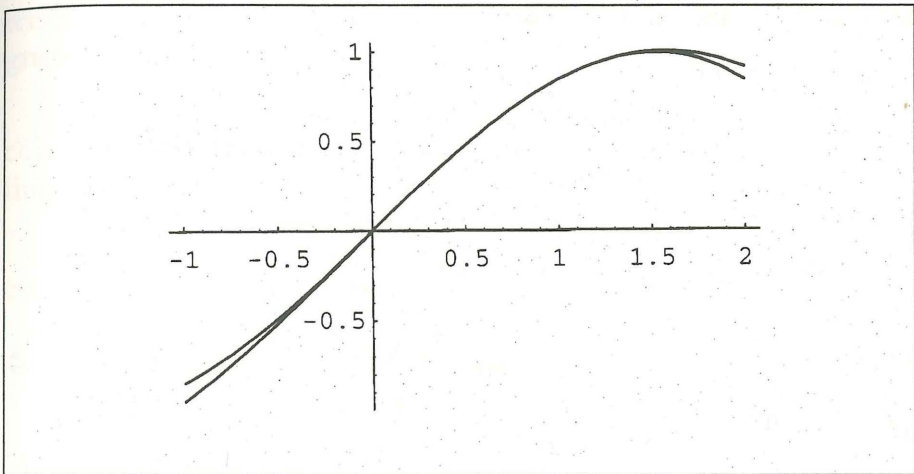


Gráfica 1

Si se sabe que estos puntos satisfacen a la función:

$$y = f(x) = \text{sen } x$$

a través de la siguiente gráfica podemos apreciar como el polinomio $P_3(x)$ que hallamos anteriormente se aproxima a la función $f(x)$:



Gráfica 2

Prácticamente en el intervalo $[0.1, 1.3]$ se superponen una sobre la otra, a no ser en los extremos del intervalo $[-1, 2]$ en que se alejan un poco una de la otra.

Otro de los métodos de *Aproximación de Funciones* que trataremos es el de *Ajuste de Datos* a través del *Método de Mínimos Cuadrados*.

4.1 Aproximación Polinómica

A través de los polinomios, hay distintas formas de aproximar una función, puesto que son funciones de fácil manejo computacional.

Una de ellas, es aproximar una función mediante un polinomio de Taylor.

Se sabe que el polinomio de Taylor se define en la forma:

$$P_n(x) = f(x_0) + \frac{f'(x_0)}{1!} (x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$$

y se sabe también que: $f(x) = P_n(x) + R_n(x)$

Donde el resto es:

$$R_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} (x-x_0)^{n+1} \quad (\text{Teorema de Taylor}).$$

Esto permite realizar la aproximación:

$$f(x) \approx P_n(x)$$

para los valores cercanos a x_0 .

Se puede observar, que toda la información que reúne $P_n(x)$ está concentrada en el punto $x = x_0$, es decir:

$$\begin{aligned} P_n(x_0) &= f(x_0) \\ P'_n(x_0) &= f'(x_0) \\ &\dots \\ P_n^{(n)}(x_0) &= f^{(n)}(x_0), \end{aligned}$$

$(n+1)$ -condiciones para construir $P_n(x)$, polinomio de grado menor o igual a n , usando valores de $f(x)$ y de sus derivadas hasta orden n en un sólo punto x_0 . Es por esta razón que este punto se conoce como punto de contacto, de orden n .

El resto $R_n(x)$ debe cumplir que $\lim_{n \rightarrow \infty} R_n(x) = 0$, lo que implica que la aproximación $f(x) \approx P_n(x)$, tiene un error razonablemente pequeño, sólo si x está cercano a x_0 .

Sin embargo, con frecuencia en aplicaciones se tiene información sobre $f(x)$, a través de sus valores discretos en puntos distintos, no necesariamente todos cercanos entre sí.

Suponiendo que tenemos los valores discretos de

$$f(x_j) \text{ con } x = x_j \in [a, b], \text{ para } j = 0, 1, \dots, n.$$

donde los puntos x_j son distintos entre sí; y si se cumple que:

$$P_n(x_j) = f(x_j), \quad \forall j = 0, 1, 2, 3, \dots, n$$

tendríamos $(n+1)$ -condiciones para construir el polinomio, $P_n(x)$, de grado menor o igual a n .

El polinomio $P_n(x)$ con las condiciones descritas anteriormente se conoce como el polinomio de interpolación de $f(x)$, y los puntos distintos entre sí,

$$x_j \in [a, b], \text{ con } j = 0, 1, \dots, n,$$

como los puntos de interpolación o nodos.

4.2 Aproximación de Funciones Usando el Polinomio de Interpolación de Lagrange

Teorema 1:

Sean $x_j \in [a, b]$, $j = 0, 1, \dots, n$ puntos distintos. Entonces existe y es único el polinomio de grado menor o igual a n , $P_n(x)$ que satisfice:

$$f(x_j) = P_n(x_j), \quad j = 0, 1, 2, \dots, n.$$

Demostración:

Para probar la existencia de $P_n(x)$, construiremos primero las funciones:

$$L_k(x) = \frac{(x-x_0)\dots(x-x_{k-1})(x-x_{k+1})\dots(x-x_n)}{(x_k-x_0)\dots(x_k-x_{k-1})(x_k-x_{k+1})\dots(x_k-x_n)}, \quad k = 0, 1, \dots, n$$

Se puede observar que para $j = 0, 1, 2, \dots, n$, se cumple:

$$L_k(x_j) = \begin{cases} 1, & \text{si } k=j \\ 0, & \text{si } k \neq j \end{cases};$$

$$\text{entonces } P_n(x) = \sum_{k=0}^n f(x_k)L_k(x)$$

vemos que $P_n(x)$, realmente es un polinomio de grado menor o igual a n , ya que es la suma de las funciones $L_k(x)$, y cada una de estas funciones es un polinomio de grado menor o igual a n .

Este polinomio se conoce con el nombre del polinomio de interpolación de Lagrange.

Para probar que es único, supongamos de que si $Q_n(x)$ es otro polinomio distinto a $P_n(x)$, con las mismas características; esto es, $f(x_j) = Q_n(x_j)$, entonces $h(x) = P_n(x) - Q_n(x)$, se anula en los $(n+1)$ -puntos: x_0, x_1, \dots, x_n , lo que es imposible ya que $h(x)$ es un polinomio de grado menor o igual a n . \square

Aplicaciones:

Ejemplo 1:

Sean:

$$\begin{aligned}x_0 &= 1.5; x_1 = 2 \\ f(x_0) &= 0.405465; f(x_1) = 0.693147.\end{aligned}$$

Con 2 puntos de interpolación se puede construir el polinomio de Lagrange de grado menor o igual a 1:

$$P_1(x) = f(x_0) L_0(x) + f(x_1) L_1(x)$$

donde:

$$L_0(x) = \frac{x - x_1}{x_0 - x_1} = \frac{x - 2}{1.5 - 2} = -2(x - 2)$$

$$L_1(x) = \frac{x - x_0}{x_1 - x_0} = \frac{x - 1.5}{2 - 1.5} = 2(x - 1.5)$$

entonces

$$\begin{aligned}P_1(x) &= 0.405465(-2)(x-2) + 2(0.693147)(x-1.5) \\ &= 0.575364x - 0.457581\end{aligned}$$

Ejemplo 2:

Sean

$$x_0 = 1.5, x_1 = 2, x_2 = 3$$

con

$$f(x_0) = 0.405465, f(x_1) = 0.693147, f(x_2) = 1.0986123.$$

Con tres puntos se puede construir el polinomio de Lagrange de grado menor o igual a 2:

$$P_2(x) = f(x_0) L_0(x) + f(x_1) L_1(x) + f(x_2) L_2(x),$$

donde:

$$L_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - 2)(x - 3)}{(-0.5)(-1.5)} = \frac{4}{3}(x - 2)(x - 3)$$

$$L_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - 1.5)(x - 3)}{(0.5)(-1)} = -2(x - 1.5)(x - 3)$$

$$L_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x - 1.5)(x - 2)}{(1.5)(1)} = \frac{2}{3}(x - 1.5)(x - 2)$$

entonces:

$$\begin{aligned} P_2(x) &= (0.405465)(4/3)(x - 2)(x - 3) - 2(0.693147)(x - 1.5)(x - 3) + \\ &\quad (2/3)(1.0986123)(x - 1.5)(x - 2) = \\ &= -0.1132658 x^2 + 0.9717943 x - 0.7973784. \end{aligned}$$

Comentario

Si comparamos los 2 ejemplos dados podemos observar que al agregar un punto de interpolación al conjunto de puntos dados, las funciones $L_k(x)$ deben construirse nuevamente.

Teorema 2: (Teorema del Error del Polinomio de Lagrange)

Sean x_0, x_1, \dots, x_n , $(n+1)$ -puntos distintos en el intervalo $[a, b]$ y $f \in C_{[a, b]}^{(n+1)}$ entonces

$$f(x) = P_n(x) + R_n(x),$$

donde

$$P_n(x) = \sum_{k=0}^n f(x_k) L_k(x)$$

polinomio de interpolación de Lagrange y

$$R_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \prod_{j=0}^n (x - x_j),$$

con $\xi(x) \in]a, b[$ para cada $x \in [a, b]$.

Comentario:

Si comparamos este Teorema con el Teorema de Taylor vemos que los restos $R_n(x)$ de ambos casos se expresan en forma parecida, sólo que el resto del polinomio de interpolación recoge la información de todos los puntos de interpolación, a través del término:

$$\prod_{j=0}^n (x - x_j),$$

mientras que el resto en el caso del polinomio de Taylor usa el factor

$$(x - x_0)^{n+1}.$$

Así, para que haya convergencia en ambos casos se debe cumplir que $\lim_{n \rightarrow \infty} R_n(x) = 0$. Esto implica en el caso de la aproximación polinómica a través del polinomio de Taylor, que x debería estar cercano a x_0 , y en el caso de la aproximación de la función $f(x)$ por el polinomio de Lagrange, que x debería estar cercano a algún punto de interpolación, es decir para $x \in [a, b]$.

Demostración:

Debemos probar que $\forall x \in [a, b]$,

$$f(x) - P_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} (x - x_0) \dots (x - x_n) = R_n(x)$$

Si $x = x_j$, $j = 0, 1, 2, \dots, n$ es decir es uno de los puntos de interpolación, entonces $R_n(x) = 0$ y $f(x_j) = P_n(x_j)$, lo que es cierto.

Sea $x \in [a, b]$, pero $x \neq x_j$, $j = 0, 1, 2, \dots, n$.

Construiremos una función:

$$F(t) = f(t) - P_n(t) - C \prod_{i=0}^n (t - x_i)$$

donde

$$C = \frac{f(x) - P_n(x)}{\prod_{j=0}^n (x - x_j)}$$

Entonces si $t = x$:

$$F(x) = f(x) - P_n(x) - C \prod_{j=0}^n (x - x_j) = 0,$$

lo mismo para $t = x_0, x_1, x_2, \dots, x_n$, $F(t) = 0$; es decir, la función $F(t)$ tiene al menos $(n+2)$ -ceros distintos en $[a, b]$.

Aplicando el Teorema de Rolle, podemos afirmar que:

- la 1ra. derivada $F'(t)$ debe tener por lo menos $(n + 1)$ -ceros en el menor de los intervalos que contenga a x y a los x_j ,
- la 2da. Derivada $F''(t)$ debe tener no menos de n ceros y finalmente,
- la derivada de orden $(n + 1)$, $F^{(n+1)}(t)$ debe tener por lo menos un cero.

Sea $\xi(x)$ tal cero.

Entonces, derivando $(n + 1)$ veces:

$$F^{(n+1)}(t) = f^{(n+1)}(t) - P_n^{(n+1)}(t) - C \frac{d^{n+1}}{dt^{n+1}} \prod_{j=0}^n (t - x_j),$$

tomando $t = \xi(x)$, se tiene

$$0 = F^{(n+1)}(\xi(x)) = f^{(n+1)}(\xi(x)) - C(n+1)!$$

y de aquí:

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

y como

$$f(x) - P_n(x) = C \prod_{j=0}^n (x - x_j),$$

se tiene que:

$$f(x) - P_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \prod_{j=0}^n (x-x_j)$$

lo que se quería demostrar. \square

Casos Particulares:

Caso 1: Interpolación Lineal

Está definida por 2 puntos de interpolación: x_0, x_1 .

Entonces

$$P_2(x) = \frac{x-x_1}{x_0-x_1} f(x_0) + \frac{x-x_0}{x_1-x_0} f(x_1) \quad y$$

$$R_2(x) = f(x) - P_2(x) = \frac{f''(\xi(x))}{2!} (x-x_0)(x-x_1) \leq \frac{M_2}{2} \max_{[x_0, x_1]} \left\{ (x-x_0)(x-x_1) \right\}$$

donde

$$|f''(x)| \leq M_2, \quad \forall x \in]a, b[\text{ y } x \in]x_0, x_1[.$$

Y además se puede ver que la función:

$$|(x-x_0)(x-x_1)|/2$$

logra su máximo valor en $x = (x_0 + x_1)/2$ y toma el valor $(x_1 - x_0)^2/8$.
entonces

$$|R_2(x)| = |f(x) - P_2(x)| \leq \frac{M_2}{8} (x_1 - x_0)^2.$$

Ejemplo 3:

Sean:

$$f(x) = \text{sen } x \text{ y } x_0, x_1, \text{ dos valores sucesivos}$$

de la tabla de $f(x)$ con $h = x_1 - x_0$, entonces usando la interpolación lineal, el error estará acotado por $h^2/8$, ya que $M_2 = \max |f''(x)| = 1$, en este caso.

Caso 2: Interpolación Cuadrática

Para la interpolación cuadrática se usan 3 puntos: x_0, x_1, x_2 .

Entonces

$$P_3(x) = f(x_0) \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + f(x_1) \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + f(x_2) \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}$$

y

$$R_3(x) = |f(x) - P_3(x)| \leq \frac{M_3}{3!} \max \left| \prod_{i=0}^2 (x - x_i) \right|,$$

$$\text{en donde } M_3 = \max_{x \in [x_0, x_2]} |f^{(3)}(x)|$$

El teorema de Lagrange nos permite aproximar $f(x)$ para algún $x \in [a, b]$, tal que $x \neq x_j, j = 0, 1, \dots, n$.

Usaremos los polinomios construidos en los **Ejemplos 1 y 2** para estimar $f(1.75)$:

Ejemplo 4:

Con 2 puntos de interpolación:

$$x_0 = 1.5, x_1 = 2$$

se construyó el polinomio:

$$P_1(x) = 0.575364x - 0.457581$$

Entonces para: $x = 1.75$

$$P_1(1.75) = 0.549306$$

Por lo tanto, el valor estimado para $f(1.75)$ es:

$$f(1.75) \approx 0.549306$$

Los valores $f(x_0), f(x_1)$ son valores discretos de $f(x) = \ln x$, entonces

$$E = \text{Error} = |f(1.75) - P_1(1.75)| \leq \frac{M_2}{8} (x_1 - x_0)^2$$

donde:

$$M_2 = \max_{[1.5, 2]} |f''(x)| \text{ y como } f'(x) = \frac{1}{x}$$

$$|f''(x)| = \left| -\frac{1}{x^2} \right| \leq \frac{1}{(1.5)^2} < 0.5$$

se tiene

$$|f(1.75) - P_1(1.75)| = 0.0103 < \frac{0.5(0.5)^2}{8} = 0.015625 < 5 \times 10^{-2}$$

por lo tanto en nuestra aproximación $f(1.75) \approx 0.549306$ podemos esperar tener dos cifras exactas lo que se confirma al comparar el valor aproximado de $f(1.75)$ con el valor exacto de $\ln(1.75) = 0.5596157879$.

Ejemplo 5:

Con 3 puntos de interpolación:

$$x_0 = 1.5, x_1 = 2, x_2 = 3$$

$$P_2(x) = -0.1132658x^2 + 0.9717943x - 0.7973784$$

entonces

$$P_2(1.75) = 0.556385113 \approx f(1.75)$$

Estimando el error:

$$E = |f(1.75) - P_2(1.75)| \leq \frac{M_3}{3!} (1.75 - 1.5)(1.75 - 2)(1.75 - 3)$$

$$M_3 = \max_{[1.5, 3]} |f^{(3)}(x)|$$

y como

$$|f^{(3)}(x)| = \left| \frac{2}{x^3} \right| \leq \frac{2}{(1.5)^3}$$

se tiene

$$E \leq \left| \frac{(0.25)^2 1.25}{(1.5)^3 3} \right| \approx 0.00771 < 5 \times 10^{-2}$$

lo que nos asegura 2 dígitos exactos.

4.3 Polinomio de Interpolación de Newton con Diferencias Divididas

Si se quisiera agregar un punto de interpolación al conjunto de puntos dados, hay que proceder a calcular las funciones $L_k(x)$ nuevamente; es decir el conocimiento de las funciones $L_k(x)$ para un conjunto de nodos, no nos ayuda a construir las funciones $L_k(x)$ para un conjunto de nodos ampliado por uno o más puntos de interpolación. Esto es una gran desventaja para los polinomios de interpolación de Lagrange. Sin embargo se puede encontrar una forma más eficiente, para un polinomio de interpolación basado en un conjunto de nodos ampliado.

Veamos:

- a. Si tenemos dos puntos de interpolación: x_0, x_1 , el polinomio de grado 1, puede presentarse de la forma:

$$P_1(x) = f(x_0)L_0(x) + f(x_1)L_1(x) =$$

$$= f(x_0) \frac{x - x_1}{x_0 - x_1} + f(x_1) \frac{x - x_0}{x_1 - x_0} = f(x_0) + \left[\frac{f(x_1) - f(x_0)}{x_1 - x_0} \right] (x - x_0)$$

- b. En caso de 3 puntos de interpolación: x_0, x_1, x_2 , el polinomio $P_2(x)$ puede presentarse en la forma:

$$P_2(x) = f(x_0)L_0(x) + f(x_1)L_1(x) + f(x_2)L_2(x) =$$

$$= f(x_0) \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + f(x_1) \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + f(x_2) \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} =$$

$$= f(x_0) + \left[\frac{f(x_1) - f(x_0)}{x_1 - x_0} \right] (x - x_0) + \left[\frac{\left[\frac{f(x_1) - f(x_0)}{x_1 - x_0} \right] - \left[\frac{f(x_2) - f(x_1)}{x_2 - x_1} \right]}{x_0 - x_2} \right] (x - x_0)(x - x_1)$$

Observando $P_1(x)$ y $P_2(x)$, podemos notar que $P_2(x) = P_1(x) + Q_2(x)$, donde:

$$Q_2(x) = \left[\frac{\left[\frac{f(x_1) - f(x_0)}{x_1 - x_0} \right] - \left[\frac{f(x_2) - f(x_1)}{x_2 - x_1} \right]}{x_0 - x_2} \right] (x - x_0)(x - x_1)$$

Además de que el 2° sumando de $P_1(x)$ y $Q_2(x)$ tienen la misma estructura, a través del concepto de diferencias divididas, que definiremos a continuación.

Diferencias Divididas

Definición:

- a. Dado 2 puntos distintos: x_i, x_{i+1} , se define la diferencia dividida de $f(x)$ de **orden 1** como:

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

Así para x_0, x_1 se tiene,

$$f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

- b. Dado 3 puntos distintos: x_i, x_{i+1}, x_{i+2} .

Se define la diferencia dividida de **orden 2** como la diferencia dividida de orden 1 de las diferencias divididas de orden 1; esto es:

$$f[x_i, x_{i+1}, x_{i+2}] = \frac{f[x_{i+1}, x_{i+2}] - f[x_i, x_{i+1}]}{x_{i+2} - x_i}$$

Así para x_0, x_1, x_2 , se tiene:

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$$

- c. En general, dado: $x_i, x_{i+1}, x_{i+2}, \dots, x_{i+k}$, k puntos distintos. Se define la diferencia dividida de **orden k** , como la diferencia dividida de orden 1 de las diferencias divididas de orden $(k-1)$, como:

$$f[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{f[x_{i+1}, \dots, x_{i+k}] - f[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i}$$

Así para x_0, x_1, \dots, x_n , $(n+1)$ -puntos distintos se tiene:

$$f[x_0, x_1, \dots, x_n] = \frac{f[x_1, x_2, \dots, x_n] - f[x_0, x_1, \dots, x_{n-1}]}{x_n - x_0}$$

Se puede observar entonces, que:

$$\begin{aligned} P_1(x) &= f(x_0) + f[x_0, x_1](x - x_0), \\ P_2(x) &= P_1(x) + f[x_0, x_1, x_2](x - x_0)(x - x_1) = \\ &= f(x_0) + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \end{aligned}$$

y en general, para $(n+1)$ -puntos de interpolación:

$$P_n(x) = P_{n-1}(x) + f[x_0, x_1, \dots, x_n] \prod_{j=0}^{n-1} (x - x_j);$$

esto es,

$$P_n(x) = f(x_0) + f[x_0, x_1] (x - x_0) + f[x_0, x_1, x_2] (x - x_0) (x - x_1) + \dots + f[x_0, x_1, \dots, x_n] \prod_{j=0}^{n-1} (x - x_j)$$

o abreviando $P_n(x) = \sum_{k=0}^n f[x_0, x_1, \dots, x_k] \prod_{j=0}^{k-1} (x - x_j).$

Esta forma del polinomio de interpolación se conoce con el nombre de Polinomio de Interpolación de Newton con Diferencias Divididas, y su construcción se simplifica mucho, a partir de la tabla de Diferencias Divididas:

Tabla de Diferencias Divididas

<i>i</i>	x_i	$f(x_i)$	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$	$f[x_i, x_{i+1}, x_{i+2}, x_{i+3}]$
0	x_0	$f(x_0)$			
			$f[x_0, x_1]$		
1	x_1	$f(x_1)$		$f[x_0, x_1, x_2]$	
			$f[x_1, x_2]$		$f[x_0, x_1, x_2, x_3]$
2	x_2	$f(x_2)$		$f[x_1, x_2, x_3]$	
			$f[x_2, x_3]$		
3	x_3	$f(x_3)$			
.....

Tabla 2

Comentario 1:

Se puede ver que el polinomio de Newton se construye usando los elementos situados en diagonal hacia abajo, a partir de $f(x_0)$ de la tabla.

Ejemplo 1:

Suponga que

$$f(x) = \int_0^x e^{-t^3} dt$$

en la siguiente tabla se muestran algunos de sus valores:

i	x_i	$f(x_i)$
0	0	0
1	0.1	0.09997500714
2	0.2	0.1996009126
3	0.3	0.2979905235

Tabla 3

Hallar $f(0.25)$ aproximadamente.

Solución:

El Polinomio de Newton de grado tres:

$$P_3(x) = \sum_{i=0}^3 f[x_0, x_1, \dots, x_i] \prod_{j=0}^{i-1} (x - x_j)$$

Donde las diferencias divididas son:

$$f[x_0, x_1] = \frac{f(x_0) - f(x_1)}{x_0 - x_1} = 0.9997500714$$

$$f[x_1, x_2] = \frac{f(x_1) - f(x_2)}{x_1 - x_2} = 0.9962590546$$

$$f[x_2, x_3] = \frac{f(x_2) - f(x_3)}{x_2 - x_3} = 0.983896109$$

$$f[x_0, x_1, x_2] = \frac{f[x_0, x_1] - f[x_1, x_2]}{x_0 - x_2} = -0.017455084$$

$$f[x_1, x_2, x_3] = \frac{f[x_1, x_2] - f[x_2, x_3]}{x_1 - x_3} = -0.061814728$$

$$f[x_0, x_1, x_2, x_3] = \frac{f[x_0, x_1, x_2] - f[x_1, x_2, x_3]}{x_0 - x_3} = -0.14786548$$

⇒

$$\begin{aligned} P_3(x) &= \sum_{i=0}^3 f[x_0, \dots, x_i] \prod_{j=0}^{i-1} (x - x_j) = \\ &= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \\ &\quad + f[x_0, x_1, x_2, x_3](x - x_0)(x - x_1)(x - x_2) = \end{aligned}$$

$$= 0 + 0.999750x + (-0.017455)x(x-0.1) + (-0.14786548)x(x-0.1)(x-0.2)$$

$$P_3(0.25) = 0.2499375179 - 6.5456565 \times 10^{-4} - 2.77247775 \times 10^{-4}$$

$$P_3(0.25) = 0.2490057044$$

El valor exacto de la integral es:

$$f(0.25) = \int_0^{0.25} e^{-t^3} dt = 0.2490277813$$

Por lo tanto el error es:

$$\text{Error} = |f(0.25) - P_3(0.25)| = 2.20769 \times 10^{-5}$$

Se logra una precisión de 4 decimales exactos.

A continuación se muestra la tabla de diferencias divididas:

Tabla de diferencias divididas

i	x_i	$f[x_i]$	$f[.]$	$f[. .]$	$f[. . .]$
0	0	0			
			0.9997500714		
1	0.1	0.09997500714		-0.017455084	
			0.9962590546		-0.14786548
2	0.2	0.1996009126		-0.061814728	
			0.983896109		
3	0.3	0.2979905235			

Tabla 4

Obteniéndose el polinomio $P_3(x)$:

$$P_3(x) = 0.9985382702x + 0.02690456x^2 - 0.14786548x^3$$

Teorema:

Si $f(x) \in C_{[a, b]}^n$, con $x_i \neq x_j \in [a, b]$, para $i, j = 0, 1, 2, \dots, n$, entonces

$$\text{existe un } \xi \in]a, b[\text{ tal que } f[x_0, x_1, \dots, x_n] = \frac{f^{(n)}(\xi)}{n!(\xi)}$$

Demostración:

Si $k = 1$ tenemos

$$f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{f'(\xi)}{1!}$$

por el Teorema del Valor Medio.

En caso general definimos la función $L_n(x) = f(x) - P_n(x)$

Como para $x = x_j$, con $j = 0, 1, \dots, n$; $L_n(x_j) = 0$, entonces

$L_n(x)$ tiene $(n+1)$ -ceros distintos x_0, \dots, x_n en $]a, b[$.

Aplicando el Teorema de Rolle podemos afirmar que

$$L_n'(x) = f'(x) - P_n'(x)$$

tiene por lo menos n ceros en $]a, b[$. Luego

$$L_n''(x) = f''(x) - P_n''(x)$$

tiene por lo menos $(n-1)$ ceros en $] a, b [$. y así sucesivamente, la n -ésima derivada $L_n^{(n)}(x)$ tiene por lo menos un cero en $] a, b [$.

Sea ξ tal cero: $L_n^{(n)}(\xi) = 0$ y como

$$0 = L_n^{(n)}(\xi) = f^{(n)}(\xi) - P_n^{(n)}(\xi) \quad y$$

$$P_n(x) = f_0 + f[x_0, x_1](x-x_0) + \dots + f[x_0, x_1, \dots, x_n](x-x_0)(x-x_1)\dots(x-x_{n-1})$$

Derivando n veces y reemplazando luego $x = \xi$ se tiene

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

Por lo tanto

$$f[x_0, \dots, x_n] = f^{(n)}(\xi) / n!. \quad \square$$

Comentario 2:

Como

$$f(x) = P_n(x) + R_n(x), \text{ con } R_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \prod_{j=0}^n (x-x_j)$$

Entonces,

$$R_n(x) \approx f[x_0, x_1, \dots, x_{n+1}] \prod_{j=0}^n (x-x_j);$$

es decir para algún $x_{n+1} \in] a, b [$, el resto es aproximadamente el término que sigue al pasar del polinomio $P_n(x)$ al polinomio $P_{n+1}(x)$, que en la práctica resulta muy útil, ya que permite determinar aproximadamente el error del polinomio de interpolación $P_n(x)$.

Aplicaciones:

Ejemplo 2:

Sea la función:

$$f(x) = \text{sen } x$$

En la siguiente tabla se muestran algunos de sus valores:

i	0	1	2	3
x_i	0.1	0.5	0.9	1.3
$f(x_i)$	0.099833	0.479425539	0.7833269	0.96355818

Tabla 5

Estimar $f(0.3) = \text{sen}(0.3)$, por medio de un polinomio de grado 2.

Solución:

El polinomio de Newton de grado dos estará dado por:

$$P_2(x) = \sum_{i=0}^2 f[x_0, \dots, x_i] \prod_{j=0}^{i-1} (x - x_j)$$

Hallaremos las diferencias divididas que intervienen en la construcción del polinomio:

$$f[x_0] = f(x_0) = 0.099833$$

$$f[x_0, x_1] = (f[x_1] - f[x_0]) / (x_1 - x_0) = 0.948981$$

$$f[x_1, x_2] = (f[x_1] - f[x_2]) / (x_1 - x_2) = 0.759753$$

$$f[x_2, x_3] = (f[x_2] - f[x_3]) / (x_2 - x_3) = 0.4505781895$$

$$f[x_0, x_1, x_2] = (f[x_0, x_1] - f[x_1, x_2]) / (x_0 - x_2) = -0.2365349312$$

$$f[x_1, x_2, x_3] = (f[x_1, x_2] - f[x_2, x_3]) / (x_1 - x_3) = -0.3864685131$$

$$f[x_0, x_1, x_2, x_3] = (f[x_0, x_1, x_2] - f[x_1, x_2, x_3]) / (x_0 - x_3) = -0.1249450599$$

Por lo tanto el polinomio de Newton es:

$$P_2(x) = \sum_{i=0}^2 f[x_0, \dots, x_i] \prod_{j=0}^{i-1} (x - x_j) = 0.099833 + 0.948981 (x - 0.1) + (-0.2365349312) (x - 0.1) (x - 0.5)$$

La tabla de diferencias divididas será:

Tabla de Diferencias Divididas

i	x_i	$f[x_i]$	$f[,]$	$f[, ,]$	$f[, , ,]$
0	0.1	0.099833			
			0.948981		
1	0.5	0.47942553		-0.2365349312	
			0.759753		-0.1249450599
2	0.9	0.7833269		-0.3864685131	
			0.4505781895		
3	1.3	0.96355818			

Tabla 6

El valor estimado es $f(0.3) \approx P_2(0.3) = 0.2990906668$, y el valor exacto es $f(0.3) = 0.2955202067$.

El error aproximado es:

$$Error \approx f[x_0, x_1, x_2, x_3] \prod_{j=0}^2 (x - x_j) = 2.998681438 \times 10^{-3} < 5 \times 10^{-3}$$

Ejemplo 3 (Método de Muller)

Este método es una aplicación del polinomio de Newton con diferencias divididas. Lo emplearemos para construir un algoritmo que nos permita encontrar las raíces de una función arbitraria $f(x)$, definida en un intervalo $[a, b]$.

El método de Muller emplea tres aproximaciones diferentes de una raíz de $f(x) = 0$:

$$x_{i-2}, x_{i-1}, x_i$$

que nos servirán para construir un polinomio de Newton de grado 2, de la forma:

$$P_2(x) = f_{i-2} + f[x_{i-2}, x_{i-1}](x-x_{i-2}) + f[x_{i-2}, x_{i-1}, x_i](x - x_{i-2})(x - x_{i-1})$$

Esta ecuación cuadrática es la única parábola que pasa por los tres puntos dados. Luego, hallaremos a los ceros de la parábola, de los cuales

elegiremos al más cercano a las aproximaciones dadas, y este pasará a ser la siguiente aproximación x_{i+1} de la raíz.

En la forma más acostumbrada de escribir polinomios obtenemos que:

$$P_2(x) = a_0 + a_1 x + a_2 x^2,$$

con,

$$a_2 = f[x_{i-2}, x_{i-1}, x_i]$$

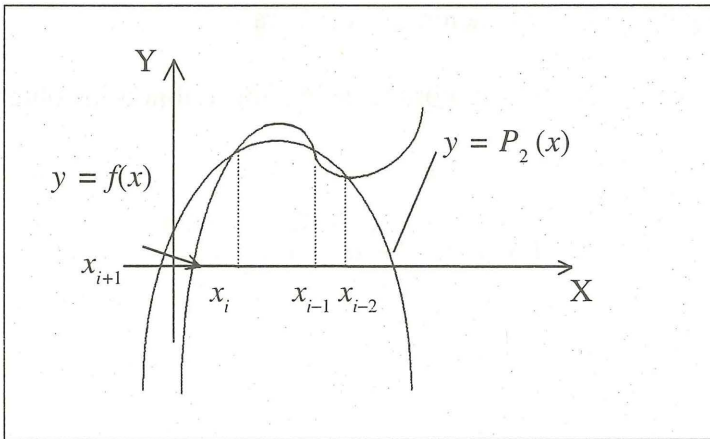
$$a_1 = f[x_{i-2}, x_{i-1}] - (x_{i-2} + x_{i-1})a_2$$

$$a_0 = f_{i-2} - x_{i-2}(f[x_{i-2}, x_{i-1}] - x_{i-1}a_2)$$

Una vez calculados los números a_0 , a_1 , a_2 , se pueden determinar las raíces de $P_2(x)$ a partir de la fórmula cuadrática, escrita en la forma:

$$x = \frac{2a_0}{-a_1 \pm (a_1^2 - 4a_0a_2)^{1/2}} \dots (1)$$

El signo que precede al radical en la ecuación anterior se selecciona de modo que el denominador tenga máxima magnitud, y la raíz correspondiente se toma como la aproximación siguiente, x_{i+1} . La razón para escribir la fórmula cuadrática en esta forma, como se explicó anteriormente, es obtener un máximo de exactitud en la fórmula. El proceso se repite entonces, usando x_{i-1} , x_i , x_{i+1} como las tres aproximaciones iniciales. Si las raíces obtenidas a partir de la ecuación (1) son reales, la situación se podría representar por medio de una gráfica:



Gráfica 3

Observar que aunque la raíz que se busca sea real, podemos hallar aproximaciones complejas. Sin embargo, en tales casos la componente imaginaria será normalmente tan pequeña en magnitud que puede despreciarse.

Este método puede usarse para hallar cualquier número prescrito de ceros reales o complejos, de una función arbitraria. El método es iterativo, converge casi cuadráticamente en la vecindad de una raíz, no requiere el cálculo de la derivada de la función, y obtiene tanto raíces reales como complejas aún cuando estas raíces no sean simples.

Además, el método es global en el sentido de que el usuario necesita suplir una aproximación inicial.

Aplicación:

Dado el siguiente polinomio:

$$f(x) = x^5 - x + 3 = 0.$$

Hallar la raíz compleja y su conjugada, correspondiente a las tres primeras aproximaciones:

$$x_0^{(0)} = 1, x_1^{(0)} = 1.2, x_2^{(0)} = 1.4$$

y aplicando el algoritmo nos lleva a la siguiente aproximación, que la denotaremos como:

$$x_2^{(1)} = 0.91616437 + 0.4052215 i,$$

entonces ahora, las nuevas aproximaciones son:

$$x_0^{(1)} = 1.2, x_1^{(1)} = 1.4, x_2^{(1)} = 0.91616437 + 0.4052215 i,$$

y así sucesivamente, en la tabla a continuación se muestran las 6 primeras iteraciones:

i	$x_2^{(i)}$
0	1.4
1	0.91616437 + 0.4052215i
2	0.91533025 + 0.5492393i
3	0.958844 + 0.628673i
4	0.97949731 + 0.62560521i
5	0.97906034 + 0.62521876i
6	0.97906187 + 0.62521908i

Tabla 7

4.4 Interpolación de una Función Basada en Puntos Igualmente Espaciados

Consiste en determinar un polinomio que interpole a $f(x)$ en puntos igualmente espaciados, que forman una sucesión creciente. Esto es:

$$x_0, x_1, x_2, \dots, x_N \in [a, b]$$

$$x_0 = a, \quad x_1 = a + h, \quad x_2 = a + 2h, \quad x_N = b$$

donde:

$$x_i = x_0 + i h, \text{ con } N = \frac{b-a}{h}$$

para $i = 0, 1, \dots, N$.

Entonces:

$$\begin{array}{ccccccccccc} \bullet & & \bullet & & \bullet & & \bullet & & \dots & & \bullet & & \bullet & & \dots & & \bullet \\ \hline a = x_0 & & x_1 & & x_2 & & \dots & & x_k & & x_{k+1} & & \dots & & x_N = b \end{array}$$

Si $x \in [a, b] \Rightarrow \exists s \in \left[0, \frac{b-a}{h}\right]$, tal que $x = x_0 + sh$.

De manera que usaremos la notación:

$$f_s = f(x) = f(x_0 + sh),$$

$$f_i = f(x_i) = f(x_0 + ih)$$

Debe observarse que el cambio lineal de variables transforma al polinomio de grado n de x , en un polinomio de grado n en s .

Recordaremos el concepto de diferencias finitas o progresivas:

La diferencia finita o progresiva de orden 1, se define como:

$$\Delta f_i = f_{i+1} - f_i.$$

De aquí, la diferencia progresiva de orden k se define como diferencia finita de orden 1 de las diferencias finitas de orden $(k-1)$.

Así:

$$\Delta^2 f_i = \Delta(\Delta f_i) = \Delta(f_{i+1} - f_i) = \Delta f_{i+1} - \Delta f_i$$

ya que el operador de diferencias es un operador lineal y por lo tanto:

$$\Delta^2 f_i = f_{i+2} - 2f_{i+1} + f_i.$$

Se puede observar el comportamiento binomial de los coeficientes.

Así, en forma simbólica se expresa:

$$\Delta^2 f_i = [(1 + \Delta) - 1]^2 f_i = \binom{2}{0}(1 + \Delta)^2 f_i - \binom{2}{1}(1 + \Delta)f_i + \binom{2}{2}(1 + \Delta)^0 f_i$$

Este concepto permitirá establecer una relación con las diferencias divididas de la siguiente forma:

Propiedad:

$$\frac{\Delta^i f_k}{h^i i!} = f[x_k, x_{k+1}, \dots, x_{k+i}], \quad \forall i \geq 0$$

Demostración:

Lo haremos por inducción:

Si $i = 0$

$$\Rightarrow f(x_k) = \frac{\Delta^0 f_k}{0! h^0} = f[x_k]$$

Para $i = n$ asumiremos que es válido.

Luego probaremos que es válido para $i = n + 1$.

Consideremos a la $(n+1)$ -ésima diferencia dividida:

$$f[x_k, x_{k+1}, \dots, x_{k+n+1}] = \{f[x_k, \dots, x_{k+n}] - f[x_{k+1}, \dots, x_{k+n+1}]\} / (x_k - x_{k+n+1})$$

Como habíamos asumido que es válida la propiedad para $i = n$, entonces:

$$\begin{aligned} f[x_k, x_{k+1}, \dots, x_{k+n+1}] &= \frac{\Delta^n f_k / h^n n! - \Delta^n f_{k+1} / h^n n!}{-(a + (k+n+1)h) + (a+kh)} = \\ &= \frac{-(\Delta^n f_{k+1} - \Delta^n f_k)}{-(n+1)h} \cdot \frac{1}{h^n n!} = \frac{\Delta^{n+1} f_k}{h^{n+1} (n+1)!} \end{aligned}$$

Lo que prueba la propiedad. \square

Por lo tanto, según el polinomio interpolante de Newton se tendría, para los $(n+1)$ -puntos $x_k, x_{k+1}, \dots, x_{k+n}$ distintos, el polinomio de Newton:

$$\begin{aligned} P_n(x) &= \sum_{i=0}^n f[x_k, x_{k+1}, \dots, x_{k+i}] \prod_{j=0}^{i-1} (x - x_{k+j}) \\ &= \sum_{i=0}^n \frac{\Delta^i f_k}{h^i i!} \cdot \prod_{j=0}^{i-1} (x - x_{k+j}) = \\ &= \sum_{i=0}^n \frac{\Delta^i f_k}{h^i i!} \cdot \prod_{j=0}^{i-1} (x - (a + (k+j)h)) \end{aligned}$$

Expresándolo en términos de s , se tendría:

$$x - x_{k+j} = a + sh - (a + (k+j)h) = (s - k - j)h$$

de modo que:

$$P_n(x) = \sum_{i=0}^n \frac{\Delta^i f_k}{h^i i!} \cdot \prod_{j=0}^{i-1} (s - k - j)h$$

y observando:

$$\begin{aligned} \frac{1}{h^i} \frac{1}{i!} \prod_{j=0}^{i-1} (s - k - j)h &= \frac{h^i}{h^i i!} (s - k - 0)(s - k - 1) \dots (s - k - i + 1) \\ &= \left(\frac{s - k}{1} \right) \left(\frac{s - k - 1}{2} \right) \dots \left(\frac{s - k - (i - 1)}{i} \right) \\ &= \prod_{j=0}^{i-1} \left[\frac{s - k - j}{j + 1} \right] \\ \Rightarrow P_n(x) &= \sum_{i=0}^n \Delta^i f_k \prod_{j=0}^{i-1} \left[\frac{s - k - j}{j + 1} \right] \end{aligned}$$

y esta expresión recibe el nombre de fórmula de Newton de diferencias progresivas.

Además, se pueden usar los coeficientes binomiales para simplificar la expresión:

$$\prod_{j=0}^{i-1} \left[\frac{s-k-j}{j+1} \right] = \frac{(s-k)(s-k-1)\dots(s-k-i+1)}{1 \cdot 2 \cdot 3 \dots i} = \frac{(s-k)!}{(s-k-i)!i!} = \binom{s-k}{i}$$

$$\Rightarrow P_n(x) = \sum_{i=0}^n \Delta^i f_k \binom{s-k}{i}$$

Esto es la forma de diferencia progresiva de Newton para el polinomio de grado menor o igual a n que interpola $f(x)$ en $x_k, x_{k+1}, \dots, x_{k+n}$.

Comentario:

Dado un conjunto de puntos de interpolación $x_{-3}, x_{-2}, x_{-1}, x_0, x_1, x_2, x_3$, se puede construir la tabla de diferencias finitas:

Tabla de Diferencias Finitas

i	x_i	f_i	Δf_i	$\Delta^2 f_i$	$\Delta^3 f_i$
-3	x_{-3}	f_{-3}			
			Δf_{-3}		
-2	x_{-2}	f_{-2}		$\Delta^2 f_{-3}$	
			Δf_{-2}		$\Delta^3 f_{-3}$
1	x_{-1}	f_{-1}		$\Delta^2 f_{-2}$	
			Δf_{-1}		$\Delta^3 f_{-2}$
0	x_0	f_0		$\Delta^2 f_{-1}$	
			Δf_0		$\Delta^3 f_{-1}$
1	x_1	f_1		$\Delta^2 f_0$	
			Δf_1		$\Delta^3 f_0$
2	x_2	f_2		$\Delta^2 f_1$	
			Δf_2		
3	x_3	f_3			

Tabla 8

La que nos será útil para construir el polinomio de Newton $P_n(x)$. Así por ejemplo, si $k = 0$, tenemos:

$$P_3(x) = \sum_{i=0}^3 \Delta^i f_0 \binom{s}{i}$$

en donde se utilizan las diferencias finitas indicadas en la tabla.

Si $k + n > N$, la fórmula anterior sería inútil, ya que está basada en los puntos $x_k, x_{k+1}, x_{k+2}, \dots, x_{k+n}$. Luego si se interpola al final de la tabla, es mejor tener una forma que use puntos de interpolación en orden decreciente de la forma: $x_k, x_{k-1}, \dots, x_{k-n}$. Partiendo nuevamente de la fórmula de Newton, interpolando en $x_{k-n}, x_{k-n+1}, \dots, x_k$, se tiene:

$$\begin{aligned} P_n(x) &= \sum_{i=0}^n f[x_{k-i}, \dots, x_k] \prod_{j=-i+1}^0 (x - x_{k+j}) = \\ &= \sum_{i=0}^n \frac{\Delta^i f_{k-i}}{h^i i!} \prod_{j=-i+1}^0 (x - x_{k+j}) = \\ &= \sum_{i=0}^n \frac{\Delta^i f_{k-i}}{h^i i!} \prod_{j=-i+1}^0 (s - k - j)h = \\ &= \sum_{i=0}^n \frac{\Delta^i f_{k-i}}{h^i i!} (-h)^i \prod_{j=0}^{i-1} (k - s - j) = \sum_{i=0}^n (-1)^i \Delta^i f_{k-i} \binom{k-s}{i} = \\ &= f_k - \binom{k-s}{1} \Delta f_{k-1} + \binom{k-s}{2} \Delta^2 f_{k-2} + \dots + (-1)^n \binom{k-s}{n} \Delta^n f_{k-n} \end{aligned}$$

que recibe el nombre de fórmula de diferencia regresiva de Newton para $P_n(x)$ en $x_k, x_{k-1}, \dots, x_{k-n}$.

Si $k = 0$

$$\Rightarrow P_n(x_0 + sh) = \sum_{i=0}^n (-1)^i \Delta^i f_{-i} \binom{-s}{i}$$

y las diferencias finitas $\Delta^i f_{-i}$ se obtienen de la Tabla de Diferencias Finitas.

El error se obtiene de manera similar al error del polinomio interpolante:

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

Aplicaciones:

Ejemplo 1:

Dada la siguiente tabla de datos:

i	0	1	2	3	4	5	6
x_i	0.1	0.3	0.5	0.7	0.9	1.1	1.3
$y_i = f_i$	0.9975	.97763	.93847	.8812	.80752	.71962	.62009

Tabla 9

- Obtener el polinomio de interpolación de diferencias progresivas, con los datos $i = 0, 1, 2$ y calcular aproximadamente $f(0.2)$, estimando el error.
- Estimar $f(1.2)$ con la forma regresiva de Newton, para los puntos $i = 6, 5, 4$.

Solución:

- La tabla de diferencias finitas es:

i	x_i	f_i	Δf_i	$\Delta^2 f_i$	$\Delta^3 f_i$	$\Delta^4 f_i$	$\Delta^5 f_i$	$\Delta^6 f_i$
0	0.1	0.9975						
			-.01987					
1	0.3	.97763		-.01929				
			-.03916		.00118			
2	0.5	.93847		-.01811		.00052		
			-.05727		.0017		-.00003	
3	0.7	.8812		-.01641		.00049		-.00006
			-.07368		.00219		-.00009	
4	0.9	.80752		-.01422		.0004		
			-.08790		.00259			
5	1.1	.71962		-.01163				
			-.09953					
6	1.3	.62009						

Tabla 10

Trabajando con los puntos: $x_0 = 0.1$, $x_1 = 0.3$; $x_2 = 0.5$, se tiene:

$$\begin{aligned}
 P_2(x) &= \sum_{i=0}^2 \Delta^i f_0 \binom{s}{i} = \Delta^0 f_0 \binom{s}{0} + \Delta^1 f_0 \binom{s}{1} + \Delta^2 f_0 \binom{s}{2} = \\
 &= 0.9975 - 0.01987s - 0.01929 \frac{s(s-1)}{2}
 \end{aligned}$$

ya que:

$$\binom{s}{0} = \frac{s!}{(s-0)!0!} = 1; \quad \binom{s}{1} = \frac{s!}{(s-1)!1!} = s; \quad \binom{s}{2} = \frac{s!}{(s-2)!2!} = \frac{s(s-1)}{2}$$

$$\text{Si } h = 0.2, x = 0.2 \Rightarrow 0.2 = x = x_0 + sh = 0.1 + s(0.2)$$

$$s = 0.5$$

$$\therefore P_2(0.2) = 0.98997625$$

El error estimado será:

$$\text{Error} = E(x) \approx \Delta^{n+1} f_0 \binom{s}{n+1} = 0.00118 \binom{s}{3}$$

$$\binom{s}{3} = \frac{s!}{(s-3)!3!} = \frac{s(s-1)(s-2)}{3!}$$

$$\Delta^3 f_0 = 1.18 \times 10^{-3}$$

$$\Rightarrow \text{Error} = E(0.2) \approx 0.00118 \frac{(0.5)(-0.5)(-1.5)}{6} = 7.375 \times 10^{-5}$$

lo que significa que podemos esperar a lo más 4 cifras exactas.

b. Sean los puntos de interpolación:

$$x_6, x_5, x_4$$

definamos:

$$x_0 = x_6 = 1.3, x_{-1} = x_5 = 1.1, x_{-2} = x_4 = 0.9;$$

entonces el polinomio de forma regresiva será:

$$\begin{aligned} P_2(x) &= \sum_{i=0}^2 (-1)^i \Delta^i f_{-i} \binom{-s}{i} = \\ &= \Delta^0 f_0 \binom{-s}{0} - \Delta f_{-1} \binom{-s}{1} + \Delta^2 f_{-2} \binom{-s}{2} = \\ &= 0.62009 \binom{-s}{0} + 0.09953 \binom{-s}{1} - 0.01163 \binom{-s}{2} \end{aligned}$$

como $x = 1.2 = x_0 + sh = 1.3 + s(0.2) \Rightarrow s = -0.5$

$$\binom{0.5}{0} = 1; \binom{0.5}{1} = 0.5; \binom{0.5}{2} = -1/8 \Rightarrow P_2(1.2) = 0.67130875$$

El error estimado será:

$$E(1.2) \approx 0.00259 \binom{0.5}{3} = 1.61875 \times 10^{-4}$$

Lo que significa que podemos esperar a lo más 3 decimales exactos.

Ejemplo 2:

En la siguiente tabla se muestran los puntos correspondientes a la Función de Distribución Acumulada de la curva Normal Estándar:

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

i	0	1	2	3	4	5	6	7	8	9
x_i	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
$y_i = F(x_i)$.500000	.503989	.507978	.511966	.515953	.519938	.523922	.527903	.531881	.535856

Tabla 11

- a. Estimar el valor de $F(0.042)$, haciendo uso de un polinomio de Newton de forma progresiva con los siguientes puntos de interpolación:
- (1) $i = 3, 4, 5, 6$.
 - (2) $i = 1, 2, 3, 4, 5, 6, 7$
- b. Estimar el error aproximado en cada uno de los casos anteriores.

Solución:

- a. Estimando $F(0.042)$:

- (1) Para $i = 3, 4, 5, 6$.

Con los puntos de interpolación:

$$x_3 = 0.03, x_4 = 0.04, x_5 = 0.05, x_6 = 0.06$$

$$\begin{aligned} P_3(x) &= \sum_{i=0}^3 \Delta^i f_3 \binom{s-3}{i} = \\ &= \Delta^0 f_3 \binom{s-3}{0} + \Delta^1 f_3 \binom{s-3}{1} + \Delta^2 f_3 \binom{s-3}{2} + \Delta^3 f_3 \binom{s-3}{3} \end{aligned}$$

como $x = 0.042$, entonces el valor s , tal que:

$$\begin{aligned} x = 0.042 &= 0 + s h = 0 + s (0.01) \\ \Rightarrow s &= 0.042 / 0.01 = 4.2 \end{aligned}$$

Y calculando los números combinatorios para $P_3(x)$, son:

$$\binom{s-3}{0} = 1; \quad \binom{s-3}{1} = \frac{(s-3)!}{(s-4)!(1!)} = s-3 = 1.2;$$

$$\binom{s-3}{2} = \frac{(s-3)!}{(s-5)!(2!)} = \frac{(s-3)(s-4)}{2} = 0.12;$$

$$\binom{s-3}{3} = \frac{(s-3)!}{(s-6)!(3!)} = \frac{(s-3)(s-4)(s-5)}{3!} = -0.032$$

A continuación se muestra la tabla de Diferencias Finitas:

i	x_i	f_i	Δf_i	$\Delta^2 f_i$	$\Delta^3 f_i$	$\Delta^4 f_i$	$\Delta^5 f_i$	$\Delta^6 f_i$	$\Delta^7 f_i$	$\Delta^8 f_i$	$\Delta^9 f_i$
0	0	0.5									
			3.989×10^{-3}								
1	0.01	.503989		0							
			3.989×10^{-3}		-1×10^{-6}						
2	0.02	.507978		-1×10^{-6}		1×10^{-6}					
			3.988×10^{-3}		0		-2×10^{-6}				
3	0.03	.511966		-1×10^{-6}		-1×10^{-6}		5×10^{-6}			
			3.987×10^{-3}		-1×10^{-6}		3×10^{-6}		-13×10^{-6}		
4	0.04	.515953		-2×10^{-6}		2×10^{-6}		-8×10^{-6}		31×10^{-6}	
			3.985×10^{-3}		1×10^{-6}		-5×10^{-6}		18×10^{-6}		-66×10^{-6}
5	0.05	.519938		-1×10^{-6}		-3×10^{-6}		10×10^{-6}		-35×10^{-6}	
			3.984×10^{-3}		-2×10^{-6}		5×10^{-6}		-17×10^{-6}		
6	0.06	.523922		-3×10^{-6}		2×10^{-6}		-7×10^{-6}			
			3.981×10^{-3}		0		-2×10^{-6}				
7	0.07	.527903		-3×10^{-6}		0					
			3.978×10^{-3}		0						
8	0.08	.531881		-3×10^{-6}							
			3.975×10^{-3}								
9	0.09	.535856									

Tabla 12

Luego

$$P_3(0.042) = 0.511966 - 1.2(3.987 \times 10^{-3}) + 0.12(-2 \times 10^{-6}) - 0.032(1 \times 10^{-6}) = 0.516750128$$

De manera similar estimamos a $x = 0.042$, a través del polinomio de grado 6:

(2) Para $i = 1, 2, 3, 4, 5, 6, 7$.

Redefiniendo los puntos:

$x_0 = x_1; x_1 = x_2; \dots, x_6 = x_7$, por lo tanto el polinomio de grado 6 de Newton será:

$$\begin{aligned}
 P_6(x) &= \sum_{i=0}^6 \Delta^i f_0 \binom{s}{i} = \\
 &= \Delta^0 f_0 \binom{s}{0} + \Delta^1 f_0 \binom{s}{1} + \Delta^2 f_0 \binom{s}{2} + \dots + \Delta^6 f_0 \binom{s}{6}
 \end{aligned}$$

Donde s es tal que: $0.042 = x_0 + sh = 0.01 + s(0.01)$

$$\Rightarrow s = (0.042 - 0.01)/0.01 = 3.2.$$

evaluando los valores de los números combinatorios se tiene:

$$\begin{aligned}
 \binom{s}{0} &= 1; \quad \binom{s}{1} = 3.2; \quad \binom{s}{2} = 3.52; \quad \binom{s}{3} = 1.408; \\
 \binom{s}{4} &= 0.0704; \quad \binom{s}{5} = -0.011264; \quad \binom{s}{6} = 3.3792 \times 10^{-3}
 \end{aligned}$$

por lo tanto $P_6(0.042) = 0.5167501488$

b. Error aproximado:

Para (1):

$$\begin{aligned}
 E(0.042) &\approx \Delta^{n+1} f_3 \binom{s-k}{n+1} = \Delta^4 f_3 \binom{1.2}{4} = \\
 &\approx -3 \times 10^{-6} \frac{(1.2)(1.2-1)(1.2-2)(1.2-3)}{4!} = -4.32 \times 10^{-8}
 \end{aligned}$$

Para (2):

$$\begin{aligned}
 E(0.042) &\approx \Delta^{n+1} f_0 \binom{s}{n+1} = \Delta^7 f_0 \binom{3.2}{7} = \\
 &\approx 18 \times 10^{-6} \frac{(3.2)(3.2-1) \dots (3.2-6)}{7!} = -2.433024 \times 10^{-8}
 \end{aligned}$$

En ambas estimaciones se obtiene una precisión del orden de 10^{-8} .

4.5 Aplicaciones de los Polinomios de Interpolación

En las técnicas del cálculo como la derivación o la integración de una función $f(x)$, nos encontraremos muchas veces frente a situaciones en las cuales no podremos obtener el valor exacto de $f'(x)$, o no podremos

hallar la antiderivada de $f(x)$. Hay funciones que aunque aparecen en términos de funciones elementales como:

$$\operatorname{sen} x / x, e^{\frac{-x^2}{2}}, \sqrt{\cos^2 x + 3\operatorname{sen}^2 x},$$

no podremos encontrar su integral, por medio de los métodos exactos. Más aún en el caso de no existir una fórmula para $f(x)$, y sólo se conocen los pares de puntos tabulados (x_i, y_i) , sería inútil querer calcular a través de los métodos exactos, la derivada o la integral de dicha función. En estos casos, en que los cálculos resultan laboriosos o imposibles, surge la necesidad de emplear los métodos numéricos para hallar una solución aproximada.

Derivación Numérica

Supongamos que x_0, x_1, \dots, x_n son $n + 1$ puntos distintos en el intervalo $[a, b]$ y sea la función $f \in C_{[a, b]}^{n+1}$ entonces por el Teorema del *Error* para el Polinomio de $f(x) \in C_{[a, b]}^{n+1}$ tiene:

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

para alguna $\xi(x)$, en donde $P_n(x)$ es el Polinomio de Lagrange para $f(x)$ basado en los puntos de interpolación: x_0, x_1, \dots, x_n .

$$P_n(x) = \sum_{i=0}^n f(x_i) L_i(x)$$

Derivando esta expresión, tenemos:

$$f'(x) = P'_n(x) + D_x \left[\frac{f^{(n+1)}(\xi(x))}{(n+1)!} \prod_{i=0}^n (x - x_i) \right]$$

$$f'(x) = \sum_{i=0}^n f(x_i) L_i'(x) + D_x \left[\prod_{i=0}^n (x - x_i) \right] \frac{f^{(n+1)}(\xi(x))}{(n+1)!} + D_x \left[\frac{f^{(n+1)}(\xi(x))}{(n+1)!} \right] \prod_{i=0}^n (x - x_i)$$

Se puede expresar el producto $\prod_{i=0}^n (x - x_i)$, como:

$$(x - x_k) \prod_{\substack{i=0 \\ i \neq k}}^n (x - x_i),$$

donde x_k es uno de los puntos de interpolación, y analizando la derivada de este producto en $x = x_k$, tendremos:

$$\begin{aligned} D_x \left[\prod_{i=0}^n (x - x_i) \right] \Big|_{x=x_k} &= D_x [(x - x_k) \prod_{\substack{i=0 \\ i \neq k}}^n (x - x_i)] \Big|_{x=x_k} = \\ &= \prod_{\substack{i=0 \\ i \neq k}}^n (x - x_i) \Big|_{x=x_k} + (x - x_k) D_x \left[\prod_{\substack{i=0 \\ i \neq k}}^n (x - x_i) \right] \Big|_{x=x_k} = \\ &= \prod_{\substack{i=0 \\ i \neq k}}^n (x_k - x_i) \end{aligned}$$

lo que nos lleva a expresar la derivada de $f(x)$, en $x = x_k$ como:

$$\Rightarrow f'(x_k) = \sum_{i=0}^n f(x_i) L_i'(x_k) + \frac{f^{(n+1)}(\xi(x_k))}{(n+1)!} \prod_{\substack{i=0 \\ i \neq k}}^n (x_k - x_i)$$

$$\forall k = 0, 1, 2, \dots, n$$

Se puede estimar a la derivada $f'(x_k)$, a través de la derivada del polinomio; esto es:

$$f'(x_k) \approx P_n'(x_k) = \sum_{i=0}^n f(x_i) L_i'(x_k)$$

y a esta ecuación se le denominará fórmula de $(n+1)$ -puntos para estimar a $f'(x_k)$.

En esta fórmula de $(n+1)$ -puntos para estimar a $f'(x_k)$, se hace uso de una combinación lineal de valores funcionales $f(x_i)$ con $i = 0, 1, 2, \dots, n$.

El error estará dado por:

$$\text{Error} = \frac{f^{(n+1)}(\xi(x_k))}{(n+1)!} \prod_{\substack{i=0 \\ i \neq k}}^n (x_k - x_i)$$

De la fórmula para el error, se observa que se obtendría mayor precisión, si se usara un número mayor de puntos de interpolación; pero aún así el número de evaluaciones funcionales y el crecimiento del error del redondeo no animan a hacerlo.

Veremos el caso particular de tres puntos, que es el que se emplea con mayor frecuencia.

Fórmulas de tres puntos: x_0, x_1, x_2

Sean x_0, x_1, x_2 los puntos de interpolación y $f(x_0), f(x_1)$, y $f(x_2)$, valores funcionales en esos puntos.

Entonces:

$$f'(x_k) \approx P_2'(x_k) = \sum_{i=0}^2 f(x_i) L_i'(x_k); \quad k = 0, 1, 2$$

donde:

$$L_0(x) = \prod_{\substack{j=0 \\ j \neq 0}}^2 \left(\frac{x - x_j}{x_0 - x_j} \right) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}$$

$$\Rightarrow L_0'(x) = (2x - x_1 - x_2) / [(x_0 - x_1)(x_0 - x_2)]$$

$$L_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} \Rightarrow L_1'(x) = \frac{2x - x_0 - x_2}{(x_1 - x_0)(x_1 - x_2)}$$

$$L_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} \Rightarrow L_2'(x) = \frac{2x - x_0 - x_1}{(x_2 - x_0)(x_2 - x_1)}$$

Por lo tanto:

$$\Rightarrow f'(x_k) \approx f(x_0) \left[\frac{2x_k - x_1 - x_2}{(x_0 - x_1)(x_0 - x_2)} \right] + f(x_1) \left[\frac{2x_k - x_0 - x_2}{(x_1 - x_0)(x_1 - x_2)} \right] +$$

$$+ f(x_2) \left[\frac{2x_k - x_0 - x_1}{(x_2 - x_0)(x_2 - x_1)} \right]$$

para $k = 0, 1, 2$

y la expresión para el error estará dada por:

$$\text{Error} = \frac{f^{(3)}(\xi(x_k))}{3!} \prod_{\substack{i=0 \\ i \neq k}}^2 (x_k - x_i)$$

donde $\xi(x_k)$ se ubica entre los puntos x_0, x_1, x_2 .

Supongamos que los puntos están igualmente espaciados con un paso h , esto es:

$$x_0, x_1 = x_0 + h, x_2 = x_0 + 2h,$$

así, por ejemplo para:

$$k = 0,$$

$$f'(x_0) = \frac{1}{2h} [-3f(x_0) + 4f(x_1) - f(x_2)] + \frac{h^2}{3!} f^{(3)}(\xi(x_0)) \dots (1)$$

donde $\xi(x_0) \in [x_0, x_0 + 2h]$

$$k = 1,$$

$$f'(x_1) = \frac{1}{2h} [-f(x_0) + f(x_0 + 2h)] - \frac{h^2}{6} f^{(3)}(\xi(x_1)) \dots (2)$$

donde $\xi(x_1) \in [x_0, x_0 + 2h]$

$$k = 2,$$

$$f'(x_2) = \frac{1}{2h} [f(x_0) - 4f(x_0 + h) + 3f(x_0 + 2h)] + \frac{h^2}{3} f^{(3)}(\xi(x_2)) \dots (3)$$

$$\text{donde } \xi(x_2) \in [x_0, x_0 + 2h]$$

Sustituyendo en (2) x_1 por x_0

y en (3) x_2 por x_0 , entonces se obtiene:

$$f'(x_0) = \frac{1}{2h} [-f(x_0 - h) + f(x_0 + h)] - \frac{h^2}{6} f^{(3)}(\xi(x_0)) \dots (4)$$

$$\xi(x_0) \in [x_0 - h, x_0 + h]$$

$$f'(x_0) = \frac{1}{2h} [f(x_0 - 2h) - 4f(x_0 - h) + 3f(x_0)] + \frac{h^2}{3} f^{(3)}(\xi(x_0)) \dots (5)$$

$$\text{donde } \xi(x_0) \in [x_0 - 2h, x_0]$$

y observar que si en (3) se cambia h por $-h$, se obtiene:

$$f'(x_0) = \frac{1}{2h} [-f(x_0 + 2h) + 4f(x_0 + h) - 3f(x_0)] + \frac{h^2}{3} f^{(3)}(\xi(x_0))$$

esto es, se reduce a la ecuación (1).

Aplicaciones:

Ejemplo 1:

Sea $f(x) = x e^x$. En la siguiente tabla se dan los valores para $f(x)$.

x_i	$f(x_i)$
1.8	10.889365
1.9	12.703199
2	14.778112
2.1	17.148957
2.2	19.855030

Tabla 13

Estimar $f'(2)$, con $h = 0.1$, $h = 0.2$

a. Para $h = 0.1$, empleando la fórmula (4), tenemos:

$$f'(2) \approx (1/(2h))[f(2.1) - f(1.9)] = 22.228790$$

ó empleando la fórmula (1), se tiene:

$$f'(2) \approx (1/(2h))[-3f(2) + 4f(2.1) - f(2.2)] = 22.032310$$

b. Para $h = 0.2$, empleando la fórmula (4), se tiene:

$$f'(2) \approx (1/0.4)[f(2.2) - f(2.18)] = 22.4141625$$

Analizando el error para (b), en el caso cuando $h = 0.2$:

$$|Error| \approx \left| - (h^2/6) f^{(3)}(\xi(x)) \right| \leq \frac{h^2}{6} \max_{[1.8, 2.2]} \{ |f^{(3)}(x)| \}$$

como sus derivadas son:

$$\begin{aligned} f(x) &= xe^x \\ f'(x) &= e^x(x+1) \\ f''(x) &= e^x(x+2) \\ f'''(x) &= e^x(x+3) \end{aligned}$$

entonces

$$\max_{[1.8, 2.2]} \{ |f^{(3)}(x)| \} = \max_{[1.8, 2.2]} \{ e^x(x+3) \} = e^{2.2}(2.2+3) = 46.93$$

Por lo tanto, el $|Error| \leq 0.312867$

El error absoluto es:

$$|f'(2) - 22.414163| = 0.2469947,$$

ya que $f'(2) = 22.1671683$.

Ejemplo 2:

Siendo:

$$f(x) = x^3 e^{-x^2} \operatorname{sen} x,$$

estimar $f'(2.19)$ utilizando las fórmulas de 3 puntos y con $h = 0.01$,

i	0	1	2	3	4
x_i	2.17	2.18	2.19	2.20	2.21
$f(x_i)$	1,132.692388	1,199.5373	1,270.49755	1,345.8372	1425.837769

Tabla 14

a. Empleando la fórmula (1), se tiene:

$$\begin{aligned} f'(2.19) &= \frac{1}{2h} [-3f(x_0) + 4f(x_0 + h) - f(x_0 + 2h)] \\ &= \frac{1}{2(0.01)} [-3f(2.19) + 4f(2.2) - f(2.21)] \\ &= \frac{1}{0.02} [146.018381] = 7300.91905 \end{aligned}$$

Como el valor exacto de $f'(2.19)$ es 7310.449609, entonces el error absoluto será 9.530559, y el error relativo es 1.30×10^{-3} .

b. Empleando la fórmula (4), se tiene:

$$\begin{aligned} f'(2.19) &= \frac{1}{2h} [f(x_0 + h) - f(x_0 - h)] = \\ &= \frac{1}{0.02} [f(2.2) - f(2.18)] \\ &= \frac{1}{0.02} [146.2999] = 7314.995 \end{aligned}$$

Entonces el error absoluto será 4.545391, y el error relativo es: 6.22×10^{-4} .

Para analizar el *Error*, en el caso de la fórmula (1) necesitamos calcular las derivadas de $f(x)$; esto es:

$$f'(x) = (3x^2 + 2x^4)e^{x^2} - \cos x$$

$$f''(x) = e^{x^2} (6x + 14x^3 + 4x^5) + \operatorname{sen} x$$

$$f'''(x) = e^{x^2} (8x^6 + 54x^2 + 48x^4 + 6) + \cos x$$

$$\begin{aligned} \max \{ |f'''(x)| \} &= \max_{[2.19, 2.21]} \left\{ e^{x^2} [8x^6 + 48x^4 + 54x^2 + 6] + \cos x \right\} \\ &= 310,181.1849 \end{aligned}$$

Luego una cota superior para el error será:

$$\text{Error} \leq \frac{\max_{[2.19, 2.21]} |f'''(x)|}{3!} \left| \prod_{\substack{j=2 \\ j \neq 2}}^4 (x_2 - x_j) \right| = 10.339373$$

Inestabilidad de las fórmulas de derivación numérica:

Las fórmulas de derivación son de la forma:

$$f'(x_k) = P'_n(x_k) + Ch^2 f^{(n+1)}(\xi)$$

con $h = x_i - x_{i-1}$ y C una constante. El error depende de h , pues si h tiende a cero, entonces:

$$\lim_{h \rightarrow 0} [Ch^2 f^{(n+1)}(\xi)] = 0$$

Se deduce de esta forma que para h suficientemente pequeña podríamos calcular $f'(x_k)$ con cualquier precisión.

Sin embargo la resta de 2 números cercanos (Fórmula (4)), así como la división sobre h número pequeño (Fórmulas (4) y (5)) aumenta considerablemente el error del cálculo. Lo que empeora el proceso de derivación numérica, convirtiéndolo en un proceso inestable, para valores de h muy pequeños.

En otras palabras la disminución de h no siempre asegura un mejor resultado.

Ejemplo 3:

Sea

$$f(x) = \text{sen}x.$$

Buscaremos $f'(0.2)$ usando la fórmula

$$f'(x_0) \approx \frac{f(x_0 + h) - f(x_0 - h)}{2h}.$$

Cuyo error es más pequeño, que de la otra fórmula de 3 puntos. Se sabe que el valor con 9 cifras exactas es:

$$f'(0.2) = \cos 0.2 = 0.980066577.$$

Usaremos distintos valores de h y trabajaremos con una aritmética de 6 dígitos.

Si $h = 0.1$:

$$f'(0.2) \approx \frac{f(0.3) - f(0.1)}{0.2} = \frac{0.295520 - 0.099833}{0.2} = 0.978435$$

La respuesta tiene una cifra decimal exacta.

Si $h = 0.01 = 10^{-2}$:

$$f'(0.2) \approx \frac{f(0.21) - f(0.19)}{0.02} = \frac{0.208460 - 0.188859}{0.02} = 0.98005$$

La respuesta coincide con el valor exacto en 4 decimales.

Si $h = 0.001 = 10^{-3}$:

$$f'(0.2) \approx \frac{f(0.201) - f(0.199)}{0.002} = \frac{0.199649 - 0.197689}{0.002} = 0.980000$$

Comparando la respuesta obtenida con la respuesta anterior (para $h = 0.01$) podemos observar que la respuesta para $h = 0.01$ es la que se aproxima mejor al valor exacto.

Si $h = 10^{-6}$

$$\begin{aligned} f'(0.2) &= \frac{f(0.200001) - f(0.199999)}{0.000002} \\ &= \frac{0.198670 - 0.198668}{0.000002} = 1.00000 \end{aligned}$$

y en este caso se han perdido todas las cifras exactas. De aquí podemos afirmar que para este ejemplo $h = 0.01$ es el mejor valor (valor óptimo de h) para aproximar la derivada pedida.

Integración Numérica

El problema de la integración numérica o cuadratura numérica, consiste en estimar al número:

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

Casos Particulares:

Caso 1: Regla del Trapecio

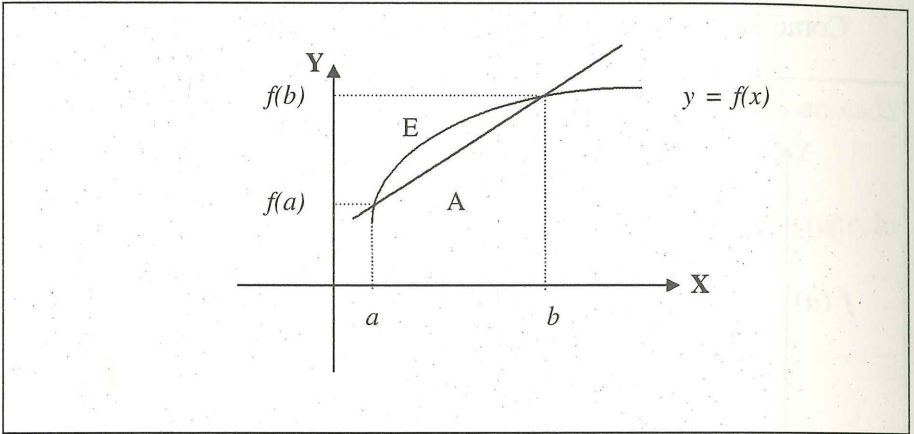
Sea $f(x)$ una función continua definida en $[a, b]$, y consideremos un polinomio de interpolación lineal para $f(x)$, basado en $x_0 = a$, y $x_1 = b$:

$$P_1(x) = f(x_0) + \frac{f(x_0) - f(x_1)}{x_0 - x_1} (x - x_0)$$

entonces la integral:

$$\int_a^b f(x) dx \approx \int_a^b P_1(x) dx$$

como se muestra en el siguiente gráfico:



Gráfica 4

donde la integral de $P_1(x)$, coincide con el área del trapecio: formado por los vértices $(a, 0)$, $(b, 0)$, $(a, f(a))$, $(b, f(b))$, la denotaremos por A , luego:

$$\int_a^b P_1(x) dx = A = \frac{(b-a)}{2} [f(a) + f(b)]$$

Entonces:

$$I(f) = \int_a^b f(x) dx = \frac{(b-a)}{2} [f(a) + f(b)] + E \approx \frac{(b-a)}{2} [f(a) + f(b)]$$

y recibe el nombre de la Regla Simple del Trapecio.

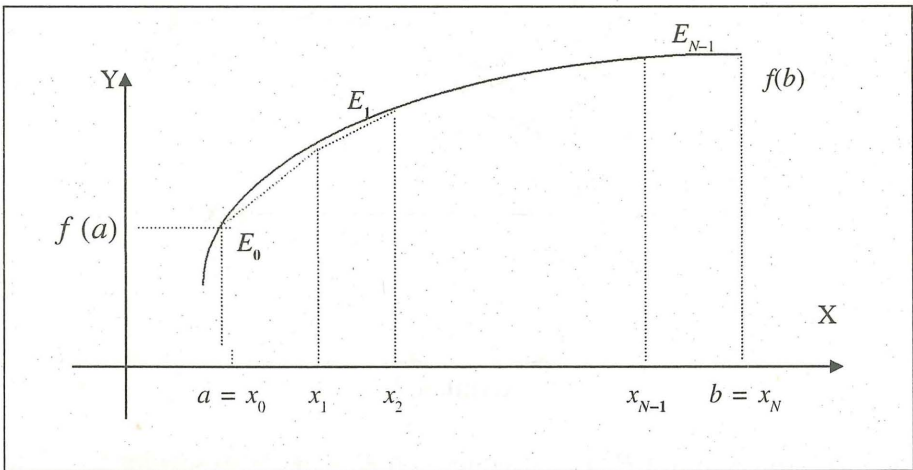
En este caso, si la longitud del intervalo $[a, b]$ no es pequeña, el error es grande, por lo que conviene dividir al intervalo $[a, b]$ en N subintervalos con una longitud cada uno de h ; esto es:

$$x_0 = a, x_1 = x_0 + h, \dots, x_i = x_0 + ih, \dots, x_N = x_0 + Nh = b.$$

Donde $h = (b - a)/N$, y cada subintervalo toma la forma de:

$$I_i = [x_i, x_{i+1}], i = 0, 1, 2, \dots, N-1.$$

Como se muestra en la gráfica a continuación:



Gráfica 5

Para cada intervalo I_i , se denota el área del trapecio correspondiente como:

$$\begin{aligned}
 A_i &= \frac{(x_i - x_{i+1})}{2} [f_i + f_{i+1}] = \frac{h}{2} [f_i + f_{i+1}] \\
 \Rightarrow I(f) &= \sum_{i=0}^{N-1} A_i + E = \frac{h}{2} \sum_{i=0}^{N-1} [f_i + f_{i+1}] + E = \\
 &= \frac{h}{2} [f_0 + f_1 + f_1 + f_2 + f_2 + f_3 + \dots + f_{N-1} + f_N] + E = \\
 &= \frac{h}{2} \left[f_0 + 2 \sum_{i=1}^{N-1} f_i + f_N \right] + E
 \end{aligned}$$

donde E es el error correspondiente a la suma de los errores en cada uno de los subintervalos I_i .

Esta última expresión recibe el nombre de Regla del Trapecio Compuesta.

Analicemos el error en el caso de la Regla Simple:

$$\text{Error} = \int_a^b f(x) dx - \left[\frac{b-a}{2} [f(a) + f(b)] \right]$$

Para analizar este resultado, utilizaremos el desarrollo de la serie de Taylor para $f(x)$, en torno a $\bar{x} = \frac{a+b}{2}$, $h = b - a$, es decir:

$$f(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + f''(\bar{x}) \frac{(x - \bar{x})^2}{2} + \dots = \sum_{j=0}^{+\infty} f^{(j)}(\bar{x}) \frac{(x - \bar{x})^j}{j!}$$

haciendo un cambio de variable: $z = x - \bar{x}$,

$$f(x) = \sum_{j=0}^{+\infty} f^{(j)}(\bar{x}) \frac{z^j}{j!}$$

Por lo tanto, si $x = a \Rightarrow z = a - \bar{x} = \frac{a-b}{2} = -\frac{h}{2}$, y

$$\text{si } x = b \Rightarrow z = b - \bar{x} = \frac{b-a}{2} = \frac{h}{2}$$

Para hallar:

$$A = \frac{(b-a)}{2} [f(a) + f(b)].$$

Emplearemos el desarrollo de Taylor obtenido, para $x = a$, y $x = b$:

$$\Rightarrow f(a) = \sum_{j=0}^{+\infty} \frac{f^{(j)}(\bar{x})}{j!} \left(\frac{-h}{2}\right)^j$$

$$f(b) = \frac{f^{(j)}(\bar{x})}{j!} \left(\frac{h}{2}\right)^j$$

$$\therefore A = \frac{b-a}{2} [f(a) + f(b)] = \frac{h}{2} \left[\sum_{j=0}^{+\infty} \frac{f^{(2j)}(\bar{x})}{2j!} \left(\frac{h}{2}\right)^{2j} \right]$$

y por otro lado, reemplazando el desarrollo de Taylor en la integral, se tiene:

$$\begin{aligned} \int_a^b f(x) dx &= \int_{-h/2}^{h/2} \sum_{j=0}^{+\infty} \frac{f^{(j)}(\bar{x})}{j!} (z)^j dx = \sum_{j=0}^{+\infty} \frac{f^{(j)}(\bar{x})}{j!} \int_{-h/2}^{h/2} z^j dz = \\ &= \sum_{j=0}^{+\infty} \frac{f^{(2j)}(\bar{x})}{(2j+1)!} 2 \left(\frac{h}{2}\right)^{(2j+1)} \end{aligned}$$

y, tomando los dos primeros términos de cada una de las series se obtiene:

$$\int_a^b f(x) dx \approx hf(\bar{x}) + \frac{1}{3!} \times \frac{1}{4} h^3 f^{(2)}(\bar{x}) \dots \quad (\text{I})$$

$$A = \frac{b-a}{2} [f(a) + f(b)] \approx hf(\bar{x}) + \frac{h^3}{8} f^{(2)}(\bar{x}) \dots \quad (\text{II})$$

$$\Rightarrow \text{Error} \approx (\text{I}) - (\text{II}) = -\frac{h^3}{12} f^{(2)}(\bar{x})$$

que es una aproximación del error en el caso de la Regla Simple del Trapecio.

Otra forma de calcular el error para la Regla del Trapecio Simple consiste en aplicar el teorema del valor promedio ponderado de las integrales, ver en Burden-Faires [2].

El error en el caso de la Regla Extendida del Trapecio es la suma de los errores de cada uno de los intervalos y está dado por:

$$E \approx \sum_{i=0}^{N-1} \frac{-h^3}{12} f^{(2)}(\bar{x}_i),$$

donde \bar{x}_i es el punto medio de cada uno de los subintervalos I_i , y se sabe por el “Teorema del Valor Medio para Sumatorias” que:

$$E \approx -\frac{N}{12} h^3 f^{(2)}(\eta) = -\frac{(b-a)^3}{12N^3} Nf^{(2)}(\eta)$$

para algún $\eta \in [a, b]$.

Esta última ecuación muestra que el error de la Regla Extendida del Trapecio es proporcional a h^2 para un intervalo fijo $[a, b]$.

Caso 2: Regla de 1/3 de Simpson

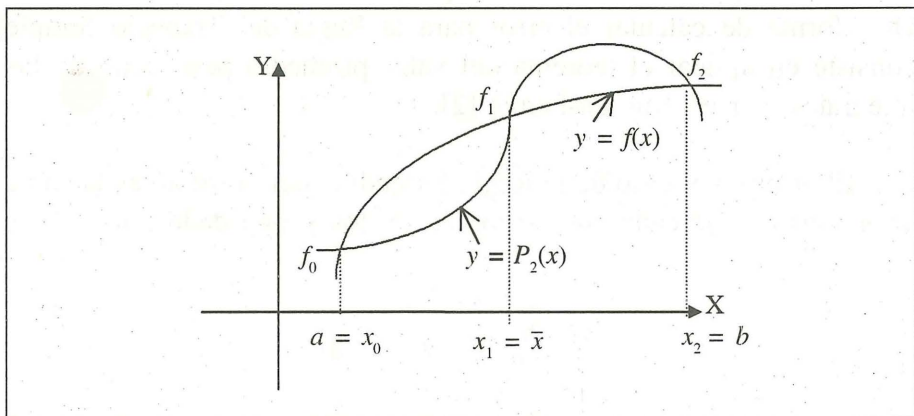
La Regla de 1/3 de Simpson se basa en la interpolación polinomial cuadrática. El polinomio de Newton hacia adelante ajustado a tres puntos equidistantes:

$$x_0 = a, x_1 = x_0 + h, x_2 = x_0 + 2h = b, \text{ donde } h = (b - a)/2$$

está dado por la ecuación:

$$P_2(x) = \sum_{i=0}^2 \Delta^i f_0 \binom{s}{i}$$

como se muestra en la gráfica a continuación:



Gráfica 6

Al integrar esta ecuación, se obtiene la Regla de 1/3 de Simpson:

$$I = \int_a^b f(x) dx = \frac{h}{3} [f(a) + 4f(x_1) + f(b)] + E$$

$$\text{donde } h = \frac{b-a}{2},$$

esta ecuación se puede escribir en forma equivalente:

$$I = \frac{h}{3} [f_0 + 4f_1 + f_2] + E \dots (1)$$

El error para la Regla de 1/3 de Simpson, se calcula en forma similar a como obtuvimos el error en el caso de la Regla del Trapecio, empleando el desarrollo de Taylor en torno de $x = x_1$, y toma la forma de:

$$E \approx -\frac{h^5}{90} f^{(4)}(x_1)$$

El error se anula si $f(x)$ es un polinomio de orden menor o igual a 3.

La Regla Extendida de 1/3 de Simpson es una aplicación repetida de la ecuación (1) para un dominio dividido en un número par de intervalos.

Si denotamos el número total de intervalos por N (par), la regla extendida de $1/3$ de Simpson se escribe como:

$$I = \int_a^b f(x) dx = \frac{h}{3} [(f_0 + 4f_1 + f_2) + (f_2 + 4f_3 + f_4) + \dots + (f_{N-2} + 4f_{N-1} + f_N)] + E$$

y esta ecuación se puede escribir en forma equivalente:

$$I = \frac{h}{3} \left[f(a) + 4 \sum_{i,\text{impar}} f_i + 2 \sum_{i,\text{par}} f_i + f(b) \right] + E \dots (2)$$

y el término del error está dado por la suma de los errores en cada uno de los intervalos; es decir:

$$E \approx \sum_{i,\text{impar}} -\frac{h^5}{90} f^{(4)}(x_i)$$

y se sabe por el “Teorema del Valor Medio para Sumatorias” que:

$$E \approx -\frac{N}{2} \times \frac{h^5}{90} f^{(4)}(\eta) = -\frac{b-a}{180} h^4 f^{(4)}(\eta)$$

para algún $\eta \in [a, b]$

Para un dominio fijo $[a, b]$ el error es proporcional a h^4 .

Caso 3: Regla de 3/8 de Simpson

La Regla de 3/8 de Simpson se obtiene al integrar un polinomio de interpolación de Newton de tercer grado. Para un dominio $[a, b]$ dividido en tres intervalos iguales, con los puntos:

$$x_0 = a, x_1 = x_0 + h, x_2 = x_0 + 2h, x_3 = x_0 + 3h = b$$

se tiene que:

$$I = \int_a^b f(x) dx = \frac{3h}{8} [f_0 + 3f_1 + 3f_2 + f_3] + E$$

donde $f_i = f(x_0 + ih)$, para $i = 0, 1, 2, 3$.

El término del error se calcula de manera similar a los casos (1) y (2), obteniéndose la expresión:

$$E \approx -\frac{3}{80} h^5 f^{(4)}(\bar{x})$$

Ejemplo 1:

Calcular la integral:

$$\int_1^3 \frac{e^x}{\sqrt{x}} dx$$

si se sabe que:

i	x_i	$y_i = f_i$
0	1	e
1	1.25	3.121858
2	1.5	3.659284
3	1.75	4.350071
4	2	5.224852
5	2.25	6.325157
6	2.5	7.704886
7	2.75	9.432862
8	3	11.596390

Tabla 15

- Usando la Regla del Trapecio, con $h = 0.25$.
- Usando la Regla de Simpson, con $h = 0.25$.
- Con la Regla de Simpson, con $h = 0.4$.

Solución:

- La Regla del Trapecio:

$$I(f) = \int_1^3 \frac{e^x}{\sqrt{x}} dx \approx \frac{h}{2} \left[f_0 + 2 \sum_{i=1}^7 f_i + f_8 \right] \approx 11.744076$$

Analizando el error:

$$\text{Error} \approx -\frac{N}{12} h^3 f^{(2)}(\eta), \text{ para algún } \eta \in]1, 3[$$

$$|\text{Error}| \leq \frac{N}{12} |h|^3 \max_{[1,3]} \{f^{(2)}(x)\}$$

Calculando las derivadas:

$$f(x) = \frac{e^x}{\sqrt{x}},$$

$$f^{(1)}(x) = e^x \left[x^{-1/2} - \frac{1}{2} x^{-3/2} \right]$$

$$f^{(2)}(x) = e^x \left[x^{-1/2} - x^{-3/2} + (3/4)x^{-5/2} \right]$$

$$f^{(3)}(x) = e^x \left[x^{-1/2} - (3/2)x^{-3/2} + (9/4)x^{-5/2} - (15/8)x^{-7/2} \right] = 0$$

$$\Leftrightarrow x^3 - (3/2)x^2 + (9/4)x - 15/8 = 0$$

$$\Leftrightarrow x_1 = 0.223213 - 1.31523i ; x_2 = 0.223213 + 1.31523i$$

$$x_3 = 1.05357$$

$$\Rightarrow \max_{[1,3]} \{f^{(2)}(x)\} = |f^{(2)}(1.05357)| = 2.029892652$$

$$|\text{Error}| \leq 0.02114471513$$

Se obtiene una precisión de por lo menos un decimal exacto en su valor absoluto.

b. La Regla de Simpson, para $h = 0.25$:

$$\int_1^3 f(x) dx \approx \frac{h}{3} \left[f_0 + 4 \sum_{i,\text{impar}} f_i + 2 \sum_{i,\text{par}} f_i + f_8 \right] \approx \frac{0.25}{3} (140.4125078) \approx 11.70104232$$

Analizando el error:

$$\text{Error} \approx -\frac{N}{2} \frac{h^5}{90} f^{(4)}(\xi(x)) \leq \frac{N}{180} h^5 \max_{[1,3]} \{f^{(4)}(x)\}$$

y por otro lado, calculando la cuarta derivada para luego maximizarla, se tiene:

$$f^{(4)}(x) = e^x \left[x^{-1/2} - 2x^{-3/2} + \frac{9}{2} x^{-5/2} - \frac{60}{8} x^{-7/2} + \frac{105}{16} x^{-9/2} \right]$$

$$\max_{[1,3]} \{f^{(4)}(x)\} \approx |f^{(4)}(1)| = 6.965597$$

y una cota superior para el error:

$$\Rightarrow |\text{Error}| \leq 3.023263 \times 10^{-4}$$

c. La Regla de Simpson, si $h=0.4$:

i	0	1	2	3	4	5
x_i	1	1.4	1.8	2.2	2.6	3
f_i	e	3.427270	4.509141	6.08466286	8.349856	11.596390

Tabla 16

Combinando ambas reglas de Simpson:

- La Regla de 1/3 aplicada a los intervalos: [1, 1.4], [1.4, 1.8];
- La Regla de 3/8 aplicada a los intervalos [1.8, 2.2], [2.2, 2.6], [2.6, 3] se obtiene:

$$\int_1^{1.8} f(x) dx \approx \frac{h}{3} [f_0 + 4f_1 + f_2] = 2.79153371$$

$$\int_{1.8}^3 f(x) dx \approx \frac{3h}{8} [f_2 + 3f_3 + 3f_4 + f_5] = 8.911363136$$

$$\Rightarrow \int_1^3 f(x) dx \approx 11.70289685$$

Analizando el error para el cálculo realizado, consideremos separadamente el error en cada una de las dos reglas utilizadas, y luego sumémoslas algebraicamente; esto es:

$$Error \approx -\frac{h^5}{90} f^{(4)}(\bar{x}_1) - \frac{3h^5}{80} f^{(4)}(\bar{x}_2)$$

$$\text{donde } \bar{x}_1 = \frac{1+1.8}{2} = 1.4, \quad \bar{x}_2 = \frac{1.8+3}{2} = 2.4$$

$$\Rightarrow Error \leq 1.976572 \times 10^{-3}$$

Nos proporciona una precisión de dos decimales exactos.

Ejemplo 2:

Calcular la integral:

$$\frac{1}{\sqrt{2\pi}} \int_{-2.5}^{1.5} e^{-t^2/2} dt$$

si se sabe que:

i	0	1	2	3	4	5	6	7	8
x_i	-2.5	-2	-1.5	-1	-0.5	0	0.5	1	1.5
f_i	.017528	.053991	.129518	.241971	.352065	.398942	.352065	.241971	.129518

Tabla 17

- Emplear la Regla del Trapecio.
- Emplear la Regla de Simpson.
- Calcular con la regla de Simpson:

$$\frac{1}{\sqrt{2\pi}} \int_{-2.5}^0 e^{-t^2/2} dt$$

Solución:

(a) Regla del Trapecio:

$$\frac{1}{\sqrt{2\pi}} \int_{-2.5}^{1.5} e^{-t^2/2} dt \approx \frac{h}{2} \left[f_0 + 2 \sum_{i=1}^7 f_i + f_8 \right] \approx 0.922023$$

Analizando el error:

$$\text{Error} \approx \frac{-Nh^3}{12} f^{(2)}(\eta) \leq \frac{-Nh^3}{12} \max_{[-2.5, 1.5]} \{f^{(2)}(x)\} = 0.03324519$$

el valor exacto de la integral es: 0.927.

Las derivadas de f son:

$$f^{(1)}(x) = \frac{-x}{\sqrt{2\pi}} e^{-x^2/2};$$

$$f^{(2)}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} [x^2 - 1]; \quad \max_{[-2.5, 1.5]} \{f^{(2)}(x)\} = |f^{(2)}(0)| = 1/\sqrt{2\pi}$$

$$f^{(3)}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} [-x^3 + 3x];$$

$$f^{(4)}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} [3 - 6x^2 + x^4]; \quad \max_{[-2.5, 1.5]} \{f^{(4)}(x)\} = |f^{(4)}(0)| = 3/\sqrt{2\pi}$$

(b) Regla de Simpson:

$$\frac{1}{\sqrt{2\pi}} \int_{-2.5}^{1.5} e^{-t^2/2} dt \approx \frac{h}{3} \left[f_0 + 4 \sum_{i, \text{impar}} f_i + 2 \sum_{i, \text{par}} f_i + f_8 \right] = 0.9269736667$$

$$\text{Error} \approx \frac{-Nh^5}{2 \times 90} f^{(4)}(\eta) \leq \frac{-Nh^5}{2 \times 90} \max_{[-2.5, 1.5]} \{|f^{(4)}(x)|\} = 0.0016623$$

(c) Regla de Simpson compuesta :

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_{-2.5}^0 e^{-t^2/2} dt \\ & \approx \frac{h}{3} [f_0 + 4f_1 + f_2] + \frac{3h}{8} [f_2 + 3f_3 + 3f_4 + f_5] = 0.49373317 \end{aligned}$$

El valor exacto de la integral es 0.4938.

$$\text{Error} \approx \left| \frac{-h^5}{90} f^{(4)}(\bar{x}) + \frac{-3h^5}{80} f^{(4)}(\bar{x}^*) \right| = 1.1441 \times 10^{-4}$$

con $\bar{x} = -2$; $\bar{x}^* = -0.75$

Lo que nos indica una precisión de tres decimales exactos.

Comentarios:

- (1) Los métodos numéricos de integración quedan definidos a través de valores funcionales.
- (2) Cuando no se conoce la función, se puede trabajar con el doble del número de puntos o con $h' = h/2$, donde h es el paso anterior, ver en Conte-Boor [5].

4.6 Ajuste de Datos

La interpolación de funciones no es la única forma de estimar a las funciones, sobre todo no resulta eficiente si se trata de trabajar con una cantidad grande de puntos de interpolación. Existen otras formas como el Método de Mínimos Cuadrados que veremos a continuación.

Sean x e y dos variables que están relacionadas a través de la función $f(x)$; es decir:

$$y = f(x)$$

y se obtienen valores experimentales, de la forma:

$$(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$$

de manera que,

$$f(x_i) = y_i + \varepsilon_i, \quad i = 1, 2, 3, \dots, N.$$

Donde los ε_i , son los errores experimentales.

Buscaremos a la función que se aproxime a $f(x)$, y puede ser escrita en la forma:

$$F(x) = F(x, c_1, c_2, \dots, c_k) = \sum_{i=1}^k c_i \Phi_i(x).$$

Donde c_1, c_2, \dots, c_k son parámetros y $\Phi_i(x)$ son funciones seleccionadas a priori.

Luego esta función $F(x_i)$, nos genera desviaciones en cada punto (x_i, y_i) , que las representaremos como:

$$e_i = y_i - F(x_i, c_1, c_2, \dots, c_k),$$

Se desea que estas desviaciones sean lo más pequeñas posibles, por lo cual se acostumbra determinar los c_j de tal manera que se minimice la función $E(c_1, c_2, \dots, c_k)$:

$$\min_{\{c_1, c_2, \dots, c_k\}} \{E(c_1, c_2, \dots, c_k)\} = \min_{\{c_1, c_2, \dots, c_k\}} \left\{ \sum_{i=1}^N (y_i - F(x_i, c_1, c_2, \dots, c_k))^2 \right\}$$

y esto lo logramos derivando respecto a cada parámetro c_j , para $j = 1, 2, 3, \dots, k$, e igualando a cero; esto es:

$$\frac{\partial \{E[c_1, \dots, c_k]\}}{\partial c_j} = 0, \quad j = 1, 2, \dots, k$$

$$\frac{\partial}{\partial c_j} \left\{ \sum_{i=1}^N [y_i - F(x_i, c_1, c_2, \dots, c_k)]^2 \right\} = 0$$

$$\Rightarrow \sum_{i=1}^N (y_i - F(x_i, c_1, c_2, \dots, c_k)) \frac{\partial F(x_i, c_1, c_2, \dots, c_k)}{\partial c_j} = 0$$

$$\therefore \sum_{i=1}^N (y_i - F(x_i, c_1, c_2, \dots, c_k)) \Phi_j(x_i) = 0$$

Obteniéndose el siguiente sistema lineal de k ecuaciones con k incógnitas que se denominan las Ecuaciones Normales del Método de Mínimos Cuadrados, con parámetros c_1, c_2, \dots, c_k :

$$\sum_{i=1}^N F(x_i, c_1, c_2, \dots, c_k) \Phi_j(x_i) = \sum_{i=1}^N y_i \Phi_j(x_i),$$

$$c_1 \sum_{i=1}^N \Phi_1(x_i) \Phi_j(x_i) + c_2 \sum_{i=1}^N \Phi_2(x_i) \Phi_j(x_i) + \dots + c_k \sum_{i=1}^N \Phi_k(x_i) \Phi_j(x_i) =$$

$$= \sum_{i=1}^N y_i \Phi_j(x_i), \forall j = 1, 2, \dots, k$$

Aplicaciones:

Ejemplo 1:

Dados los siguientes datos:

i	1	2	3	4
x_i	1	2	3	4
y_i	7	11	17	27

Tabla 18

Obtener:

- Un ajuste de mínimos cuadrados lineal.
- Un ajuste de mínimos cuadrados de la forma: $y = Ae^{Mx}$.

Solución:

- Ajuste lineal:

Sea

$$y = F(x, c_1, c_2) = c_1 \Phi_1(x) + c_2 \Phi_2(x),$$

donde $\Phi_1(x) = 1, \Phi_2(x) = x$, y de ahí las Ecuaciones Normales son:

$$4c_1 + c_2 \sum_{j=1}^4 x_j = \sum_{j=1}^4 y_j$$

$$c_1 \sum_{j=1}^4 x_j + c_2 \sum_{j=1}^4 x_j^2 = \sum_{j=1}^4 x_j y_j,$$

donde

$$\sum_{j=1}^4 x_j = 10 ; \quad \sum_{j=1}^4 x_j^2 = 30; \quad \sum_{j=1}^4 y_j = 62; \quad \sum_{j=1}^4 x_j y_j = 188$$

$$\Rightarrow \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 6.6 \end{bmatrix}$$

Por lo tanto el ajuste lineal pedido será:

$$y = F(x, c_1, c_2) = -1 + 6.6x.$$

b. Ajuste de la forma $y = Ae^{Mx}$:

Aplicando logaritmos a la ecuación, esto es: $z = \ln y = \ln A + Mx$, entonces:

i	1	2	3	4
x_i	1	2	3	4
y_i	7	11	17	27
$z_i = \ln y_i$	1.94591	2.3979	2.83321	3.29584

Tabla 19

Sea

$$z = F(x, c_1, c_2) = c_1 \Phi_1(x) + c_2 \Phi_2(x),$$

donde

$$\Phi_1(x) = 1, \quad \Phi_2(x) = x,$$

$$c_1 = \ln A; \quad c_2 = M.$$

Entonces

$$z = F(x, c_1, c_2) = c_1 + c_2 x,$$

y haciendo uso de las Ecuaciones Normales:

$$\begin{aligned} 4c_1 + c_2 \sum_{j=1}^4 x_j &= \sum_{j=1}^4 z_j \\ c_1 \sum_{j=1}^4 x_j + c_2 \sum_{j=1}^4 x_j^2 &= \sum_{j=1}^4 x_j z_j ; \end{aligned}$$

donde

$$\begin{aligned} \sum_{j=1}^4 x_j z_j &= 28.4247 ; \quad \sum_{j=1}^4 z_j = 10.47286 ; \\ \Rightarrow \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} 1.49694 \\ 0.44851 \end{bmatrix} \end{aligned}$$

luego la función pedida será:

$$f(x) = 4.468 e^{0.44851x}$$

Ejemplo 2:

La siguiente tabla nos muestra un conjunto de datos que corresponden a las variables:

Densidad de partículas por área, que emergen de una lámina delgada a un ángulo α , con la perpendicular:

i	1	2	3	4	5	6	7	8
x (grados)	5	15	30	40	50	75	105	150
x (rad)	.087266	.261799	.5235987	.698132	.872665	1.308997	1.832596	2.617994
y (part/área)	8×10^5	2×10^4	800	500	107	26	8	5

Tabla 20

- Obtener la parábola de mínimos cuadrados.
- Construir un ajuste de datos de la forma: $f(x) = \alpha x^\beta$.

Solución:

(a) Ajuste Cuadrático:

Se define:

$$F(x, c_1, c_2) = c_1 \Phi_1(x) + c_2 \Phi_2(x) + c_3 \Phi_3(x);$$

Donde: $\Phi_1(x) = 1$, $\Phi_2(x) = x$, $\Phi_3(x) = x^2$; $\forall x \in [0.087, 2.62]$

Se obtienen las Ecuaciones Normales:

$$(1) \dots c_1 \sum_{i=1}^8 \Phi_1(x_i) \Phi_1(x_i) + c_2 \sum_{i=1}^8 \Phi_1(x_i) \Phi_2(x_i) + c_3 \sum_{i=1}^8 \Phi_1(x_i) \Phi_3(x_i) = \sum_{i=1}^8 y_i \Phi_1(x_i);$$

$$(2) \dots c_1 \sum_{i=1}^8 \Phi_2(x_i) \Phi_1(x_i) + c_2 \sum_{i=1}^8 \Phi_2(x_i) \Phi_2(x_i) + c_3 \sum_{i=1}^8 \Phi_2(x_i) \Phi_3(x_i) = \sum_{i=1}^8 y_i \Phi_2(x_i);$$

$$(3) \dots c_1 \sum_{i=1}^8 \Phi_3(x_i) \Phi_1(x_i) + c_2 \sum_{i=1}^8 \Phi_3(x_i) \Phi_2(x_i) + c_3 \sum_{i=1}^8 \Phi_3(x_i) \Phi_3(x_i) = \sum_{i=1}^8 y_i \Phi_3(x_i);$$

que toman la forma:

$$(1) \dots c_1 \sum_{i=1}^8 1 + c_2 \sum_{i=1}^8 x_i + c_3 \sum_{i=1}^8 x_i^2 = \sum_{i=1}^8 y_i;$$

$$(2) \dots c_1 \sum_{i=1}^8 x_i + c_2 \sum_{i=1}^8 x_i^2 + c_3 \sum_{i=1}^8 x_i^3 = \sum_{i=1}^8 x_i y_i;$$

$$(3) \dots c_1 \sum_{i=1}^8 x_i^2 + c_2 \sum_{i=1}^8 x_i^3 + c_3 \sum_{i=1}^8 x_i^4 = \sum_{i=1}^8 x_i^2 y_i;$$

esto es:

$$\begin{bmatrix} 8 & \sum_{i=1}^8 x_i & \sum_{i=1}^8 x_i^2 \\ \sum_{i=1}^8 x_i & \sum_{i=1}^8 x_i^2 & \sum_{i=1}^8 x_i^3 \\ \sum_{i=1}^8 x_i^2 & \sum_{i=1}^8 x_i^3 & \sum_{i=1}^8 x_i^4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^8 y_i \\ \sum_{i=1}^8 x_i y_i \\ \sum_{i=1}^8 x_i^2 y_i \end{bmatrix}$$

$$\begin{bmatrix} 8 & 8.203048 & 13.525016 \\ 8.203048 & 13.525016 & 27.507976 \\ 13.525016 & 27.507976 & 62.088152 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 821446 \\ 75971.88478 \\ 8113.249016 \end{bmatrix}$$

$$[c_1 \ c_2 \ c_3]' = [500086.2387, -772398.8014, 233403.0905]'$$

entonces: $\hat{y}_i(\text{estimación}) = F(x, c_1, c_2, c_3)$

i	\hat{y}_i
1	434459.5322
2	313870.1531
3	159647.9944
4	74607.85171
5	3787.6090
6	-111051.5474
7	-131545.8856
8	77670.5221

Tabla 21

Observamos que no es un buen ajuste.

(b) Ajuste de la forma: $f(x) = \alpha x^\beta$:

Aplicando logaritmos a la ecuación dada, con el objeto de transformarla en una ecuación lineal, se tiene:

$$\ln f(x) = \ln \alpha + \beta \ln x$$

y si se define: $z = \ln y$, y $w = \ln x$,

$$\text{entonces: } z_i = \ln \alpha + \beta w_i$$

lo que nos lleva a definir:

$$F(w, c_1, c_2) = c_1 \Phi_1(w) + c_2 \Phi_2(w)$$

Donde $\Phi_1(w) = 1$, y $\Phi_2(w) = w$

$$\text{y } c_1 = \ln \alpha, c_2 = \beta$$

Se transforman los datos originales:

i	$w_i = \ln x_i$	$z_i = \ln y_i$
1	-2.438794	13.592367
2	-1.340178	9.903488
3	-0.6470297	6.684612
4	-0.359347	6.214608
5	-0.136204	4.672829
6	0.269261	3.258097
7	0.605734	2.079442
8	0.962408	1.609438

Tabla 22

Se construyen las Ecuaciones Normales:

$$(1) \dots c_1 \sum_{i=1}^8 \Phi_1(w_i) \Phi_1(w_i) + c_2 \sum_{i=1}^8 \Phi_1(w_i) \Phi_2(w_i) = \sum_{i=1}^8 z_i \Phi_1(w_i);$$

$$(2) \dots c_1 \sum_{i=1}^8 \Phi_2(w_i) \Phi_1(w_i) + c_2 \sum_{i=1}^8 \Phi_2(w_i) \Phi_2(w_i) = \sum_{i=1}^8 z_i \Phi_2(w_i);$$

que toman la forma:

$$(1) \dots c_1 \sum_{i=1}^8 1 + c_2 \sum_{i=1}^8 w_i = \sum_{i=1}^8 z_i;$$

$$(2) \dots c_1 \sum_{i=1}^8 w_i + c_2 \sum_{i=1}^8 w_i^2 = \sum_{i=1}^8 w_i z_i;$$

esto es:

$$\begin{bmatrix} 8 & \sum_{i=1}^8 w_i \\ \sum_{i=1}^8 w_i & \sum_{i=1}^8 w_i^2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^8 z_i \\ \sum_{i=1}^8 w_i z_i \end{bmatrix}$$

$$\begin{bmatrix} 8 & -3.084150 \\ -3.084150 & 9.675767 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 48.014881 \\ -49.930418 \end{bmatrix}$$

$$[c_1 \quad c_2]^T = [4.574590 \quad -3.702208]^T$$

$$\Rightarrow \alpha = 96.988245; \quad \beta = -3.702208$$

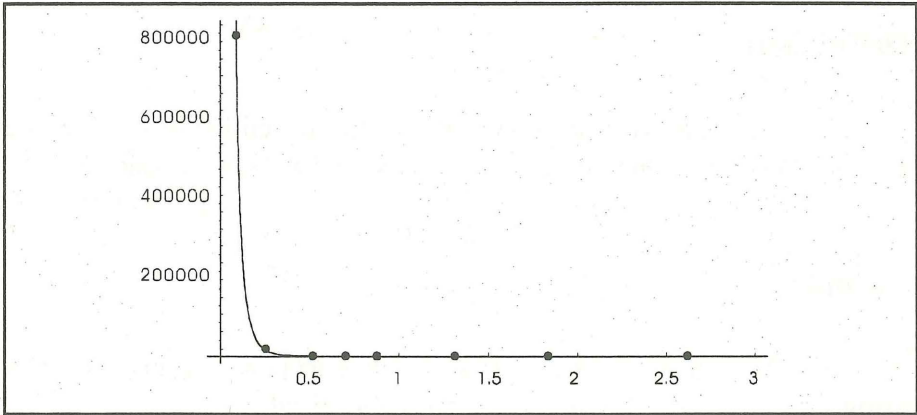
$$\Rightarrow \hat{y} = f(x) = 96.988245x^{-3.702208}$$

luego se obtiene: \hat{y}_i (estimación de y_i) = $F(w_i, c_1, c_2)$:

i	\hat{y}_i
1	808967.28
2	13852.33
3	1064.25
4	366.855
5	160.588
6	35.7922
7	10.2989
8	2.74987

Tabla 23

Este sería el mejor ajuste de los dos



Gráfica 7

4.7 Producto Escalar y Ortogonalidad de las Funciones

La motivación para esta discusión proviene del problema de aproximación de mínimos cuadrados por polinomios. Puesto que al resolver un sistema de ecuaciones lineales, a través del método de mínimos cuadrados, hay una tendencia a que los errores del cálculo crezcan, por la cantidad de operaciones que se deben de realizar, y una forma de evitar esto es por medio de los polinomios ortogonales.

Sean g y h dos funciones definidas en $[a, b]$, y $w(x)$ la función peso positiva y definida en el mismo intervalo.

Definición:

Se define el producto escalar, de dos funciones cualesquiera $g(x)$ y $h(x)$ en $[a, b]$ como:

a. En el caso continuo:

$$\langle g, h \rangle = \int_a^b g(x)h(x)w(x)dx$$

b. En el caso discreto:

$$\langle g, h \rangle = \sum_{i=1}^n g(x_i)h(x_i)w(x_i)$$

Donde suponemos que los puntos $x_1, x_2, \dots, x_n \in [a, b]$.

Definición:

Sean dos funciones $g(x)$ y $h(x)$ definidas en $[a, b]$ se dice que $g(x)$ y $h(x)$ son mutuamente ortogonales en $[a, b]$, si y sólo si:

$$\langle g, h \rangle = 0$$

Ejemplo 1:

Sean las funciones $g(x) = 1$, $h(x) = x$, y $w(x) = 1$, $g(x)$ y $h(x)$ son ortogonales en $[-1, 1]$ ya que su producto escalar es:

$$\langle g, h \rangle = \int_{-1}^1 g(x)h(x)dx = \int_{-1}^1 xdx = 0$$

Definición:

Decimos que $P_0(x), P_1(x), P_2(x), \dots$ es una sucesión finita o infinita de polinomios ortogonales en $[a, b]$ si los $P_i(x)$ son todos mutuamente ortogonales en $[a, b]$ y cada $P_i(x)$ es un polinomio de grado exacto i . En otras palabras,

(1) Para cada i ,

$$P_i(x) = \alpha_i x^i + \text{un polinomio de grado } < i,$$

con $\alpha_i \neq 0$.

(2) Siempre que $i \neq j$, se tiene $\langle P_i, P_j \rangle = 0$.

Ejemplo 2:

Sean las funciones:

$$P_0(x) = 1, P_1(x) = x, P_2(x) = 3(x^2 - 1/3)/2, \text{ definidas en } [-1, 1]$$

Éstas forman una sucesión de tres polinomios ortogonales en $[-1, 1]$ ya que,

$$\langle P_0, P_1 \rangle = 0$$

$$\langle P_0, P_2 \rangle = \int_{-1}^1 1(3/2)(x^2 - 1/3)dx = 0,$$

$$\langle P_1, P_2 \rangle = \int_{-1}^1 x(3/2)(x^2 - 1/3)dx = 0.$$

Algunas Propiedades de los Polinomios Ortogonales:

Sea $P_0(x), P_1(x), P_2(x), \dots, P_k(x)$ una sucesión finita de polinomios ortogonales. Entonces se pueden demostrar las siguientes propiedades:

Propiedad 1: Si $P(x)$ es cualquier polinomio de grado menor o igual que k , $P(x)$ se puede expresar:

$$P(x) = d_0 P_0(x) + d_1 P_1(x) + d_2 P_2(x) + \dots + d_k P_k(x) = \sum_{i=0}^k d_i P_i(x)$$

con los coeficientes d_0, d_1, \dots, d_k unívocamente determinados por $P(x)$. Específicamente si

$$P(x) = a_k x^k + \text{un polinomio de grado} < k$$

y si el coeficiente inicial de $P_k(x)$ es α_k , entonces $d_k = a_k / \alpha_k$.

Esta propiedad se deduce a partir de (1), por inducción sobre k .

Ejemplo 3:

Empleando la sucesión de polinomios ortogonales del ejemplo 2, el polinomio de grado 2:

$$P(x) = a_0 + a_1 x + a_2 x^2,$$

se puede escribir como:

$$P(x) = d_0 P_0(x) + d_1 P_1(x) + d_2 P_2(x)$$

donde:

$$\begin{aligned} d_0 &= (a_0 + a_2/3) \\ d_1 &= a_1 \\ d_2 &= 2a_2/3 \end{aligned}$$

Combinando la propiedad 1 con (2), se obtiene la propiedad 2.

Propiedad 2: Si $P(x)$ es un polinomio de grado menor que k , entonces $P(x)$ es ortogonal a $P_k(x)$, esto es,

$$\langle P, P_k \rangle = 0.$$

Consideremos a la sucesión de polinomios ortogonales: $P_0(x)$, $P_1(x)$, $P_2(x)$, dada en el ejemplo 2. Luego para $P(x) = 1 + x$, se tiene:

$$\langle P, P_2 \rangle = \int_{-1}^1 (1+x)(3/2)(x^2 - 1/3) dx = 0$$

Esta propiedad tiene consecuencias importantes, como las siguientes propiedades:

Propiedad 3:

$P_k(x)$ tiene k ceros reales simples, y todos están en el intervalo $[a, b]$ esto es $P_k(x)$ es de la forma:

$$P_k(x) = \alpha_k (x - \xi_{1,k})(x - \xi_{2,k}) \dots (x - \xi_{k,k})$$

para k -puntos distintos $\xi_{1,k}, \xi_{2,k}, \dots, \xi_{k,k}$ en $]a, b[$.

Una demostración sencilla de la propiedad 3 es como sigue:

Sea $k > 0$ y sean $\xi_{1,k}, \xi_{2,k}, \dots, \xi_{r,k}$ puntos dados del intervalo $[a, b]$ en los cuales $P_k(x)$ cambia de signo.

Afirmamos que entonces $r \geq k$, pues si r fuera menor que k , entonces con $\bar{x} \in]\max_i \xi_{i,k}, b[$, se define al polinomio de grado menor que k :

$$P(x) = P_k(\bar{x})(x - \xi_{1,k})(x - \xi_{2,k}) \dots (x - \xi_{r,k}) \dots$$

el cual, en todo punto de $]a, b[$, tiene el mismo signo que $P_k(x)$. Por consiguiente, por una parte, por la propiedad 2,

$$\langle P, P_k \rangle = \int_a^b P(x)P_k(x)w(x)dx = 0$$

mientras que por otra parte:

$$P(x) P_k(x) w(x) > 0 \quad \forall x \in]a, b[, \text{ excepto } x = \xi_{1,k}, \xi_{2,k}, \dots, \xi_{r,k}$$

y estos dos hechos se contradicen entre sí. En consecuencia debemos tener $r \geq k$; esto es $P_k(x)$ es un polinomio de grado k , y cada $\xi_{i,k}$ es una raíz de $P_k(x)$, r no puede ser mayor que k , por consiguiente r debe ser igual a k ; esto es, los k puntos distintos $\xi_{i,k}$, $i = 1, 2, \dots, k$ son exactamente los ceros de $P_k(x)$.

Propiedad 4: Los polinomios ortogonales satisfacen una relación de recurrencia de tres términos. Si tomamos,

$$A_i = \alpha_{i+1} / \alpha_i, \quad \forall i, \\ P_{-1}(x) = 0$$

y definimos:

$$S_i = \langle P_i, P_i \rangle \neq 0, \quad \text{para } i = 0, 1, 2, \dots, k-1;$$

entonces esta relación de recurrencia se puede escribir como:

$$P_{i+1}(x) = A_i(x - B_i)P_i(x) - C_iP_{i-1}(x), \quad i = 0, 1, 2, \dots, k-1,$$

donde

$$B_i = \langle xP_i(x), P_i(x) \rangle / S_i, \quad i = 1, 2, \dots, k-1 \quad \text{y} \quad C_i = \begin{cases} \text{arbitrario}, & i = 0 \\ \frac{A_i S_i}{A_{i-1} S_{i-1}}, & i > 0 \end{cases}$$

Ver la demostración en Stoer-Bulirsch [23].

Esta propiedad se puede usar para generar sucesiones de polinomios ortogonales (previsto que los números S_i y B_i puedan calcularse y que los S_i no sean cero). En tal proceso, es corriente escoger los coeficientes principales α_i , o en forma equivalente, los números A_i , de modo que la sucesión resultante sea particularmente simple en algún sentido.

Casos Particulares

Caso 1. Polinomios de Legendre:

Si el producto escalar está dado por:

$$\langle g, h \rangle = \int_{-1}^1 g(x)h(x)dx, \quad \text{con } w(x) = 1$$

Entonces los polinomios ortogonales resultantes están asociados con el nombre de Legendre.

Comenzando con $P_0(x) = 1$, se obtiene

$$S_0 = \int_{-1}^1 1 dx = 2, \quad B_0 = \int_{-1}^1 x(1) dx = 0$$

Por consiguiente, a partir de la propiedad 4 definiendo $A_0 = 1$, obtenemos

$$P_1(x) = x$$

Además,

$$S_1 = \int_{-1}^1 x^2 dx = 2/3, \quad B_1 = \int_{-1}^1 x x^2 dx = 0, \quad C_1 = S_1 / S_0 = 1/3$$

de modo que, de nuevo por la propiedad 4,

$$P_2(x) = x^2 - 1/3.$$

Otra vez,

$$S_2 = \int_{-1}^1 (x^2 - 1/3)^2 dx = 8/45, \quad B_2 = \int_{-1}^1 x(x^2 - 1/3)^2 dx = 0, \quad C_2 = 4/15$$

de modo que:

$$P_3(x) = x^3 - 3x/5.$$

Se acostumbra normalizar a los polinomios de Legendre de manera que:

$$P_k(1) = 1, \quad \forall k.$$

Con esta normalización, los coeficientes de la relación de recurrencia se transforman en:

$$A_k = (2k + 1)/(k + 1), \quad B_k = 0, \quad C_k = k/(k + 1), \quad k = 0, 1, 2, \dots,$$

tal que:

$$P_{k+1}(x) = \frac{(2k+1)xP_k(x) - kP_{k-1}(x)}{k+1}$$

La siguiente tabla muestra los primeros polinomios normalizados de Legendre:

k	$P_k(x)$
0	1
1	x
2	$(3/2)(x^2 - 1/3)$
3	$(5/2)(x^3 - (3/5)x)$
4	$(35/8)(x^4 - (6/7)x^2 + 3/35)$

Tabla 24

Caso 2: Polinomios de Hermite ($H_k(x)$):

Resulta que cuando se usa el producto escalar:

$$\langle g, h \rangle = \int_{-\infty}^{\infty} g(x)h(x)e^{-x^2} dx$$

$$\text{con } w(x) = e^{-x^2}, \text{ para } x \in]-\infty, \infty[$$

Estos polinomios satisfacen la relación de recurrencia:

$$H_{k+1}(x) = 2xH_k(x) - 2kH_{k-1}(x), \quad k = 0, 1, 2, \dots$$

En la siguiente tabla se muestran los primeros polinomios de Hermite:

k	$H_k(x)$
0	1
1	$2x$
2	$4x^2 - 2$
3	$8x^3 - 12x$
4	$16x^4 - 48x^2 + 12$

Tabla 25

4.8 Aproximación de Mínimos Cuadrados por Polinomios Ortogonales

Sea $f(x)$ una función definida sobre algún intervalo $[a, b]$ y supongamos que queremos aproximar a $f(x)$ sobre $[a, b]$ por un polinomio $P(x)$ de grado menor o igual a k . Si medimos la diferencia entre $f(x)$ y $P(x)$ por el producto escalar:

$$\langle f(x) - P(x), f(x) - P(x) \rangle = \begin{cases} \int_a^b [f(x) - P(x)]^2 w(x) dx, & \text{caso continuo} \\ \sum_{n=1}^N [f(x_n) - P(x_n)]^2 w(x_n), & \text{caso discreto} \end{cases} \quad \dots \quad (*)$$

Buscaremos un polinomio de grado menor o igual que k para el cual (*) ésta diferencia sea tan pequeña como sea posible, tal polinomio se llama una aproximación ponderada de mínimos cuadrados a $f(x)$ por polinomios de grado menor o igual que k .

El problema de hallar tal polinomio se resuelve de la siguiente manera:

supongamos que tenemos, una sucesión:

$$P_0(x), P_1(x), \dots, P_k(x)$$

de polinomios mutuamente ortogonales en $[a, b]$, de tal manera que por la propiedad 1 se puede escribir a $P(x)$, como:

$$P(x) = d_0 P_0(x) + d_1 P_1(x) + \dots + d_k P_k(x).$$

Para hallar los coeficientes adecuados d_0, d_1, \dots, d_k ; se construye la función:

$$E(d_0, d_1, \dots, d_k) = \left\langle f(x) - \sum_{i=0}^k d_i P_i(x), f(x) - \sum_{i=0}^k d_i P_i(x) \right\rangle$$

Los coeficientes d_0, d_1, \dots, d_k que buscamos, serán aquellos que minimicen a la función $E(d_0, d_1, \dots, d_k)$.

Realizando un trabajo previo de cálculo en forma similar a como se hizo en "Ajuste de Datos", se muestra que los "mejores" coeficientes d_0, d_1, \dots, d_k deben satisfacer las Ecuaciones Normales:

$$d_0 \langle P_0, P_i \rangle + d_1 \langle P_1, P_i \rangle + \dots + d_k \langle P_k, P_i \rangle = \langle f, P_i \rangle; \quad i = 0, 1, 2, \dots, k.$$

las cuales, debido a la ortogonalidad de los $P_i(x)$, se reducen a,

$$d_i \langle P_i, P_i \rangle = \langle f, P_i \rangle, \quad i = 0, 1, 2, \dots, k.$$

Por consiguiente, como:

$$\langle P_i, P_i \rangle \neq 0, \quad \forall i = 0, 1, 2, \dots, k.$$

Entonces los mejores coeficientes están dados simplemente por:

$$d_i = \frac{\langle f, P_i \rangle}{\langle P_i, P_i \rangle}, \quad i = 0, 1, 2, \dots, k.$$

Ejemplo 1:

Calcular el polinomio $P(x)$ empleando los Polinomios de Legendre de grado menor o igual que 3, que minimiza:

$$\int_{-1}^1 [e^x - P(x)]^2 dx$$

sobre todos los polinomios $P(x)$ de grado menor o igual que 3.

Solución:

En este caso,

$$f(x) = e^x,$$

y el producto escalar se expresa por

$$\langle g, h \rangle = \int_{-1}^1 g(x)h(x)dx$$

Entonces:

$$\langle f, P_0 \rangle = \int_{-1}^1 e^x dx = e - 1/e$$

$$\langle f, P_1 \rangle = \int_{-1}^1 x e^x dx = 2/e;$$

$$\langle f, P_2 \rangle = \frac{3}{2} \int_{-1}^1 e^x (x^2 - 1/3) dx = e - 7/e;$$

$$\langle f, P_3 \rangle = \frac{5}{2} \int_{-1}^1 e^x (x^3 - \frac{3}{5}x) dx = -5e + 37/e;$$

y se tiene que $\langle P_i, P_i \rangle = \frac{2}{2i+1}$, para $i = 0, 1, 2, 3$

$$\therefore \langle P_0, P_0 \rangle = 2, \quad \langle P_1, P_1 \rangle = \frac{2}{3}, \quad \langle P_2, P_2 \rangle = \frac{2}{5}, \quad \langle P_3, P_3 \rangle = \frac{2}{7}$$

$$\Rightarrow d_0 = \frac{e-1/e}{2}, \quad d_1 = 3/e, \quad d_2 = \frac{5(e^2-7)}{2e}, \quad d_3 = \frac{7(37-5e^2)}{2e}$$

Luego hallamos que la aproximación de mínimos cuadrados a e^x sobre $[-1, 1]$ por el polinomio cúbico es:

$$P(x) = 1.1752012P_0(x) + 1.103638324P_1(x) + 0.3578143506P_2(x) + 0.07045563P_3(x).$$

Si reemplazamos $P_i(x)$ por expresiones equivalentes en potencias de x usando la tabla de los primeros polinomios de Legendre y reordenamos, obtenemos:

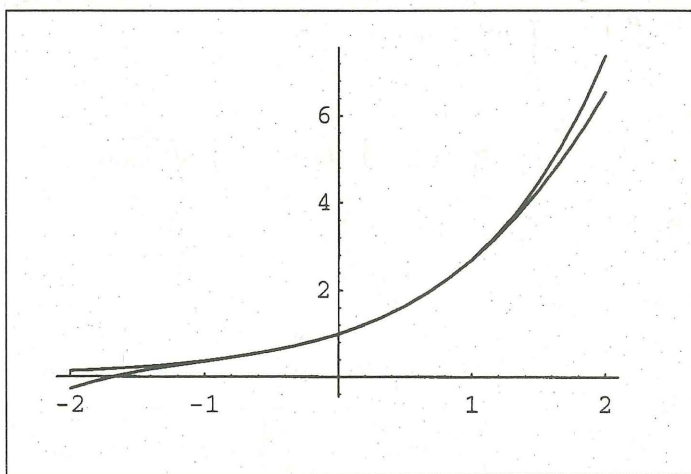
$$P(x) = 0.9962940183 + 0.9979548730x + 0.5367215260x^2 + 0.1761390842x^3,$$

sobre $[-1, 1]$.

Este polinomio nos proporcionará el valor de la integral:

$$\int_{-1}^1 [e^x - P(x)]^2 dx \approx 2.2289 \times 10^{-5}$$

lo cual nos indica una buena aproximación. Esto también se puede apreciar en la siguiente gráfica:



Gráfica 8

En la cual se observan que las curvas en el intervalo $[-1, 1]$, prácticamente se superponen una sobre la otra.

Ejemplo 2:

Calcular el polinomio $P(x)$ empleando los Polinomios de Legendre de grado menor o igual que 3, que minimiza:

$$\int_{-1}^1 [f(x) - P(x)]^2 dx; \quad f(x) = \sqrt{1-x}$$

sobre todos los polinomios $P(x)$ de grado menor o igual que 2.

Solución:

$$\text{Sea } P(x) = \sum_{i=0}^2 d_i P_i(x) = d_0 P_0(x) + d_1 P_1(x) + d_2 P_2(x);$$

Calculando:

$$d_0 = \frac{\langle f, P_0 \rangle}{\langle P_0, P_0 \rangle} = \frac{\int_{-1}^1 f(x)P_0(x)dx}{\int_{-1}^1 P_0^2(x)dx} = \frac{\int_{-1}^1 f(x)dx}{2} = \frac{\int_{-1}^1 \sqrt{1-x}dx}{2} = 0.9428090,$$

$$d_1 = \frac{\langle f, P_1 \rangle}{\langle P_1, P_1 \rangle} = \frac{\int_{-1}^1 f(x)P_1(x)dx}{\int_{-1}^1 P_1^2(x)dx} = \frac{\int_{-1}^1 xf(x)dx}{\int_{-1}^1 x^2 dx} = \frac{\int_{-1}^1 x\sqrt{1-x}dx}{2/3} = -0.56568,$$

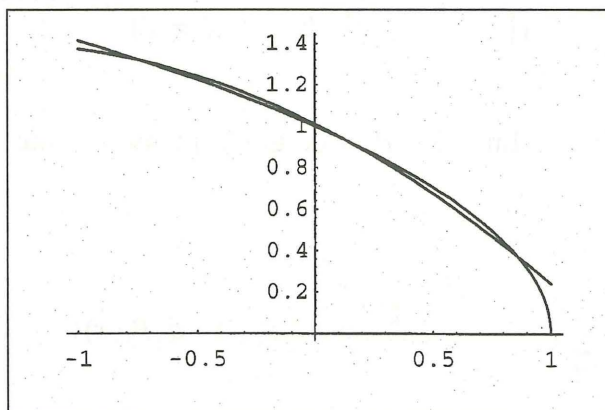
$$d_2 = \frac{\langle f, P_2 \rangle}{\langle P_2, P_2 \rangle} = \frac{\int_{-1}^1 f(x)P_2(x)dx}{\int_{-1}^1 P_2^2(x)dx} = \frac{\int_{-1}^1 [(3/2)(x^2 - 1/3)]f(x)dx}{\int_{-1}^1 [(3/2)(x^2 - 1/3)]^2 dx} =$$

$$= \frac{\int_{-1}^1 [(3/2)(x^2 - 1/3)]\sqrt{1-x}dx}{0.4} = -0.134998,$$

entonces el polinomio de grado 2 que buscamos toma la forma:

$$P(x) = 0.942809 - 0.56568x - 0.134998[(3/2)(x^2 - 1/3)]$$

Se muestra en la gráfica siguiente a las funciones $f(x)$ y $P(x)$:



Gráfica 9

4.9 Cuadratura Gaussiana

La Cuadratura Gaussiana es un método que se emplea para calcular integrales en forma aproximada haciendo uso de polinomios ortogonales.

En los métodos vistos anteriormente de Integración Numérica, la integral queda definida a través de valores funcionales, y para mejorar la aproximación se puede elegir de forma apropiada el conjunto de puntos.

Una de estas formas consiste en elegir los puntos como las raíces de los polinomios ortogonales, lo cual nos conduce al método de Cuadratura Gaussiana.

Para una función arbitraria $f(x)$ definida en el intervalo $[-1, 1]$, se tiene que su aproximación a través del polinomio interpolante de Lagrange, basado en los puntos x_1, x_2, \dots, x_n , raíces del polinomio de Legendre de grado n , será:

$$f(x) = \sum_{i=1}^n f(x_i) L_i(x) + \frac{f^{(n)}(\xi(x))}{n!} \prod_{i=1}^n (x - x_i)$$

De manera que al calcular la integral, tenemos:

$$\begin{aligned} \int_{-1}^1 f(x) dx &= \sum_{i=1}^n f(x_i) \int_{-1}^1 L_i(x) dx + \frac{1}{n!} \int_{-1}^1 f^{(n)}(\xi(x)) \prod_{i=1}^n (x - x_i) dx = \\ &= \sum_{i=1}^n w_i f(x_i) + Error \end{aligned}$$

$$\text{en donde } w_i = \int_{-1}^1 L_i(x) dx$$

y el término del error esta dado por:

$$Error = \frac{1}{n!} \int_{-1}^1 f^{(n)}(\xi(x)) \prod_{i=1}^n (x - x_i) dx$$

El cual se puede presentar en la forma:

$Error \approx C f^{(2n+2)}(\xi)$, donde ξ está entre x_1, x_2, \dots, x_n y C es una constante. Ver en Stoer-Bulirsch [23].

Por lo tanto el valor aproximado de la integral estará dado por:

$$\int_a^b f(x) dx \approx \sum_{i=1}^n w_i f(x_i)$$

La Cuadratura Gaussiana se preocupa en escoger los puntos de evaluación de una manera óptima. El procedimiento de escoger los valores x_1, x_2, \dots, x_n en el intervalo $[-1, 1]$ y las constantes w_1, w_2, \dots, w_n se basa en la idea de minimizar el error obtenido. Ver en Stoer-Bulirsch [23].

Como los polinomios de Legendre están definidos en el intervalo $[-1, 1]$, al calcular la integral en un intervalo $[a, b]$:

$$\int_a^b f(x) dx$$

es necesario hacer un cambio de variable de la forma:

$$x = \frac{(b-a)z + b + a}{2}, \quad dx = \frac{(b-a)}{2} dz$$

de manera que la integral queda expresada como:

$$\begin{aligned} \int_a^b f(x) dx &= \int_{-1}^1 f\left[\frac{(b-a)z + b + a}{2}\right] \left(\frac{b-a}{2}\right) dz = \\ &= \int_{-1}^1 g(z) \left(\frac{b-a}{2}\right) dz \end{aligned}$$

con

$$g(z) = f\left[\frac{(b-a)z + b + a}{2}\right]$$

Por lo tanto, se tiene:

$$\int_a^b f(x) dx \approx \left(\frac{b-a}{2}\right) \sum_{k=1}^n w_k g(z_k)$$

en donde los valores z_k son las raíces del polinomio de Legendre.

Comentario:

Si $f(x)$ es un polinomio de grado menor o igual que $(2n-1)$ entonces la regla de la Cuadratura Gaussiana es exacta.

Supongamos que $f(x)$ es un polinomio de grado menor o igual que $(2n-1)$ y que se va a integrar en $[-1, 1]$.

Si $P_n(x)$ es el polinomio de Legendre de grado n , entonces $f(x)$ se puede expresar como:

$$f(x) = c(x)P_n(x) + r(x)\dots(1)$$

en donde $c(x)$ y $r(x)$ son polinomios de grado menor que n y por las propiedades de los polinomios ortogonales, se tiene:

$$\begin{aligned} c(x) &= \sum_{i=0}^{n-1} c_i P_i(x) \\ \Rightarrow \int_{-1}^1 f(x) dx &= \int_{-1}^1 c(x) P_n(x) dx + \int_{-1}^1 r(x) dx = \\ &= \langle c(x), P_n(x) \rangle + \int_{-1}^1 r(x) dx = \int_{-1}^1 r(x) dx \end{aligned}$$

Por otro lado como x_i es raíz de $P_n(x)$, reemplazando en (1) tenemos:

$$f(x_i) = r(x_i)$$

Además $r(x)$ se puede expresar a través del polinomio de interpolación de Lagrange en forma exacta como:

$$r(x) = \sum_{i=1}^n r(x_i) L_i(x) + \frac{r^{(n)}(\xi(x))}{n!} \prod_{i=1}^n (x - x_i)$$

y como

$$r^{(n)}(x) = 0,$$

se tiene,

$$r(x) = \sum_{i=1}^n r(x_i) L_i(x)$$

Por lo tanto se tiene:

$$\begin{aligned} \int_{-1}^1 f(x) dx &= \int_{-1}^1 r(x) dx = \\ &= \int_{-1}^1 \sum_{i=1}^n r(x_i) L_i(x) dx = \sum_{i=1}^n r(x_i) \int_{-1}^1 L_i(x) dx = \sum_{i=1}^n f(x_i) w_i \end{aligned}$$

Donde los valores x_i son las raíces o ceros de $P_n(x)$, polinomio de Legendre de orden n , y los valores w_i están dados por:

$$w_i = \int_{-1}^1 L_i(x) dx$$

Tabla para las raíces del Polinomio $P_n(x)$ de grado n de Legendre y para las constantes w_i .

n	Raíces($\pm x_i$)	coeficientes(w_i)
2	0.5773502692	1
3	0 0.7745966692	0.8888888888888888 0.5555555555555555
4	0.3399810436 0.8611363116	0.652145155 0.3478548451
5	0 0.5384693101 0.9061798459	0.568888888888 0.4786286705 0.2369268850
6	0.238619186 0.661209387 0.932469514	0.467913935 0.360761573 0.171324492
8	0.183434642 0.525532410 0.796666478 0.960289857	0.362683783 0.313706646 0.222381034 0.101228536
10	0.148874339 0.433395394 0.679409568 0.865063367 0.973906528	0.295524225 0.269266719 0.219086363 0.149451349 0.066671344

Tabla 26

Aplicaciones:

Ejemplo 1:

Use la Cuadratura Gaussiana con $n = 3$ puntos para calcular aproximadamente:

$$\int_{-1}^1 \frac{2}{1+2x^2} dx.$$

Solución:

Aplicando la fórmula de la Cuadratura Gaussiana con $n = 3$, tenemos:

$$\begin{aligned} \int_{-1}^1 \frac{2}{1+2x^2} dx &\approx \sum_{i=1}^3 f(x_i)w_i = \\ &= f(0)w_1 + f(0.7745966692)w_2 + f(-0.7745966692)w_3 \\ &= 2.7878787878 \end{aligned}$$

y el valor exacto es 2.702043.

Ejemplo 2:

Use la cuadratura gaussiana con $n = 4$ puntos para calcular aproximadamente:

$$\int_0^3 \frac{e^x \operatorname{sen} x}{1+x^2} dx$$

Solución:

En este caso el intervalo es diferente del intervalo $[-1, 1]$, luego es necesario hacer el cambio de variable:

$$\begin{aligned} [0, 3] &\xrightarrow{h} [-1, 1]; \\ z = h(x) &= \frac{2x - (a+b)}{b-a} \Rightarrow z = \frac{2x-3}{3} \end{aligned}$$

Por lo tanto, aplicando la fórmula de la Cuadratura Gaussiana, tendremos:

$$\int_0^3 \frac{e^x \operatorname{sen} x}{1+x^2} dx \approx \frac{3}{2} \sum_{i=1}^4 g(z_i) w_i = 2.881695213$$

donde

$$g(z) = \frac{e^{(3z+3)/2} \operatorname{sen} \left(\frac{3z+3}{2} \right)}{1 + \left(\frac{3z+3}{2} \right)^2},$$

y los valores z_i , y w_i :

$$z_1 = 0.3399810436, z_2 = -z_1, z_3 = 0.8611363116, z_4 = -z_3$$

$$w_1 = 0.652145155, w_2 = -w_1, w_3 = 0.3478548451, w_4 = -w_3$$

El valor exacto es 2.881637.

Ejemplo 3:

Use la Cuadratura Gaussiana con $n = 3$ puntos para calcular aproximadamente:

$$\int_{-1}^1 \sqrt{1-x^2} dx$$

Solución:

Utilicemos la fórmula de la Cuadratura Gaussiana:

$$\begin{aligned} \int_{-1}^1 \sqrt{1-x^2} dx &\approx \sum_{i=1}^3 f(x_i) w_i \\ &= f(0) w_1 + f(0.7745966692) w_2 + f(-0.7745966692) w_3 \\ &= 1.591617 \end{aligned}$$

El valor exacto es 1.570796, lográndose una precisión de un decimal exacto, habiéndose empleado en el cálculo sólo tres puntos.

Ejemplo 4:

Use la Cuadratura Gaussiana con $n = 4$ puntos para calcular aproximadamente:

$$\int_{0.5}^{1.5} \left[\frac{1}{x} + \tan x \right] dx$$

Solución:

Empleando la fórmula de la cuadratura gaussiana para $n = 4$:

$$\int_{0.5}^{1.5} \left[\frac{1}{x} + \tan x \right] dx \approx \sum_{i=1}^4 f(x_i) w_i.$$

$$[0.5, 1.5]: \xrightarrow{h} [-1, 1]: z = h(x) = \frac{2x - (a+b)}{b-a}$$

$$\Rightarrow z = \frac{2x-2}{1}$$

$$\Rightarrow x = \frac{z+2}{2}$$

$$\int_{-1}^1 f\left(\frac{z+2}{2}\right) dx \approx \frac{1}{2} \sum_{i=1}^4 g(z_i) w_i = 3.56813984$$

$$\text{donde } g(z) = \frac{2}{z+2} + \tan \left[\frac{z+2}{2} \right],$$

$$z_1 = 0.3399810436, z_2 = -z_1, z_3 = 0.8611363116, z_4 = -z_3$$

$$w_1 = 0.652145155, w_2 = -w_1, w_3 = 0.3478548451, w_4 = -w_3$$

El valor exacto es 3.61681

4.10 Ejercicios

1. Aproxime:

$$y = \frac{1+x}{1+2x+3x^2}$$

en $[0, 5]$ mediante el polinomio de interpolación de Lagrange $P_4(x)$.

- Halle $P_4(x)$, empleando puntos igualmente espaciados, con $h = 1.25$
- Grafique la distribución del error.

2. Dada

$$f(x) = (x+1)/(1+2x+3x^2),$$

y la tabla:

i	0	1	2	3
x_i	1	1.8	2.5	3
$f(x_i)$	1/3	0.1955	0.1414	2/17

- Estime $f(2.41)$, con un polinomio de Lagrange de grado 3.
- Empleando dos polinomios de Newton diferentes de grado 2, estime $f(2.41)$ y analice el error en cada caso, haciendo uso del Teorema para el Error de Interpolación.

3. Dada la siguiente tabla de población:

Año	1930	1940	1950	1960	1970	1980
Población (en miles)	123,203	131,669	150,697	179,323	203,212	226,505

Calcule aproximadamente la población correspondiente a los años:

- 1935
- 1975

Empleando en cada estimación un polinomio apropiado de Newton de grado 4, usando diferencias divididas.

4. a. Demuestre que:

$$f [x_0, x_1, x_2, \dots, x_n] = f^{(n+1)} (\xi(x)) / (n + 1)!$$

para algún $\xi(x)$.

- b. dada la siguiente tabla de valores:

x_i	$f(x_i)$
0.0	1.00
0.2	0.98007
0.4	0.92106
0.6	0.82534
0.8	0.69671
1.0	0.54030

Halle distintos polinomios de grado 3 (usando diferencias finitas), para calcular aproximadamente $f(0.43)$, y analizando el error en cada caso diga ¿Cuál elegiría Ud?

5. a. Analice cómo un error individual en una entrada tabular se propaga en la tabla de diferencias finitas, para esto considere los puntos:

x_i , para $i = -4, -3, -2, -1, 0, 1, 2, 3, 4$ y $f_i = 0, \forall i \neq 0, f_0 = \varepsilon$. Construya la tabla de diferencias finitas y determine la ley que gobierna el crecimiento de este error aislado.

- b. Por medio de la tabla de diferencias finitas, determine el error que se presenta en la siguiente lista de valores:

x_i	$f(x_i)$
1	1
2	4
3	9
4	16
5	26
6	36
7	49
8	64
9	81

y después de corregido el error, calcule aproximadamente por el polinomio de Newton:

(i) $f(2.3)$

(ii) $f(8.1)$

empleando un polinomio de grado 4.

6. Se quiere hallar una raíz de $f(x) = 0$ en el intervalo $[a, b]$, para lo cual se buscan tres aproximaciones x_i, x_{i+1}, x_{i+2} de dicha raíz. Luego con los puntos $(x_i, f(x_i)), (x_{i+1}, f(x_{i+1})), (x_{i+2}, f(x_{i+2}))$, se construye una parábola empleando el polinomio de interpolación de Newton, con la finalidad de hallar la siguiente aproximación para la raíz de $f(x) = 0$, la cual corresponderá a la intersección de la parábola con el eje X. Al hallar esta nueva aproximación, se usarán las tres nuevas aproximaciones: $x_{i+1}, x_{i+2}, x_{i+3}$. Halle dos iteraciones de este método, para la raíz de la ecuación:

$$e^{-x} - \cos x = 0, \text{ en el intervalo } [0, 2].$$

7. En los estudios de radiación inducida por polimerización, una fuente de rayos gama, fue empleada para dar medidas de dosis de radiación. Sin embargo, el dosaje varió con la posición del aparato, y se registraron los siguientes datos:

Posición (pulgadas)	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
Dosaje, 10^3 rad/hra.	1.90	2.39	2.71	2.98	3.20	3.22	3.20	2.98	2.74

Ajuste distintos polinomios de interpolación para los grados $n = 3, 5$ para hallar el dosaje aproximado correspondiente a 1.75 pulgs. ¿Cuál de ellos considera que se ajusta menor? ¿Porqué?

8. S. Chen y S. Saxena realizaron un estudio experimental para la emisión de Tungsteno en función de la temperatura (1973), encontrando los siguientes datos:

Temperatura ($^{\circ}\text{K}$)	300	400	500	600	700	800	900	1000	1100
e	0.024	0.035	0.046	0.058	0.067	0.083	0.097	0.111	0.125

Ajuste distintos polinomios de interpolación para los grados $n = 3, 5$ para obtener la emisión aproximada correspondiente a $t = 650^\circ\text{K}$. ¿Cuál de ellos considera que se ajusta mejor? ¿Porqué?

9. a. Construya la tabla de diferencias hacia delante a partir de la siguiente tabla de valores:

k	x_k	$f(x_k)$
0	0	1.792
1	20	1.308
2	40	0.801
3	60	0.549
4	80	0.406
5	100	0.317

por medio de la fórmulas de Newton hacia delante, escriba los polinomios de interpolación ajustados a:

i. x_1, x_2, x_3

ii. x_4, x_5, x_6

iii. x_2, x_3, x_4, x_5 .

- b. Halle una expresión aproximada del error en cada una de las fórmulas de interpolación obtenidas en la parte (a).

10. Sea la función:

$$f(x) = e^{-x} - \cos(x), \text{ definida en } [0.7, 1.6]$$

- Halle el polinomio de interpolación de la forma progresiva de Newton, de grado 3, con $h = 0.3$.
- Haga uso del polinomio obtenido en la parte (a), para hallar aproximadamente una raíz de $f(x)$, por medio de uno de los métodos para resolver una ecuación no lineal.
- Halle la raíz de $f(x)$, aplicando directamente uno de los métodos para resolver una ecuación no lineal.
- Compare los resultados obtenidos en los ítems (b) y (c).

11. Dada la siguiente tabla de valores:

i	0	1	2	3	4	5	6
x_i	-1.5	-1	-0.5	0	0.5	1	1.5
$F(x_i)$	0.0668	0.1587	0.3085	0.5	0.6915	0.8413	0.9332

que corresponden a la función $F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$

- Estime $F(0.58)$, por medio de un polinomio de Lagrange de grado 2 y halle una cota para el error.
- Halle dos polinomios diferentes de grado 2, haciendo uso de la tabla de diferencias finitas, para estimar $F(0.58)$ y elija aquel que considere el más apropiado.
- Se sabe que la función $F(x)$ dada, corresponde a la función de distribución acumulada de una variable aleatoria x que tiene Distribución Normal para parámetros $\mu = 0$ y $\sigma^2 = 1$.
Por una propiedad de las probabilidades, se sabe que:

$$\text{si } x \sim N(0,1) \Rightarrow y = x^2 \sim \chi^2(1)$$

Estime la probabilidad de que $0 < y < 0.81$, esto es:

$P[0 < y < 0.81]$, haciendo uso de un polinomio de grado 4.

- Para calibrar un medidor de orificio se miden la velocidad v de un fluido y la caída de presión P . Los datos experimentales se muestran en la siguiente tabla:

v	3.83	4.17	4.97	6.06	6.71	7.17	7.51	7.98
P	30	35.5	50.5	75	92	105	115	130

- Calcular aproximadamente el valor de $P(6.45)$, haciendo uso de polinomios de Lagrange de grado $n = 2, 3, 4$.
 - Calcular aproximadamente el valor de la presión correspondiente a la velocidad de 4.4, haciendo uso de los polinomios de Newton de grados: $n = 5, 6, 7$.
- Para un circuito con un voltaje impuesto $s(t)$ e inductancia L , la 1ª. Ley de Kirchhoff da la relación:

$$\varepsilon = L di/dt + Ri$$

donde R es la resistencia del circuito e i la corriente. Suponga que medimos la corriente para varios valores de t y obtenemos:

t	1	1.01	1.02	1.03	1.04
i	3.1	3.12	3.14	3.18	3.24

donde t se mide en segundos, i en amperes, la inductancia L es una constante 0.98 henries, y la resistencia es 0.142 ohm. Aproxime el voltaje ε en los valores $t = 1, 1.01, 1.02, 1.03, 1.04$; usando las fórmulas de 3 puntos.

14. A partir de la siguiente tabla:

x_i	1.2	1.3	1.4	1.5	1.6
$f(x_i) = \sinh(x_i)$	1.5095	1.6984	1.9043	2.1293	2.3756

- a. Hallar $f'(1.4)$, empleando las fórmulas de 3 puntos. Hallar también $f''(1.4)$. Compare sus resultados con los resultados de $f'(1.4) = 2.1509$ y $f''(1.4) = \cosh(1.4) = 1.9043$, que son correctos hasta las cifras dadas.
- b. Emplee la regla del Trapecio para hallar.

$$\int_{1.2}^{1.6} \sinh(x) dx \text{ y estime una cota para el error.}$$

- c. Emplee la regla de Simpson para hallar:

$$\int_{1.2}^{1.6} \sinh(x) dx \text{ y estime una cota para el error.}$$

- d. Emplee el método de Cuadratura Gaussiana para $n = 3$, y calcule aproximadamente:

$$\int_{1.2}^{1.6} \sinh(x) dx$$

15. Las fuerzas del viento ejercidas sobre el mástil de un velero de carreras varían en función de la distancia sobre la cubierta del bote.

- a. Dada la siguiente tabla:

z_i (distancia)	0	3	6	9	12	15	18	21	24	27	30
$f(z_i)$ (fuerza)	0	61.4	73.13	70.56	63.43	55.18	47.1	39.83	33.42	27.89	23.2

Emplee el método del Trapecio para calcular aproximadamente la integral y una cota superior para el error.

- b. Emplee el método de Cuadratura Gaussiana con $n = 4$ puntos para calcular aproximadamente la integral.

16. a. Emplee un cambio de variable, $u = x^{-1}$, en la integral:

$$\int_a^{\infty} f(x) dx, \quad a > 0$$

con la finalidad de poder aplicar los métodos de integración numérica.

- b. Suponga que un cuerpo de masa m viaja verticalmente hacia arriba empezando en la superficie de la tierra $x = R$. Si se desprecian toda las fuerzas de resistencia excepto la gravedad, entonces la velocidad de escape v está dada por:

$$v^2 = 2gR \int_1^{\infty} z^{-2} dz \quad \text{donde } z = x/R,$$

y g es la fuerza del campo gravitacional en la superficie de la tierra. Si $g = 0.00609$ millas/seg² y $R = 3,960$ millas, aproxime la velocidad de escape v , usando el resultado de la parte (a) y la regla de Simpson con $h = 0.1$. Analice el error.

17. Calcule la siguiente integral impropia:

$$\int_0^2 (e^x / \sqrt{|x^2 - 1|}) dx$$

- a. Emplee la regla del trapecio, con $N = 12$ intervalos, utilice la fórmula del error para hallar una cota para el error.
- b. Emplee la regla de Simpson, con $N = 12$ intervalos, calcule una cota para el error como en (a).

Sugerencia: Para $x = 1$, emplee valores próximos tanto a la izquierda, como a la derecha de $x = 1$; es decir: $1 \pm \varepsilon$, para un ε pequeño.

18. Calcule aproximadamente:

$$\int_{-3}^1 (x \operatorname{sen}(x+4) - \ln(x+5)) dx,$$

empleando la Cuadratura Gaussiana con $n = 4$.

19. Calcule aproximadamente la integral:

$$\int_0^{0.5} x e^x dx$$

- Haciendo uso del método de Simpson, con $N = 5$. Calcule el error aproximadamente.
 - Haciendo uso del método del Trapecio, con $N = 5$. Calcule el error aproximadamente.
20. a. Extienda el método de mínimos cuadrados al caso en que y es función lineal de 2 variables, esto es:

$$y = F(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2$$

y halle las ecuaciones normales correspondientes, si se sabe que $(x_{1i}, x_{2i}, y_i)_{(1 \leq i \leq n)}$ es la muestra de datos, de tamaño n .

- Un estudio de ingeniería mecánica indica que el fluido, a través de un tubo es función del diámetro del tubo y de su pendiente. En el siguiente cuadro se muestran los datos experimentales:

Experimento	Diámetro, pies	Pendiente pies/pies	Flujo pies 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
9	3	0.05	200

Se sabe que el flujo (Q) está relacionado con la pendiente (S) y el diámetro (D) a través de la siguiente expresión

$$Q = a_0 D^{a_1} S^{a_2}$$

Emplee las fórmulas obtenidas en (a) para aplicarlas a la ecuación:

$$\ln Q = \ln a_0 + a_1 \ln D + a_2 \ln S$$

y estime el valor de Q en un tubo con:

$$D = 2.5 \text{ pies}$$

$$S = 0.025 \text{ pies/pie}$$

21. Dados los datos:

x_i	y_i
0.2	0.050446
0.3	0.098426
0.6	0.33277
0.9	0.72660
1.1	1.0972
1.3	1.5697
1.4	1.8487
1.6	2.5015

- Hallar una aproximación polinómica de mínimos cuadrados de grado 2.
- Hallar una aproximación de mínimos cuadrados de la forma bx^a .
- Comparar ambos ajustes y decir ¿Cuál sería mejor?

22. Dados los siguientes datos:

x_i	0	0.5	1	1.5	2	2.5
y_i	5.02	5.21	6.49	9.54	16.02	24.53

Utilídense mínimos cuadrados para ajustarlos:

- A una parábola.
- A una curva de la forma: $y = ae^{bx}$

¿Qué ajuste considera mejor? ¿Porqué?

23. A partir de la siguiente tabla de valores:

x_i	0	$\pi/9$	$2\pi/9$	$\pi/3$	$4\pi/9$	$5\pi/9$	$2\pi/3$	$7\pi/9$	$8\pi/9$	π
$f(x_i) = \cos x_i$	1	0.9397	0.766	0.5	0.1736	-0.1736	-0.5	-0.766	-0.9397	-1

- Construir la tabla de diferencias finitas y emplear un polinomio de Newton de grado 3 para estimar $f(7\pi/8)$. Analizar el error.

- b. Emplear el método de mínimos cuadrados para hallar un polinomio de grado 2, en el intervalo $[\pi/2, \pi]$, que se aproxime a la función dada. Estime $f(7\pi/8)$, compare este resultado con el de la parte (a).
- c. Calcule aproximadamente la integral:

$$\int_{4\pi/9}^{\pi} f(x) dx$$

empleando la regla de Simpson. Analizar el error.

24. Los siguientes datos muestran la concentración de oxígeno disuelto (y) en función de la temperatura (x) para el caso en que la concentración de cloruro es igual a 20000 mg/L.

i	0	1	2	3	4	5	6
x_i	5	10	15	20	25	30	35
$y_i=f(x_i)$	10.5	9.2	8.2	7.4	6.7	6.1	5.6

- a. Hallar el valor aproximado de $f'(x_i)$, para $i = 0, 1, 2, 3, 4, 5, 6$; empleando las fórmulas de tres puntos.
- b. Basándose en los valores hallados en (a), estimar $f''(x_2)$.

- c. Calcular aproximadamente el valor de $\int_5^{35} f(x) dx$, por medio de

la regla de Simpson.

- d. Construya un polinomio de Newton de grado 5 y estime $f(16)$. Calcule una cota superior para el error de esta estimación.
- e. Emplee el método de mínimos cuadrados, para hallar un ajuste de tipo cúbico para estimar $f(16)$ y calcule el error cuadrático medio.
- f. Compare el resultado de (d) con (e).

25. Haciendo uso del método de mínimos cuadrados, ajustar los datos de la siguiente tabla:

x_i	0	0.5	1	1.5	2	2.5
y_i	5.02	5.21	6.49	9.54	16.02	24.53

a la curva de la forma $f(x) = ae^x + be^{-x}$.

26. Dada la siguiente tabla:

x_i	1	2	3	4	5	6	7	8	9	10	11
$f(x_i)$	0	0.6	1.77	1.92	3.31	3.52	4.59	5.31	5.79	7.06	7.17

Hallar la aproximación de mínimos cuadrados, tomando:

$$F(x) = c_1 + c_2x + c_3\text{sen}[50(x-1)].$$

27. a. Emplear la aproximación de mínimos cuadrados por polinomios ortogonales de Legendre para hallar un polinomio de grado 2 a fin de minimizar la integral:

$$\int_{-1}^1 [f(x) - P(x)]^2 dx, \text{ donde } f(x) = \cos x$$

- b. Analice las gráficas de $P(x)$ y $f(x)$.

Ecuaciones Diferenciales Ordinarias

Las ecuaciones diferenciales se presentan con mucha frecuencia en el campo de las Ciencias e Ingeniería, como modelos matemáticos que corresponden a las leyes que aparecen en las áreas de mecánica, dinámica, electrónica, termodinámica, etc. Muchas de estas ecuaciones no pueden ser resueltas por los métodos clásicos.

Por lo tanto los métodos numéricos son la única alternativa para estos casos, así como también son útiles en general, ya que simplifican el proceso de determinar una solución aproximada.

El problema clásico del valor inicial es encontrar una función $y = y(x)$ que satisfaga la ecuación diferencial de 1er. orden:

$$y' = f(x, y)$$

que tome el valor inicial $y(x_0) = y_0$, y el objetivo en esta sección será hallar la solución aproximada para esta ecuación.

5.1 El Método de Euler

El método de Euler se usa rara vez en la práctica, sin embargo su derivación es tan simple que puede usarse para ilustrar las técnicas involucradas en la construcción de algunas de las técnicas más avanzadas, sin tener que usar la enorme cantidad de álgebra que acompaña a estas construcciones.

El objetivo del método es obtener una aproximación al problema del valor inicial bien planteado,

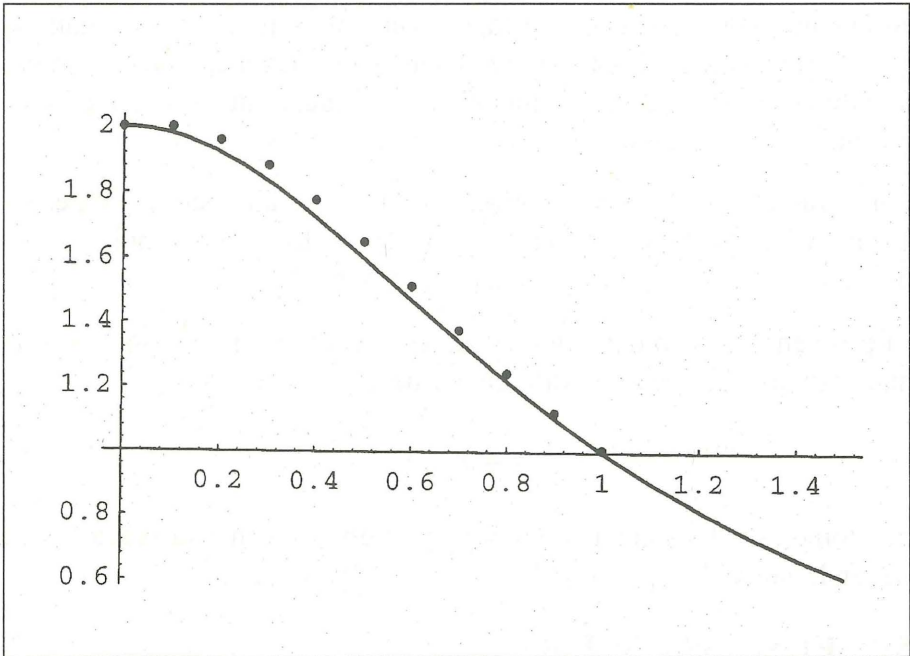
$$y' = f(x, y), \quad a \leq x \leq b, \quad y(x_0) = y_0$$

Lo que interesa es el comportamiento de la función $y(x)$, en el intervalo $[a, b]$. Para esto se define una partición de puntos:

$$x_0, x_1, \dots, x_N, \quad x_i = x_0 + ih, \text{ para } i = 0, 1, 2, \dots, N.$$

donde $h = (b - a) / N$.

El algoritmo numérico construye valores discretos aproximados y_i de la función y en $x = x_i$, $\forall i = 0, 1, 2, \dots, N$. Como se muestra en la gráfica:



Gráfica 1

la línea curva continua representa a la función $y(x)$, solución de la ecuación diferencial, y los puntos discretos a los pares (x_i, y_i) , en donde los y_i son los valores estimados de la función y en x_i .

Una de las formas de construir este algoritmo es a través de la Fórmula de Taylor para $n = 1$ en torno de x_i .

$$\Rightarrow y(x) = y(x_i) + (x - x_i) y'(x_i) + (x - x_i)^2 y''(\xi(x))/2$$

$$\text{ó } y(x) = y_i + (x - x_i) y'_i + (x - x_i)^2 y''(\xi(x))/2.$$

donde $\xi(x)$ está dada en el intervalo entre x y x_i .

Si

$$x = x_{i+1} \Rightarrow y(x_{i+1}) = y(x_i) + h y'(x_i) + h^2 y''(\xi(x_{i+1}))/2$$

$$\therefore y(x_{i+1}) \approx y(x_i) + h y'(x_i) \quad \text{pero } y'_i = y'(x_i) = f(x_i, y_i) \\ y_{i+1} \approx y_i + h y'_i$$

luego,

$$y_{i+1} \approx y_i + h f(x_i, y_i), \quad i = 0, 1, 2, \dots, N$$

que recibe el nombre de la Fórmula de Euler, con un error:

$$E = h^2 y''(\xi(x_{i+1})) / 2,$$

donde $\xi(x_{i+1})$ está entre x_i y x_{i+1} .

Algoritmo de Euler

Objetivo:

Hallar una solución aproximada de la solución del problema de valor inicial:

$$y' = f(x, y) \quad a \leq x \leq b$$

con condición inicial $y(x_0) = y_0$, en $(N + 1)$ puntos uniformemente espaciados en el intervalo $[a, b]$:

Paso 1:

Tomar:

$$h = (b - a) / N$$

$$x_0 = a;$$

$$y_0 = y(x_0);$$

Paso 2: Hacer para $i = 0, 1, 2, 3, \dots, N$.

$$x_{i+1} = x_0 + ih$$

$$y(x_{i+1}) \approx y_{i+1} = y_i + h f(x_i, y_i)$$

Aplicaciones:**Ejemplo 1:**

Aplicar el método de Euler a la ecuación diferencial:

$$y' = f(x, y) = -xy^2, \text{ si } y(0) = 2$$

para calcular aproximadamente $y(1)$ con $h = 0.1$.

Solución:

Se tiene que la Fórmula de Euler es:

$$y_{i+1} = y_i + hf(x_i, y_i), \quad \forall \quad i = 0, 1, 2, \dots, 10.$$

$$y_{i+1} = y_i + h(-x_i y_i^2), \quad \text{con } h = 0.1$$

Para $i=0$, $x_0 = 0$, $y_0 = 2$:

$$y(0.1) \approx y_1 = 2 + 0.1(0) = 2$$

Para $i=1$, $x_1 = 0.1$, $y_1 = 2$:

$$y(0.2) \approx y_2 = 2 - 0.1(0.1)(4) = 1.96$$

y así sucesivamente se calculan los siguientes valores:

i	1	2	3	4	5
x_i	0.1	0.2	0.3	0.4	0.5
y_i	2	1.96	1.883168	1.776778349	1.650500697

i	6	7	8	9	10
x_i	0.6	0.7	0.8	0.9	1
y_i	1.514293069	1.376708059	1.244035303	1.120225396	1.007283952

El valor pedido es $y(1) \approx 1.007283952$.

Ejemplo 2:

Resolver la ecuación diferencial:

$$y' = f(x, y) = -xy^2 ; \quad \text{si } y(2) = 1;$$

Para obtener aproximadamente $y(3)$, con $h= 0.1$, $h = 0.05$

Solución:

Método de Euler:

$$y_{i+1} = y_i + hf(x_i, y_i) , \quad \forall i = 0, 1, 2, \dots, N.$$

(a) Para $h=0.1$

Para $i = 0$,

$$y_1 = y_0 + 0.1f(x_0, y_0) = 1 + 0.1(-2(1)^2)=0.8$$

Para $i = 1$,

$$y_2 = y_1 + 0.1f(x_1, y_1) = 0.8 + 0.1(-2.1(0.8)^2) = 0.6656,$$

y así sucesivamente, los resultados se muestran en la siguiente tabla:

i	0	1	2	3	4	5	6	7	8	9	10
x_i	2	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3
y_i	1	0.8	0.6656	0.5681	0.4938	0.4353	0.3880	0.3488	0.3160	0.2880	0.2640
Valor Exacto	1	0.8299	0.7042	0.6079	0.5319	0.4706	0.4202	0.3781	0.3425	0.3120	0.2857

(b) Para $h=0.05$

i	0	2	4	6	8	10	12	14	16	18	20
x_i	2	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3
y_i	1	0.8170	0.6869	0.5897	0.5142	0.4539	0.4048	0.3640	0.3297	0.3003	0.2751
Valor Exacto	1	0.8299	0.7042	0.6079	0.5319	0.4706	0.4202	0.3781	0.3425	0.3120	0.2857

De los resultados obtenidos se puede apreciar que estos métodos involucran dos tipos de errores; uno propiamente del método y el otro que proviene del error de redondeo que se comete en los cálculos que se realizan.

El método de Euler se conoce como el método de Taylor de orden 1. A continuación veremos el método de Taylor de orden superior.

5.2 Método de Taylor de Orden Superior

Sea la ecuación diferencial

$$y' = f(x, y), \quad \text{para } a \leq x \leq b.$$

con condición inicial $y(x_0) = y_0$ con $x_0 = a$.

Sea la partición

$$x_i = x_0 + ih, \quad \forall i = 0, 1, 2, \dots, N; \quad \text{con } x_0 = a \quad \text{y} \quad x_N = b.$$

Usando el desarrollo de Taylor en torno de x_i , tenemos:

$$y(x) = \sum_{j=0}^{+\infty} \frac{y^{(j)}(x_i)(x-x_i)^j}{j!} \quad \therefore$$

si $x = x_{i+1}$ entonces $y(x_{i+1}) = y_{i+1} = \sum_{j=0}^{+\infty} \frac{y^{(j)}(x_i)h^j}{j!}$

donde $y^{(j)}$ son las derivadas totales de $f(x, y)$, luego:

$$y' = f(x, y)$$

Luego la segunda derivada total será:

$$y'' = f_x + f_y y'$$

y la tercera derivada total:

$$y''' = f_{.xx} + 2f f_{xy} + f^2 f_{yy} + (f_y)^2 f + f_x f_y$$

En general, la j -ésima derivada total de $y(x)$:

$$y^{(j)} = f^{(j-1)}, \quad j=1, 2, \dots$$

Si se trunca en n la serie de Taylor para la función $y(x)$, obtenemos una aproximación de $y(x_{i+1})$, esto es:

$$y(x_{i+1}) \approx y_{i+1} = \sum_{j=0}^n y^{(j)}(x_i) \frac{h^j}{j!} = \sum_{j=0}^n y_i^{(j)} \frac{h^j}{j!}$$

que recibe el nombre de la fórmula de Taylor de orden n y el error está dado por:

$$Error = y^{(n+1)}(\xi) \frac{h^{(n+1)}}{(n+1)!}, \quad \xi \in]x_i, x_{i+1}[$$

Aplicaciones:

Ejemplo 1:

Resolver aproximadamente:

$$y' = f(x, y) = x^2 - y^2, \quad \text{si } y(-1) = 1,$$

para $y(0)$, con $h = 0.2$, empleando el método de Taylor para $n=2$.

Solución:

Sea la partición:

$$x_0 = -1, \quad x_1 = -0.8, \quad x_2 = -0.6, \quad x_3 = -0.4, \quad x_4 = -0.2, \quad x_5 = 0,$$

y el método de Taylor de orden 2 es:

$$y_{i+1} = y_i + hy'_i + h^2y''_i, \quad \forall i = 0, 1, 2, \dots, 5.$$

$$y_{i+1} = y_i + hf(x_i, y_i) + h^2 f'(x_i, y_i) / 2, \quad i = 0, 1, 2, \dots, 5.$$

Pero la segunda derivada total de $y(x)$ es:

$$\begin{aligned} y''(x) &= f''(x, y) = f''_x + f''_y y' = 2x - 2yy' = \\ &= 2x - 2y(x^2 - y^2) = \\ &= 2x - 2yx^2 + 2y^3 \end{aligned}$$

Luego:

$$y''(x_{i+1}) \approx y''(x_i) + 0.2(x_i^2 - y_i^2) + 0.2^2(2x_i - 2y_i x_i^2 + 2y_i^3) / 2;$$

para $i = 0, 1, 2, \dots, N$

Para $i = 0, \quad x_0 = -1, \quad y_0 = 1$

$$y_1 = y_0 + 0.2y'_0 + 0.2^2 y''_0 / 2 = 0.96$$

Para $i=1 \quad x_1 = -0.8, \quad y_1 = 0.96$

$$y_2 = y_1 + 0.2y'_1 + 0.2^2 y''_1 / 2 = 0.88249344$$

y así sucesivamente se obtienen los demás valores:

i	1	2	3	4	5
x_i	-0.8	-0.6	-0.4	-0.2	0
y_i	0.96	0.88249344	0.789517847	0.695482719	0.611086799

Ejemplo 2

Resolver la ecuación diferencial:

$$y' = f(x, y) = -xy^2; \quad \text{si } y(2) = 1$$

Para $y(3)$ aproximadamente usando el método de Taylor para $n = 2$, con $h = 0.1$ y $h = 0.05$.

Solución:Método de Taylor con $n = 2$:

$$y_{i+1} = y_i + y'_i h + y''_i h^2 / 2 \quad \forall i = 0, 1, 2, \dots, N.$$

Luego:

$$y'' = f_{xx} + f_{yy} y' = -y^2 - 2xy(-xy^2) = -y^2 + 2x^2y^3;$$

Por lo tanto:

$$y_{i+1} = y_i + h(-x_i y_i^2) + (h^2 / 2) (-y_i^2 + 2x_i^2 y_i^3); \quad \forall i = 0, 1, 2, \dots, N.$$

(a) Para $h=0.1$ Para $i=0$, $x_0 = 2$, $y_0 = 1$:

$$y_1 = 1 + 0.1(-2(1)^2) + (0.1^2/2)(-1^2 + 2(2)^2(1)^3) = 0.835$$

Para $i = 1$, $x_1 = 2.1$, $y_1 = 0.835$:

$$y_2 = 0.835 + 0.1(-2.1(0.835)^2) + (0.1^2/2)(-0.835^2 + 2(2.1)^2(0.835)^3) \\ = 0.7108$$

y así sucesivamente, se obtienen los siguientes resultados:

i	0	1	2	3	4	5	6	7	8	9	10
x_i	2	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3
y_i	1	0.835	0.7108	0.6145	0.5380	0.4761	0.4250	0.3823	0.3462	0.3153	0.2886
Valor Exacto	1	0.8299	0.7042	0.6079	0.5319	0.4706	0.4202	0.3781	0.3425	0.3120	0.2857

(b) Para $h=0.05$, se tiene:

i	0	2	4	6	8	10	12	14	16	18	20
x_i	2	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3
y_i	1	0.8310	0.7056	0.6093	0.5332	0.4718	0.4212	0.3790	0.3433	0.3127	0.2863
Valor Exacto	1	0.8299	0.7042	0.6079	0.5319	0.4706	0.4202	0.3781	0.3425	0.3120	0.2857

5.3 Métodos de Runge-Kutta

Como se puede apreciar de los dos métodos anteriores; el método de Euler resulta ser muy impreciso, a no ser si el paso h , fuese muy pequeño, y el método de Taylor de orden superior logra mejorar la precisión, pero en cambio la cantidad de cálculos que se necesita realizar para obtener las derivadas totales es muy grande, lo que los hace poco eficientes. Los métodos de Runge-Kutta surgen como una alternativa a la precisión de los métodos de Taylor y se aplican realizando un número menor de operaciones.

Consideremos a la ecuación diferencial:

$$y' = f(x, y), \quad \text{con } x \in [a, b],$$

con condición inicial:

$$y(x_0) = y_0,$$

y sea la partición:

$$x_i = x_0 + ih, \quad i = 0, 1, 2, \dots, N; \quad \text{con } x_0 = a, \quad x_N = b.$$

Todos los métodos de Runge-Kutta realizan el siguiente paso:

$$\begin{aligned} \int_{x_i}^{x_{i+1}} y' dx &= \int_{x_i}^{x_{i+1}} f(x, y) dx \\ \Rightarrow y(x_{i+1}) - y(x_i) &= \int_{x_i}^{x_{i+1}} f(x, y) dx \\ y(x_{i+1}) &= y(x_i) + \int_{x_i}^{x_{i+1}} f(x, y) dx, \quad \forall i = 0, 1, 2, \dots, N \\ y_{i+1} &= y_i + \int_{x_i}^{x_{i+1}} f(x, y) dx, \quad \forall i = 0, 1, 2, \dots, N \end{aligned}$$

y el trabajo se reduce al cálculo aproximado de la integral, por medio de uno de los métodos de integración.

Método de Runge-Kutta de orden 2

Para el cálculo de la integral se emplea la regla del Trapecio, esto es:

$$\int_{x_i}^{x_{i+1}} f(x, y) dx \approx \frac{h}{2} [f(x_i, y_i) + f(x_{i+1}, y_{i+1})]$$

$$\Rightarrow y_{i+1} = y_i + \frac{h}{2} [f(x_i, y_i) + f(x_{i+1}, y_{i+1})]$$

en donde y_{i+1} , también aparece incluida en su propio cálculo, por lo que es conveniente realizar una estimación de ella, y esto se hace empleando el método de Euler, esto es:

$$\hat{y}_{i+1} = y_i + h y_i' = y_i + h f(x_i, y_i)$$

$$\Rightarrow y_{i+1} = y_i + \frac{h}{2} [f(x_i, y_i) + f(x_{i+1}, \hat{y}_{i+1})]$$

y si definimos:

$$k_1 = h f(x_i, y_i);$$

$$k_2 = h f(x_{i+1}, y_i + h f(x_i, y_i));$$

$$\Rightarrow y_{i+1} = y_i + [k_1 + k_2] / 2, \quad \forall i = 0, 1, 2, 3, \dots, N$$

El error está dado aproximadamente por la expresión:

$$E \approx -h^3 f^{(2)}(\xi) / 12 \quad \text{donde } \xi \text{ está entre } x_i \text{ y } x_{i+1}$$

Método de Runge-Kutta de 4° orden

En este caso, se calcula la integral por medio de la regla de 1/3 de Simpson:

$$\int_{x_i}^{x_{i+1}} f(x, y) dx \approx \frac{1}{6} [k_1 + 2k_2 + 2k_3 + k_4]; \quad \forall i = 0, 1, 2, 3, \dots, N.$$

$$\therefore y_{i+1} = y_i + \frac{1}{6} [k_1 + 2k_2 + 2k_3 + k_4],$$

donde:

$$k_1 = hf(x_i, y_i);$$

$$k_2 = hf\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}\right);$$

$$k_3 = hf\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}\right);$$

$$k_4 = hf(x_{i+1}, y_i + k_3);$$

y el error está dado por:

$$E \approx Ch^5 y^{(5)}(\xi),$$

donde ξ está entre x_i y x_{i+1} , y C es una constante.

Algoritmo de Runge-Kutta de orden Cuatro

Objetivo: Determinar los valores aproximados de la función $y(x)$ en los puntos seleccionados

$$x_i = x_0 + ih; \text{ para } i = 0, 1, 2, \dots, N$$

con

$$x_0 = a;$$

$$x_N = b;$$

y

$$h = \frac{b-a}{N}$$

Paso 1. Para $i = 0, 1, 2, \dots, N$

Tomar

$$y(x_{i+1}) \approx y_{i+1} = y_i + \frac{1}{6} [k_1 + 2k_2 + 2k_3 + k_4]$$

con

$$k_1 = hf(x_i, y_i)$$

$$k_2 = hf\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}\right)$$

$$k_3 = hf\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}\right)$$

$$k_4 = hf(x_{i+1}, y_i + k_3)$$

Aplicaciones:

Ejemplo 1:

Resolver la ecuación diferencial:

$$y' = f(x, y) = x^2 - y^2, \quad \text{si } y(-1) = 1, \quad \text{para } y(-0.6), \quad \text{con } h=0.2.$$

Empleando:

- Runge-Kutta de orden 2.
- Runge-Kutta de orden 4.

Solución:

- Runge-Kutta de orden 2

Los puntos seleccionados:

$$x_0 = -1, \quad x_1 = -0.8, \quad x_2 = -0.6,$$

Algoritmo:

$$y(x_{i+1}) \approx y_{i+1} = y_i + 1/2 [k_1 + k_2]; \quad i = 0, 1, 2$$

donde:

$$k_1 = h f(x_i, y_i)$$

$$k_2 = h f(x_{i+1}, y_i + k_1)$$

Para $i=0$, $x_0 = -1$, $y_0 = 1$,

$$y(x_1) = y(-0.8) \approx y_1 = y_0 + 1/2 [k_1 + k_2],$$

donde:

$$k_1 = h f(x_0, y_0) = 0.2f(-1, 1) = 0$$

$$k_2 = 0.2f(x_1, y_0 + k_1) = 0.2f(-0.8, 1) = -0.072,$$

por lo tanto:

$$y(-0.8) \approx y_1 = 0.964,$$

y en forma similar se calcula:

$$y(-0.6) \approx y_2$$

Para $i = 1$, $x_1 = -0.8$, $y_1 = 0.964$:

$$y(x_2) = y(-0.6) \approx y_2 = y_1 + 1/2[k_1 + k_2]$$

donde:

$$k_1 = h f(x_1, y_1) = 0.2f(-0.8, 0.964) = -0.057859$$

$$k_2 = h f(x_2, y_1 + k_1) = 0.2f(-0.6, 0.906141) = -0.092218$$

por lo tanto,

$$y_2 \approx 0.888961$$

i	1	2
x_i	-0.8	-0.6
y_i	0.964	0.888961

b. Runge-Kutta de orden 4

Los puntos seleccionados son:

$$x_0 = -1, \quad x_1 = -0.8, \quad x_2 = -0.6$$

Algoritmo:

$$y(x_{i+1}) \approx y_{i+1} = y_i + 1/6[k_1 + 2k_2 + 2k_3 + k_4], \quad i = 0, 1, 2.$$

donde:

$$\begin{aligned} k_1 &= h f(x_i, y_i); \\ k_2 &= h f(x_i + h/2, y_i + k_1/2); \\ k_3 &= h f(x_i + h/2, y_i + k_2/2); \\ k_4 &= h f(x_{i+1}, y_i + k_3); \end{aligned}$$

Para $i=0$, $x_0 = -1$, $y_0 = 1$:

$$y(-0.8) \approx y_1 = y_0 + (1/6)[k_1 + 2k_2 + 2k_3 + k_4];$$

donde:

$$\begin{aligned} k_1 &= h f(x_0, y_0) = 0.2f(-1, 1) = 0; \\ k_2 &= h f(x_0 + h/2, y_0 + k_1/2) = 0.2f(-0.9, 1) = -0.038 \\ k_3 &= h f(x_0 + h/2, y_0 + k_2/2) = 0.2f(-0.9, 0.981) = -0.030472, \\ k_4 &= h f(x_0 + h, y_0 + k_3) = 0.2f(-0.8, 0.969528) = -0.059997 \end{aligned}$$

por lo tanto:

$$y(-0.8) \approx y_1 = 0.967176,$$

y en forma similar se obtiene:

$$y_2 \approx y(-0.6)$$

Para $i = 1$, $x_1 = -0.8$, $y_1 = 0.967176$:

$$y(-0.6) \approx y_2 = y_1 + 1/6[k_1 + 2k_2 + 2k_3 + k_4];$$

donde:

$$\begin{aligned}k_1 &= 0.2f(-0.8, 0.967176) = -0.059086; \\k_2 &= 0.2f(-0.7, 0.937633) = -0.077831; \\k_3 &= 0.2f(-0.7, 0.928261) = -0.074334; \\k_4 &= 0.2f(-0.6, 0.892843) = -0.087434;\end{aligned}$$

por lo tanto:

$$y(-0.6) \approx 0.892035$$

Ejemplo 2:

Resolver la ecuación diferencial:

$$y' = f(x, y) = x(y+1),$$

si $y(0) = 1$, para $y(0.5)$, con $h = 0.1$.

Empleando:

- Runge-Kutta de orden 2.
- Runge-Kutta de orden 4.

Solución:

- Runge-Kutta de orden 2

Los puntos seleccionados son:

$$x_0 = 0, x_1 = 0.1, x_2 = 0.2, \dots, x_5 = 0.5$$

Método de Runge-Kutta:

$$y(x_{i+1}) \approx y_{i+1} = y_i + 1/2[k_1 + k_2]; \quad i = 0, 1, 2, \dots, 5$$

donde:

$$\begin{aligned}k_1 &= h f(x_i, y_i) \\k_2 &= h f(x_{i+1}, y_i + k_1)\end{aligned}$$

Para $i = 0$, $x_0 = 0$, $y_0 = 1$:

$$y_1 = y_0 + 1/2 [k_1 + k_2],$$

donde:

$$k_1 = h f(x_0, y_0) = 0.1f(0, 1) = 0$$

$$k_2 = 0.1f(x_1, y_0 + k_1) = 0.1f(0.1, 1) = 0.1 \times 0.1 \times 2 = 0.02,$$

$$y(0.1) \approx 1.01,$$

y en forma similar se calcula y (0.2):

Para $i = 1$, $x_1 = 0.1$, $y_1 = 1.01$;

$$y_2 = y_1 + 1/2[k_1 + k_2]$$

donde:

$$k_1 = h f(x_1, y_1) = 0.1f(0.1, 1.01) = 0.1 \times 0.1 \times (1.01 + 1) = 0.0201$$

$$k_2 = h f(x_2, y_1 + k_1) = 0.1f(0.2, 1.0301) = 0.040602$$

$$y(0.2) \approx 1.040351$$

y de esta manera se obtienen los siguientes resultados:

i	0	1	2	3	4	5
x_i	0	0.1	0.2	0.3	0.4	0.5
y_i	1	1.01	1.040351	1.09197	1.16645	1.26610
Valor Exacto	1	1.010025	1.04040268	1.09206	1.16657	1.266297

b. Runge-Kutta de Orden 4

$$y(x_{i+1}) \approx y_{i+1} = y_i + 1/6[k_1 + 2k_2 + 2k_3 + k_4], \quad i = 0, 1, 2, \dots, 5.$$

donde:

$$k_1 = h f(x_i, y_i);$$

$$k_2 = h f(x_i + h/2, y_i + k_1/2);$$

$$k_3 = h f(x_i + h/2, y_i + k_2/2);$$

$$k_4 = h f(x_{i+1}, y_i + k_3);$$

Para $i=0$, $x_0=0$, $y_0 = 1$,

$$y(0.1) \approx y_1 = y_0 + (1/6)[k_1 + 2k_2 + 2k_3 + k_4];$$

donde:

$$\begin{aligned} k_1 &= h f(x_0, y_0) &&= 0.1f(0, 1) &&= 0; \\ k_2 &= h f(x_0 + h/2, y_0 + k_1/2) &&= 0.1f(0.05, 1) &&= 0.01 \\ k_3 &= h f(x_0 + h/2, y_0 + k_2/2) &&= 0.1f(0.05, 1.005) &&= 0.010025 \\ k_4 &= h f(x_0 + h, y_0 + k_3) &&= 0.1f(0.1, 1.010025) &&= 0.02010025 \end{aligned}$$

$$y(0.1) \approx y_1 = 1.010025042,$$

y en forma similar se obtiene:

$$y_2 \approx y(0.2):$$

Para $i= 1$, $x_1 =0.1$, $y_1 = 1.010025042$:

$$y(0.2) \approx y_2 = y_1 + 1/6[k_1 + 2k_2 + 2k_3 + k_4];$$

donde:

$$\begin{aligned} k_1 &= 0.1f(0.1, 1.010025042) = 0.0201002504; \\ k_2 &= 0.03030112751 \\ k_3 &= 0.03037763409 \\ k_4 &= 0.04080805352 \end{aligned}$$

$$y(0.2) \approx y_2 = 1.04040268$$

y en forma similar se obtienen los siguientes resultados:

i	0	1	2	3	4	5
x_i	0	0.1	0.2	0.3	0.4	0.5
y_i	1	1.010025	1.04040268	1.092055	1.16657	1.266297
Valor exacto	1	1.010025	1.04040268	1.09206	1.16657	1.266297

5.4 Métodos Predictor-Corrector

Estos métodos usan dos fórmulas:

- Una para predecir el siguiente valor de y_i , esto es y_{i+1}
- y la otra para hacer una corrección que proporcione un valor más exacto de y_{i+1} .

Existen varios de estos métodos, veremos el método Predictor-Corrector de Euler:

El Método Predictor-Corrector de Euler

Emplea el par Predictor-Corrector:

$$y_{i+1}^{(0)} = y_i + hy_i' \quad \dots \quad \text{Fórmula Predictora.}$$

$$y_{i+1}^{(k)} = y_i + \frac{1}{2}h(y_i' + y_{i+1}^{\prime}) \quad \dots \quad \text{Fórmula Correctora.}$$

para $k = 0, 1, 2, 3, \dots$

en donde $y_i' = f(x_i, y_i)$, $y_{i+1}^{\prime} = f(x_{i+1}, y_{i+1}^{(k-1)})$.

Con la 1^{ra}. Fórmula se estima y_{i+1} y esta estimación lleva entonces al valor y_{i+1}^{\prime} , por tanto al valor corregido y_{i+1} y este proceso puede repetirse nuevamente.

Ejemplo:

Resolver por el método de Euler para el primer valor $y(x_1)$, la ecuación diferencial:

$$y' = f(x_i, y_i) = -xy^2, \quad \text{si } y(0) = 2, \quad \text{y si } h = 0.2$$

$$x_0 = 0, \quad x_1 = 0.2$$

Para $i=0$, $x_0 = 0$, $y_0 = 2$:

Fórmula Predictora (FP):

$$y_1^{(0)} = y_0 + hy_0' = y_0 + hf(x_0, y_0) = 2 + 0.2f(0, 2) = 2$$

Fórmula Correctora (FC):

Para $k=1$

$$\begin{aligned} y_1^{(1)} &= y_0 + (h/2)[y_0' + y_1'] = \\ &= y_0 + (h/2)[f(x_0, y_0) + f(x_1, y_1^{(0)})] = 1.92, \end{aligned}$$

y en forma similar, se repite esta segunda fórmula, para:

$$\begin{aligned} k=2, & \quad y_1^{(2)} = 1.926272, \\ k=3, & \quad y_1^{(3)} = 1.925789524, \\ k=4, & \quad y_1^{(4)} = 1.925826694, \\ k=5, & \quad y_1^{(5)} = 1.925823831, \\ k=6, & \quad y_1^{(6)} = 1.925824051, \\ k=7, & \quad y_1^{(7)} = 1.925824035, \\ k=8, & \quad y_1^{(8)} = 1.925824036. \end{aligned}$$

El valor exacto es $y(0.2) = 1.923076923$, lo que da un error absoluto de 2.7471×10^{-3} .

5.5 Sistemas de Ecuaciones Diferenciales

Un sistema de n ecuaciones diferenciales de primer orden, es de la forma:

$$y_j' = f_j(x, y_1, y_2, y_3, \dots, y_n), \quad j = 1, 2, \dots, n, \quad x \in [a, b],$$

con las condiciones iniciales:

$$y_j(x_0) = a_j, \quad \forall j = 1, 2, 3, \dots, n; \quad \text{con } x_0 = a.$$

Para determinar los valores de las funciones $y_j(x)$ en los puntos seleccionados se emplean cualesquiera de los métodos dados para resolver ecuaciones de 1er. orden.

Aplicaciones:**Ejemplo 1:**

Resolver el sistema:

$$y' = f_1(x, y, z) = -2y/\sqrt{y^2 + z^2}$$

$$z' = f_2(x, y, z) = 1 - 2z/\sqrt{y^2 + z^2}$$

para $y(0.2)$, $z(0.2)$, con $h = 0.1$, si $y(0) = 1$, $z(0) = 0$, empleando el método de Euler.

Solución:

Luego los puntos seleccionados son:

$$x_0 = 0, x_1 = 0.1, x_2 = 0.2$$

Aplicando el Método de Euler:

$$y(x_{i+1}) \approx y_{i+1} = y(x_i) + hy'(x_i) = y(x_i) + hf_1(x_i, y(x_i), z(x_i));$$

$$z(x_{i+1}) \approx z_{i+1} = z(x_i) + hz'(x_i) = z(x_i) + hf_2(x_i, y(x_i), z(x_i));$$

Para $i = 0, 1, 2$.

Para $i = 0$, $x_0 = 0$, $y(0) = 1$, $z(0) = 0$,

$$y(0.1) \approx y_1 = y(0) + 0.1f_1(0, y(0), z(0)) = 1 + 0.1f_1(0, 1, 0) = 0.8$$

$$z(0.1) \approx z_1 = z(0) + 0.1f_2(0, y(0), z(0)) = 0.1;$$

Para $i = 1$, $x_1 = 0.1$, $y(0.1) \approx 0.8$, $z(0.1) \approx 0.1$;

$$y(0.2) \approx y_2 = y(0.1) + 0.1f_1(0.1, 0.8, 0.1) = 0.601544;$$

$$z(0.2) \approx z_2 = z(0.1) + 0.1f_2(0.1, 0.8, 0.1) = 0.175193;$$

Ejemplo 2:

Resolver el sistema:

$$y' = f_1(x, y, z) = z;$$

$$z' = f_2(x, y, z) = x + y;$$

para $y(0.5)$, $z(0.5)$, con $h=0.1$, si $y(0)=1$, $z(0)=-1$, empleando el método de Euler.

Solución:

Los puntos seleccionados son:

$$x_0 = 0, x_1 = 0.1, x_2 = 0.2, \dots, x_5 = 0.5$$

El Método de Euler:

$$\begin{aligned} y(x_{i+1}) &\approx y_{i+1} = y(x_i) + hy(x_i) = y(x_i) + hf_1(x_i, y(x_i), z(x_i)); \\ z(x_{i+1}) &\approx z_{i+1} = z(x_i) + hz(x_i) = z(x_i) + hf_2(x_i, y(x_i), z(x_i)); \end{aligned}$$

para $i = 0, 1, 2, 3, 4, 5$.

Para $i=0$, $x_0 = 0$, $y(0) = 1$, $z(0) = -1$;

$$\begin{aligned} y(0.1) &\approx y_1 = y(0) + 0.1f_1(0, y(0), z(0)) = 1 + 0.1f_1(0, 1, -1) = 0.9; \\ z(0.1) &\approx z_1 = z(0) + 0.1f_2(0, y(0), z(0)) = -0.9; \end{aligned}$$

Para $i=1$, $x_1 = 0.1$, $y(0.1) \approx 0.9$, $z(0.1) \approx -0.9$;

$$\begin{aligned} y(0.2) &\approx y_2 = y(0.1) + 0.1f_1(0.1, 0.9, -0.9) = 0.81 \\ z(0.2) &\approx z_2 = z(0.1) + 0.1f_2(0.1, 0.9, -0.9) = -0.8; \end{aligned}$$

y en forma similar se obtienen los siguientes resultados:

i	0	1	2	3	4	5
x_i	0	0.1	0.2	0.3	0.4	0.5
$y(x_i)$	1	0.9	0.81	0.73	0.6601	0.6005
$z(x_i)$	-1	-0.9	-0.8	-0.699	-0.596	-0.48999

5.6 Ecuaciones Diferenciales de Orden Mayor

Una ecuación diferencial de orden m es de la forma:

$$y^{(m)} = f(x, y, y', y'', \dots, y^{(m-1)}), \quad x \in [a, b]$$

con condiciones iniciales:

$$y^{(j)}(x_0) = \alpha_{j+1}, \quad j = 0, 1, 2, \dots, m-1; \quad \text{con } x_0 = a.$$

En este caso se redefinen las funciones de la siguiente forma:

$$\begin{aligned} y_1 &= y \\ y_2 &= y' = y_1' \\ y_3 &= y'' = y_2' \\ &\dots\dots\dots \\ y_m &= y^{(m-1)} = y_{m-1}' \Rightarrow y_m' = y^{(m)} \end{aligned}$$

$$\therefore y_j(x_0) = y^{(j-1)}(x_0) = \alpha_j, \quad j = 1, 2, 3, \dots, m.$$

En estas condiciones la ecuación de orden m , se convierte en:

$$y_m' = f(x, y_1, y_2, \dots, y_m)$$

una ecuación diferencial de 1er. orden, y se genera un sistema de m ecuaciones diferenciales de 1er. orden que se resuelve por cualesquiera de los métodos dados.

Aplicaciones:

Ejemplo 1:

Hallar el valor aproximado de y (0.4) en la ecuación diferencial:

$$y'' = f(x, y, y') = 0.1(1-y^2) y' - y$$

Con condiciones iniciales:

$$y(0) = 1, y'(0) = 0,$$

y paso $h= 0.2$.

Mediante el método de Euler.

Solución:

Redefiniendo:

$$\begin{aligned}y_1 &= y; \\y_2 &= y' = y_1' = f_1(x, y_1, y_2); \\y_3 &= y'' = y_2' = f_2(x, y_1, y_2);\end{aligned}$$

por lo tanto se tiene el sistema:

$$\begin{aligned}y_1' &= f_1(x, y_1, y_2) = y_2; \\y_2' &= f_2(x, y_1, y_2) = y_3 = 0.1(1 - y_1^2)y_2 - y_1;\end{aligned}$$

con condiciones iniciales:

$$y_1(0) = y(0) = 1, \quad y_2(0) = y'(0) = 0;$$

Los puntos seleccionados son:

$$x_0 = 0, \quad x_1 = 0.2, \quad x_2 = 0.4;$$

Empleando el Método de Euler, tenemos:

$$\begin{aligned}y_1(x_{i+1}) &\approx y_1(x_i) + hy_1'(x_i) = y_1(x_i) + hf_1(x_i, y_1(x_i), y_2(x_i)); \\y_2(x_{i+1}) &\approx y_2(x_i) + hy_2'(x_i) = y_2(x_i) + hf_2(x_i, y_1(x_i), y_2(x_i));\end{aligned}$$

Para $i = 0, 1, 2$.

$$\text{Para } i = 0, \quad x_0 = 0, \quad y_1(0) = 1, \quad y_2(0) = 0;$$

$$\begin{aligned}y_1(0.2) &\approx y_1(0) + 0.2f_1(0, y_1(0), y_2(0)) = 1 + 0.2f_1(0, 1, 0) = 1; \\y_2(0.2) &\approx y_2(0) + 0.2f_2(0, y_1(0), y_2(0)) = -0.2;\end{aligned}$$

$$\text{Para } i = 1, \quad x_1 = 0.2, \quad y_1(0.2) \approx 1, \quad y_2(0.2) \approx -0.2;$$

$$\begin{aligned}y_1(0.4) &\approx y_1(0.2) + 0.2f_1(0.2, 1, -0.2) = 0.96; \\y_2(0.4) &\approx y_2(0.2) + 0.2f_2(0.2, 1, -0.2) = -0.4.\end{aligned}$$

Ejemplo 2:

Hallar el valor aproximado de y (2.3) en la ecuación diferencial:

$$y'' = f(x, y, y') = -y' / y^2,$$

con condiciones iniciales:

$$y(2) = 2, \quad y'(2) = 0.5, \quad \text{con } h = 0.1.$$

Mediante el método de Runge-Kutta de orden 2

Solución:

Redefiniendo:

$$\begin{aligned} y_1 &= y \\ y_2 &= y' = y'_1 = f_1(x, y_1, y_2) \\ y_3 &= y'' = y''_2 = f_2(x, y_1, y_2) \end{aligned}$$

por lo tanto se tiene el sistema:

$$\begin{aligned} y'_1 &= f_1(x, y_1, y_2) = y_2 \\ y'_2 &= f_2(x, y_1, y_2) = y'' = -y_2 / y_1^2 \end{aligned}$$

Con condiciones iniciales:

$$y_1(2) = y(2) = 2, \quad y_2(2) = y'(2) = 0.5$$

Método de Runge-Kutta de orden 2

$$x_0 = 2, \quad x_1 = 2.1, \quad x_2 = 2.2, \quad x_3 = 2.3$$

$$\begin{aligned} y_1(x_{i+1}) &\approx y_1(x_i) + 1/2[k_1 + k_2]; \quad i = 0, 1, 2, 3. \\ y_2(x_{i+1}) &\approx y_2(x_i) + 1/2[l_1 + l_2]; \quad i = 0, 1, 2, 3. \end{aligned}$$

donde:

$$\begin{aligned}
 k_1 &= h f_1(x_i, y_1(x_i), y_2(x_i)) \\
 l_1 &= h f_2(x_i, y_1(x_i), y_2(x_i)) \\
 k_2 &= h f_1(x_{i+1}, y_1(x_i) + k_1, y_2(x_i) + l_1) \\
 l_2 &= h f_2(x_{i+1}, y_1(x_i) + k_1, y_2(x_i) + l_1)
 \end{aligned}$$

Para $i = 0$, $x_0 = 2$, $y_1(2) = 2$, $y_2(2) = 0.5$:

$$\begin{aligned}
 k_1 &= h f_1(x_i, y_1(x_i), y_2(x_i)) = 0.1 f_1(2, 2, 0.5) = 0.1(0.5) = 0.05 \\
 l_1 &= h f_2(x_i, y_1(x_i), y_2(x_i)) = 0.1 f_2(2, 2, 0.5) = -1 / 80 \\
 k_2 &= h f_1(x_{i+1}, y_1(x_i) + k_1, y_2(x_i) + l_1) = 0.1 f_1(2.1, 2.05, 0.4875) = 0.04875 \\
 l_2 &= h f_2(x_{i+1}, y_1(x_i) + k_1, y_2(x_i) + l_1) = 0.1 f_2(2.1, 2.05, 0.4875) = -0.01160023795
 \end{aligned}$$

$$\begin{aligned}
 y_1(2.1) &\approx 2.049375; \\
 y_2(2.1) &\approx 0.487949881
 \end{aligned}$$

y de esta manera se obtienen los siguientes resultados:

i	0	1	2	3
x_i	2	2.1	2.2	2.3
$y_1(x_i)$	2	2.049375	2.09759	2.14472
$y_2(x_i)$	0.5	0.487950	0.476731	0.466252
$y_1(x_i)$ (v. Exacto)	2	2.04939	2.09762	2.14476
$y_2(x_i)$ (v. Exacto)	0.5	0.48795	0.47673	0.466252

Algoritmo de Runge-Kutta de Orden 4 para un Sistema de Dos Ecuaciones Diferenciales de Primer Orden:

Objetivo: Determinar los valores aproximados de las funciones

$$y(x), z(x), x \in [a, b]$$

si se tiene:

$$\begin{aligned}
 y'(x) &= f(x, y, z) \\
 z'(x) &= g(x, y, z)
 \end{aligned}$$

Con condiciones iniciales: $y(x_0) = y_0$, $z(x_0) = z_0$; en los puntos seleccionados:

$$x_i = x_0 + ih, \quad i = 0, 1, \dots, N$$

con:

$$\begin{aligned} x_0 &= a \\ x_N &= b \\ h &= \frac{b-a}{N} \end{aligned}$$

Hacer para $i = 0, 1, 2, \dots, N$

Tomar:

$$z(x_{i+1}) \approx z_{i+1} = z_i + \frac{1}{6} [l_1 + 2l_2 + 2l_3 + l_4]$$

$$y(x_{i+1}) \approx y_{i+1} = y_i + \frac{1}{6} [k_1 + 2k_2 + 2k_3 + k_4]$$

con:

$$k_1 = hf(x_i, y_i, z_i)$$

$$l_1 = hg(x_i, y_i, z_i)$$

$$k_2 = hf\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}, z_i + \frac{l_1}{2}\right)$$

$$l_2 = hg\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}, z_i + \frac{l_1}{2}\right)$$

$$k_3 = hf\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}, z_i + \frac{l_2}{2}\right)$$

$$l_3 = hg\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}, z_i + \frac{l_2}{2}\right)$$

$$k_4 = hf(x_i + h, y_i + k_3, z_i + l_3)$$

$$l_4 = hg(x_i + h, y_i + k_3, z_i + l_3)$$

5.7 Ejercicios

1. Un proyectil de masa $m = 0.11\text{kg}$. se lanza verticalmente hacia arriba con una velocidad inicial $v(0) = 8\text{m/seg}$. y se va frenando

debido a la fuerza de la gravedad y a la resistencia del aire, donde $g = 9.8\text{m/seg}^2$ y $k = 0.002\text{ kg/m}$.

La ecuación diferencial para la velocidad v está dada por:

$$mv' = -mg - kv^2.$$

Calcular $v(1\text{ seg})$.

- Empleando el método de Euler, con $h = 0.1\text{seg}$.
 - Empleando el método de Taylor de orden 2, con $h = 0.2$.
 - Empleando el método de Runge-Kutta de orden 2, con $h = 0.2$.
 - Empleando el método Predictor-Corrector de Euler para $v(0.1)$, con $h = 0.1$.
2. Una pieza metálica con una masa de 0.1 kg y 25°C se calienta internamente de forma eléctrica a razón de $q = 3000\text{ w}$.

La ecuación diferencial de la temperatura que se obtiene es:

$$dT/dt = 20 - t^2, \text{ si } T(0) = 298.$$

Calcular $T(1)$.

- Empleando el método de Euler, con $h = 0.1$.
 - Empleando el método de Taylor de orden 2, con $h = 0.2$.
 - Empleando el método de Runge-Kutta de orden 2, con $h = 0.2$.
 - Empleando el método Predictor - Corrector de Euler para $T(0.1)$, con, $h = 0.1$.
3. Para la ecuación:

$$y' = 3y + e^{(1-x)}$$

si $y(0) = 1$, $h = 0.5$, calcular aproximadamente $y(2)$.

- Utilice el método de Taylor de orden 2.
- Use el método de Runge-Kutta de orden 2.
- Emplee el método de Predictor-Corrector de Euler, utilizando la fórmula correctora varias veces, para hallar $y(0.5)$, con una tolerancia de 10^{-5} .

4. Emplee el método Predictor-Corrector de Euler para hallar $y(1.05)$ e $y(1.1)$, si

$$y' = \left[\frac{y}{x} \right]^2 + \left[\frac{y}{x} \right] (1) = 1, \quad h=0.05 \quad \text{con } TOL = 10^{-4}.$$

5. En un circuito eléctrico la diferencia de voltajes entre los puntos A y B, V_{AB} es:

$$V_{AB} = L \frac{di}{dt} + Ri + \frac{q}{C}$$

donde:

i = corriente

L = inductancia

R = resistencia

q = carga

C = capacitancia

Derivando respecto de t y recordando que $\frac{dq}{dt} = i$ obtenemos:

$$L \frac{d^2i}{dt^2} + R \frac{di}{dt} + \frac{1}{C} i = \frac{dv}{dt}$$

Usando el método de Runge-Kutta de orden 4, determine como varía la corriente entre 0 y 0.1 seg., si $h = 0.02$ y

$$V_{AB} = 15 \text{ volts.}$$

$$C = 1,000 \text{ Microfarads}$$

$$L = 50 \text{ milihenries}$$

$$R = 4.7 \text{ ohms.}$$

6. Calcular $y(0.5)$, para la ecuación siguiente por medio del método de Runge-Kutta de orden 4, con $h = 0.1$,

$$y'' - 0.01(y')^2 + 2y = \text{sen } x$$

si $y(0) = 0, \quad y'(0) = 1.$

7. Calcular $y(0.5)$, para la ecuación siguiente por medio del método de Runge-Kutta de orden 4, con $h = 0.1$,

$$(1+x^2)y'' - 2(x+1)y' + 6y = 0$$
$$\text{si } y(0) = -1/2, \quad y'(0) = 0.$$

8. Calcular $y(0.5)$, para la ecuación siguiente por medio del método de Runge-Kutta de orden 4, con $h = 0.1$,

$$y'' - 0.2y' + 0.003y \operatorname{sen}(x) = 0$$
$$\text{si } y(0) = 0, \quad y'(0) = 1.$$

9. Resolver la siguiente ecuación diferencial:

$$y'' = 1 + y^2, \quad \text{si } y(0) = y'(0) = 0.$$

para $y(0.01)$, con $h = 0.01$ por el método Predictor-Corrector de Euler.

10. Resolver la ecuación diferencial:

$$y'' + 2ty' + t^2y = e^t, \quad \text{si } y(0) = 1, \quad y'(0) = -1, \quad h = 0.1$$

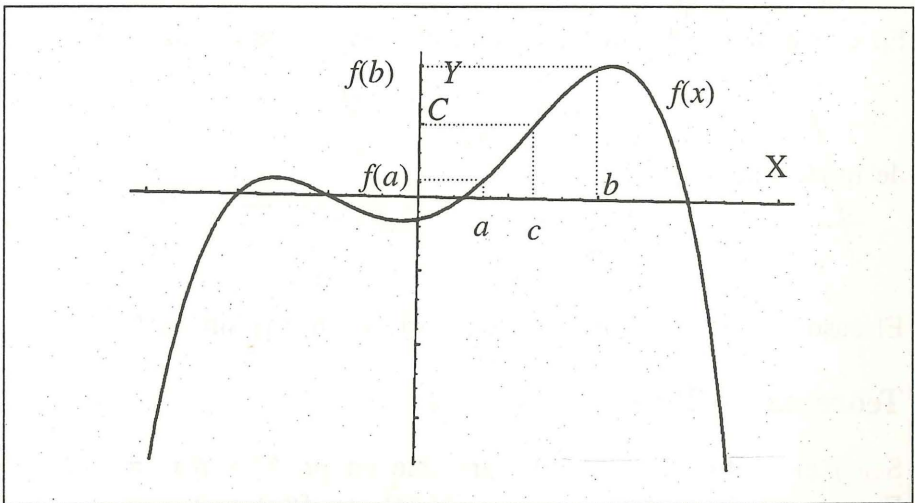
para obtener aproximadamente $y(0.4)$, empleando el método de Euler.

Apéndice A

Repaso de los Resultados Básicos del Cálculo

Teorema del Valor Intermedio

Sea $y = f(x)$ una función continua en $[a, b]$ y C un número real tal que $f(a) \leq C \leq f(b)$ ó $f(b) \leq C \leq f(a)$, entonces existe un $c \in [a, b]$ tal que $f(c) = C$.



Gráfica 1

Demostración

Sea $f(a) < f(b)$ y $C \in [f(a), f(b)]$

Considere el conjunto S de puntos de $[a, b]$ tal que $f(x) \leq C$ para $x \in S$. Claramente $a \in S$ y $b \notin S$ y b es una cota superior de S , es decir el conjunto S es acotado, por lo tanto existe un número que es la mínima cota superior de S . Representaremos esta mínima cota superior por c , y probaremos que $f(c) = C$.

Como c es la mínima cota superior de S , todo intervalo $]a, c[$, a la izquierda de c , contiene elementos x , tales que $f(x) \leq C$ (ver figura).

Entonces

$$\lim_{x \rightarrow c-} f(x) \leq C$$

Por otra parte $f(x) \geq C$ para todo $x \in]c, b[$ por la selección de c ; entonces

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

En consecuencia las dos desigualdades anteriores producen

$$C \leq \lim_{x \rightarrow c} f(x) \leq C$$

de modo que:

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

El caso donde $f(a) > f(b)$ se demuestra en forma similar. \square

Teorema de Rolle

Sea $f(x)$ continua en $[a, b]$, derivable en $]a, b[$ y $f(a) = f(b) = 0$. Entonces hay un número $c \in]a, b[$ tal que $f'(c) = 0$.

Demostración

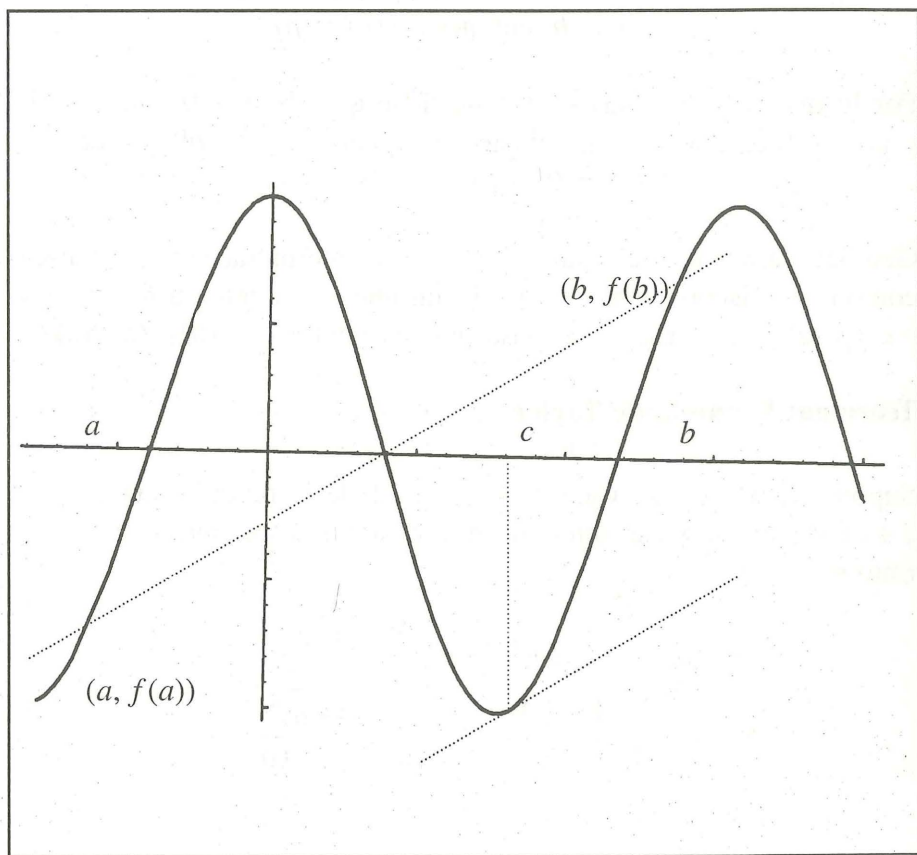
Si $f(x) = 0$ para todo $x \in [a, b]$, entonces $f'(c) = 0$ para todo $c \in]a, b[$.
 Sea $f(x) > 0$ para algún $x \in [a, b]$, entonces $f(x)$ tiene un valor máximo positivo. Sea $f(c)$ el máximo de f en $c \in]a, b[$. Entonces $f'(c) = 0$, así que

$c \neq a$, $c \neq b$. Por lo tanto $a < c < b$ y $f'(c) = 0$. Si $f(x) < 0$ para algún $x \in [a, b]$, entonces $f(x)$ tiene un valor mínimo negativo. Sea $f(c)$ el mínimo de $f(x)$ en $[a, b]$, entonces $c \neq a$, $c \neq b$. Y por lo tanto $a < c < b$ y $f'(c) = 0$. \square

Teorema del Valor Medio

Sea $f(x)$ continua en el intervalo cerrado en $[a, b]$ y derivable en el intervalo $]a, b[$. Entonces, por lo menos para un número $c \in [a, b]$ se tiene:

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



Gráfica 2

Demostración:

Consideremos la recta que pasa por los puntos $(a, f(a))$ y $(b, f(b))$ (Ver la gráfica). Esta recta tiene la pendiente

$$m = \frac{f(b) - f(a)}{b - a}$$

y es la gráfica de la función $g(x) = m(x - a) + f(a)$, continua y derivable en $[a, b]$. La derivada de $g(x)$ es una función constante: $g'(x) = m$. Ahora,

$$\text{si } x = a, \text{ entonces: } g(a) = f(a)$$

$$\text{si } x = b, \text{ entonces: } g(b) = f(b)$$

Por lo que si $h(x) = f(x) - g(x)$, se tiene que: $h(a) = 0$; $h(b) = 0$ y por el Teorema de Rolle existe un número $c \in]a, b[$ tal que,
 $h'(c) = 0$, ó $f'(c) = g'(c) = m$.

Geoméricamente este resultado se puede interpretar (ver la gráfica) como que existe un punto $c \in]a, b[$ tal que la tangente a $f(x)$ en $x = c$ es paralela a la recta que pasa por los puntos $(a, f(a))$, $(b, f(b))$.

Teorema: Fórmula de Taylor

Supóngase que la derivada de orden $n+1$ de la función $f(x)$ existe y es continua sobre el intervalo que contiene a los puntos a y b . Entonces

$$f(b) = f(a) + f'(a)(b-a) + f''(a) \frac{(b-a)^2}{2!} + \dots \\ + f^{(n)}(a) \frac{(b-a)^n}{n!} + f^{(n+1)}(z) \frac{(b-a)^{n+1}}{(n+1)!}$$

para algún número z comprendido entre a y b .

Demostración:

Usaremos un artificio que consiste en empezar por introducir dos funciones auxiliares $F(x)$ y $G(x)$ que se definen en la siguiente forma:

$$F(x) = f(b) - f(x) - f'(x)(b-x) - \dots - \frac{f^{(n)}(x)}{n!}(b-x)^n \dots \quad (1)$$

y

$$G(x) = \frac{(b-x)^{n+1}}{(n+1)!} \dots \quad (2)$$

Notar que $F(b)=G(b)=0$. Al calcular $F'(x)$ se obtiene

$$\begin{aligned} F'(x) = & -f'(x) + [f'(x) - f^{(2)}(x)(b-x)] + \left[f^2(x)(b-x) - \frac{1}{2!} f^{(3)}(x)(b-x)^2 \right] + \\ & \left[\frac{1}{2!} f^{(3)}(x)(b-x)^2 - \frac{1}{3!} f^{(4)}(x)(b-x)^3 \right] + \\ & + \dots \\ & \left[\frac{1}{(n-1)!} f^{(n)}(x)(b-x)^{n-1} - \frac{1}{n!} f^{(n+1)}(x)(b-x)^n \right]. \end{aligned}$$

Al examinar con cuidado este resultado, se ve que todos los términos, con excepción del último, se cancelan por pares, así, es que

$$F'(x) = -\frac{f^{(n+1)}(x)}{n!}(b-x)^n \quad \Lambda \quad (3)$$

Y la derivada de $G(x)$ es:

$$G'(x) = -\frac{(b-x)^n}{n!} \quad \Lambda \quad (4)$$

Ahora aplicaremos la fórmula de Cauchy del Valor Medio que afirma que

$$[F(b) - F(a)] G'(z) = [G(b) - G(a)] F'(z) \dots \quad (5)$$

para alguna z de $] a, b [$. Puesto que $F(b)=G(b)=0$, mientras que $G'(z) \neq 0$ por la expresión (4), se sigue de la ecuación (5) que

$$F(a) = \frac{F'(z)}{G'(z)} G(a) = \frac{-f^{(n+1)}(z)(b-z)^n / n! (b-a)^{n+1}}{-(b-z)^n / n! (n+1)!};$$

esto es,

$$F(a) = \frac{f^{(n+1)}(z)}{(n+1)!} (b-a)^{n+1} \quad \dots (6)$$

Pero la sustitución $x = a$ en la ecuación (1) produce

$$F(a) = f(b) - f(a) - f'(a)(b-a) - \dots - \frac{f^{(n)}(a)}{n!} (b-a)^n \quad \dots (7)$$

Para concluir, se obtiene la Fórmula de Taylor (7) cuando se igualan los segundos miembros de las ecuaciones (6) y (7). Nótese también que no importa si $a < b$ ó $a > b$ porque la fórmula de Cauchy (5) no se afecta por un intercambio de a y b . \square

Teorema de Taylor para la función de dos variables

Sea

$$f : \mathfrak{R}^2 \rightarrow \mathfrak{R},$$

con segundas derivadas continuas en un disco D con el centro en (x_0, y_0) . Entonces para cualquier punto $(x, y) \in D$ se tiene:

$$f(x, y) = f(x_0, y_0) + (x-x_0)f'_x(x_0, y_0) + (y-y_0)f'_y(x_0, y_0) + \frac{1}{2} \left[(x-x_0)^2 f''_{xx}(c, d) + 2(x-x_0)(y-y_0)f''_{xy}(c, d) + (y-y_0)^2 f''_{yy}(c, d) \right]$$

donde (c, d) es un punto en D .

(Demostración: Leer en el Libro de Tom M. Apostol, II Volumen[1], Editorial Reverté, S.A.)

Apéndice B

Programas en Pascal

Los siguientes programas en Pascal se encuentran grabados en el disquete que viene incluido en el libro:

- B.1 Programa de Bisección(Bisección.pas)
- B.2 Programa de Punto Fijo(Pfijo.pas)
- B.3 Programa de Newton(Newton.pas)
- B.4 Programa de Punto Fijo para Dos Variables(Pfijo2var.pas)
- B.5 Programa de Newton para Dos Variables(Newto2var.pas)
- B.6 Programa de Bairstow(Bairstow.pas)
- B.7 Programa de Eliminación Gaussiana(Eligauss.pas)
- B.8 Programa de Factorización Triangular(Factrian.pas)
- B.9 Programa de Jacobi(Jacobi.pas)
- B.10 Programa de Gauss-Seidel(Gausseide.pas)
- B.11 Programa para el Método de Potencia(Potencia.pas)
- B.12 Programa de Diferencias Divididas(Difdiv.pas)
- B.13 Programa de Diferencias Finitas(Difin.pas)
- B.14 Programa de la Regla del Trapecio(Trapecio.pas)
- B.15 Programa de la Regla de Simpson(Simpson.pas)

Apéndice C

Respuestas a los Ejercicios Seleccionados

C.1 Capítulo 1

1. a. $P_3(0.5) = 0.41\widehat{6}$. La cota es aproximadamente 0.015625.
b. $P_4(\pi/3) = 0.5017962015$. La cota es aproximadamente 0.0090885.
c. $P_5(2) = 2.71\widehat{6}$. La cota es aproximadamente 0.003775.
3. a. La respuesta con una aritmética de punto flotante de 4 dígitos es:

$$x_1 = 140.5; \quad x_2 = 0.$$

Las raíces exactas son:

$$x_1 = 140.4822042; \quad x_2 = 0.01779584834$$

- b. La respuesta con una aritmética de punto flotante de 4 dígitos es:

$$x_1 = 217.8; \quad x_2 = 0.025.$$

Las raíces exactas son:

$$x_1 = 217.7859954; \quad x_2 = 0.01400457359.$$

5. a. $P_3(0.5) = 0.541\widehat{6}$. La cota es aproximadamente: 0.0427874
b. Tiene rapidez de convergencia del orden $\{1/n^2\}_{(n \geq 1)}$.
7. El valor es: $k = 4$.

C.2 Capítulo 2

1. a. $r_1 = 0$; $r_2 \in [1.5, 2]$ $r_3 \in [-2, -1.5]$
 b. $r_2 = 1.895$

3. a. $r_1 \in [0.5, 0.65]$ $r_2 \in [0.7, 0.9]$
 b. $r_1 = 0.6137$ $r_2 = 0.8353$
 c. $n = 54$ iteraciones $n = 56$ iteraciones

6. i.
 a. $r_1 \in [-0.5, -0.4]$ $r_2 \in [2.2, 2.3]$
 b. $r_1 = -0.452276$ $r_2 = 2.271335$
 iii.
 a. $r_1 \in [-1.5, -1.4]$
 b. $r_1 = -1.4016670$
 v.
 a. $r_1 \in [1.2, 1.3]$
 b. $r_1 = 1.2926957$ $r_2 = 0$

7. i.
 a. $r_1 \in [1, 1.5]$ $r_2 \in [2.5, 3]$
 b. $r_1 = 1.4190544$ $r_2 = 2.7063271$
 ii.
 a. $r_1 \in [0, 1]$ $r_2 \in [3, 4]$
 b. $r_1 = 0.5085208$ $r_2 = 3.6857152$

10. a. $r_1 = 1.829$, $r_1 \in [1.5, 2]$.
 b. Se obtienen 4 iteraciones, $r_1 = 1.82938$.

17. a. $r = 1.4973003891$
 b. $r = 1.29269571937$

20. a. $r = -4.89402$

26. $x = 0.771844506347$ $y = 0.419643377608$

28. a. Tiene una solución en $[1, 2] \times [0, 1]$
 b. $F(x, y) = 1/x$ $G(x, y) = \ln(e^x + 2)$
 $x = 1.39755189887$ $y = 0.715536933411$

30. $x = 0.49924130683$ $y = 1.64747087153$
31. a. Las raíces se ubican en las regiones:
 $(x_1, y_1) \in [1.8, 2.2] \times [-1, 0]$; $(x_2, y_2) \in [2.8, 3] \times [1, 2]$
 b. $x_1 = 1.90540878804$, $y_1 = -0.78332131893$
33. a. Las raíces se ubican en las regiones:
 $(x_1, y_1) \in [1.5, 2] \times [3.5, 4]$; $(x_2, y_2) \in [3.5, 4] \times [1, 2]$
 b. $x_1 = 1.87596433203$; $y_1 = 3.9961500828$
 $x_2 = 3.73859492514$; $y_2 = 1.01142138082$
35. a. $P(x) = x^2 + 1.3135618233x + 1.13292$
 b. $P(x) = x^2 + 3.900502326x - 8.36456121118$
37. a. $r_{1,2} = 1.9999999994 \pm 0.999999988i$
 $P(x) = x^2 - 4x + 5$
 b. $P(x) = x^2 - 6x + 10$
 $r_{1,2} = 2.9999999994 \pm 0.99999998i$

C.3 Capítulo 3

2. a. $x = \begin{bmatrix} 1.694915 \\ 1.694915 \\ 1.694915 \end{bmatrix}$; b. $x = \begin{bmatrix} 1.694915 \\ 1.694915 \\ 1.694915 \end{bmatrix}$

3. $x = \begin{bmatrix} 0.871970 \\ 10.052118 \\ -11.092794 \\ -5.595156 \\ 5.184936 \\ -4.040099 \\ 5.168957 \end{bmatrix}$

6. a. $L = \begin{bmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.5 & 0 & 1 \end{bmatrix}$ $U = \begin{bmatrix} 4 & 2 & 2 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

$$b. \quad y = \begin{bmatrix} 1 \\ -1.5 \\ 0.5 \end{bmatrix} \Rightarrow x = \begin{bmatrix} 0.3875 \\ -0.3750 \\ 0.1 \end{bmatrix}; \quad y = \begin{bmatrix} 1 \\ -0.5 \\ -1.5 \end{bmatrix} \Rightarrow x = \begin{bmatrix} 0.4625 \\ -0.125 \\ -0.3 \end{bmatrix}$$

$$7. \quad a. \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ -0.5 & -0.2 & 1 \end{bmatrix} \quad U = \begin{bmatrix} 2 & -3 & 1 \\ 0 & 2.5 & -1.5 \\ 0 & 0 & -2.8 \end{bmatrix}$$

$$b. \quad A^{-1} = \begin{bmatrix} 0.142857 & 0.571429 & -0.142857 \\ -0.285714 & 0.357143 & -0.214286 \\ -0.145857 & -0.071429 & -0.357143 \end{bmatrix}$$

$$c. \quad x = \begin{bmatrix} -0.285714 \\ -0.928571 \\ -0.214286 \end{bmatrix}; \quad x = \begin{bmatrix} 2.428571 \\ -1.357143 \\ -2.928571 \end{bmatrix}; \quad x = \begin{bmatrix} 1.0 \\ 1.0 \\ 1.0 \end{bmatrix}; \quad x = \begin{bmatrix} -0.142857 \\ 0.2825714 \\ 0.142857 \end{bmatrix}$$

10. i. 50 iteraciones aproximadamente.

$$ii. \quad x = \begin{bmatrix} 0.234937 \\ 0.107158 \\ 0.332713 \\ 1.339615 \\ 0.693818 \\ 0.724895 \end{bmatrix}$$

14. a. La matriz B de Jacobi:

$$B_J = \begin{bmatrix} 0 & 0 & 5/8 \\ -1/3 & 0 & -1/3 \\ 1/2 & 1/6 & 0 \end{bmatrix} \Rightarrow \|B_J\|_\infty = 2/3$$

La matriz B de Gauss-Seidel:

$$B_G = \begin{bmatrix} 0 & 0 & 0.625 \\ 0 & 0 & -0.5416 \\ 0 & 0 & 0.222 \end{bmatrix} \Rightarrow \|B_G\|_u = 0.625$$

Aplicando el Método de Potencia en cada caso, obtenemos que para la matriz de Jacobi diverge, y para la matriz de Gauss-Seidel se obtienen las siguientes iteraciones:

i	$\lambda^{(i)}$	$u_1^{(i)}$	$u_2^{(i)}$	$u_3^{(i)}$
1	0.625	0.7298894	-0.632493	0.2592567
2	0.222	0.7298894	-0.632493	0.2592567

b. El Método más conveniente es el de Gauss-Seidel con $x^{(0)}=(0, 0, 0)$

i	$x_1^{(i)}$	$x_2^{(i)}$	$x_3^{(i)}$
0	0	0	0
1	0.25	-0.08383	0.111111
2	0.319444	-0.143518	0.1358024
3	0.334876	-0.156893	0.141289
4	0.338306	-0.159865	0.142509
5	0.339068	-0.160526	0.1427797
6	0.339237	-0.160672	0.1428399
7	0.3392749	-0.160705	0.1428533
8	0.3392833	-0.1607122	0.1428563
9	0.3392852	-0.1607138	0.1428570
10	0.3392856	-0.160714	0.1428571

c. La cota superior es 0.006063.

16. $TOL = 0.001$

a. $\lambda^{(7)} = 5.236068$; $\mathbf{u}^{(7)} = \begin{bmatrix} 0.7071064 \\ 0.6319947 \\ 0.3171485 \end{bmatrix}$

b. $A^{-1} = \begin{bmatrix} 0.75 & -0.5 & -0.25 \\ -0.5 & 2/3 & 1/6 \\ -0.25 & 1/6 & 5/12 \end{bmatrix}$

El valor propio dominante de A^{-1} , con $TOL = 0.001$, es $\lambda = 1.309$, y su recíproco $\lambda^{-1} = 0.763942$, que corresponde al valor propio de menor magnitud de A .

19. a. La matriz B de Jacobi:

$$B_J = \begin{bmatrix} 0 & 0 & -1/2 \\ -1/5 & 0 & -3/5 \\ 1/7 & -2/7 & 0 \end{bmatrix} \Rightarrow \|B_J\|_u = 4/5$$

La matriz B de Gauss-Seidel:

$$B_G = \begin{bmatrix} 0 & 0 & -0.5 \\ 0 & 0 & -0.5 \\ 0 & 0 & 1/14 \end{bmatrix} \Rightarrow \|B_G\|_u = 0.5$$

Más rápido converge el método de Gauss-Seidel.

b.

Con Jacobi: $\mathbf{x} = \begin{bmatrix} 0.423076 \\ -0.076925 \\ 0.653845 \end{bmatrix}$. Se logra con 14 iteraciones

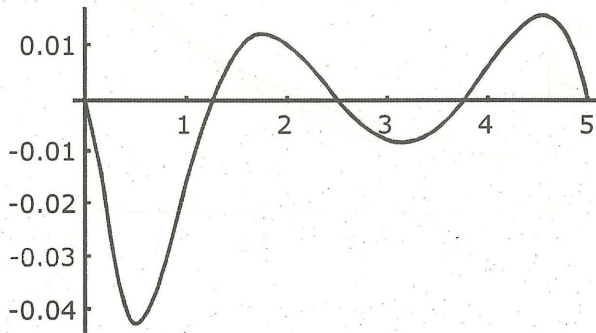
Con Gauss-Seidel:

$$x = \begin{bmatrix} 0.423196 \\ -0.0768039 \\ 0.653829 \end{bmatrix} =. \text{ Se logra con 7 iteraciones}$$

C.4 Capítulo 4

1. a. $P_4(x) = 0.00758052x^4 - 0.100041x^3 + 0.481618x^2 - 1.04067x + 1$

b. La gráfica de la distribución del error:



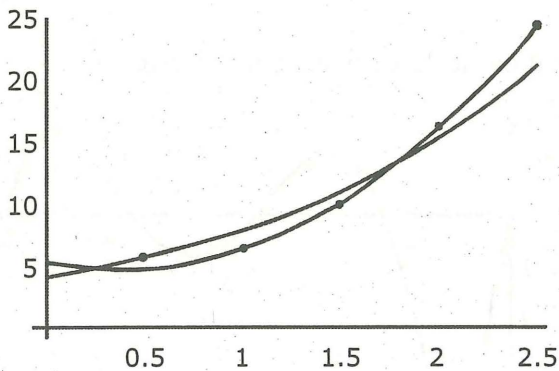
2. a. $P_3(x) = -0.0192604x^3 + 0.165417x^2 - 0.519128x + 0.706304$
 $P_3(2.41) = 0.146369$ donde $\epsilon \approx 0.00033$

b. $P_2(x) = 0.0633373x^2 - 0.349636x + 0.619632$, para $i = 0, 1, 2$
 $f(2.41) \approx P_2(2.41) = 0.144878$ donde $\epsilon \approx 0.00182$
 $P_2^*(x) = 0.0248165x^2 - 0.183997x + 0.446289$, para $i = 1, 2, 3$
 $f(2.41) \approx P_2^*(2.41) = 0.146993$ donde $\epsilon \approx 0.00029$

3. a. $P_4(x) = -0.0557125x^4 + 433.282583x^3 - 1.263568864 \times 10^6 x^2 + 1.637654903 \times 10^9 x - 7.958951155 \times 10^{11}$
 $P_4(1935) = 126578$

b. $P_4(x) = 0.076983x^4 - 604.398833x^3 + 1.779398717 \times 10^6 x^2 - 2.328259703 \times 10^9 x + 1.142382253 \times 10^{12}$
 $P_4(1975) = 213952$

14. a. $f'(1.4) \approx 2.1545$ $f''(1.4) \approx 1.9125$
 b. $I = 0.7674478048$ $\varepsilon_T < 7.919 \times 10^{-4}$
 c. $I = 0.7668093294$ $\varepsilon_S < 5.279 \times 10^{-7}$
 d. $I = 0.7668089022$, y el valor exacto de la integral es:
 0.766808904.
22. a. $F(x, c_0, c_1, c_2) = 5.347143 - 3.541143x + 4.457143x^2$
 b. $y = 4.024658e^{0.667848x}$



La parábola es mejor ajuste, está más cerca de la muestra de puntos.

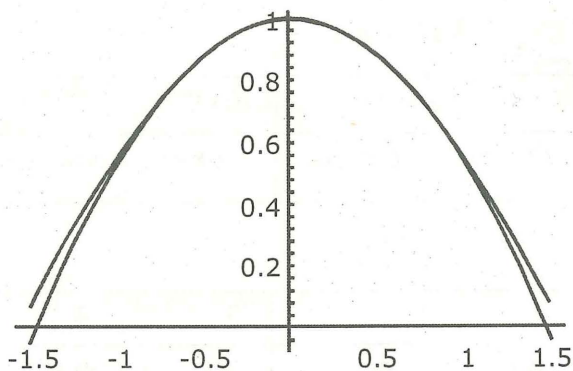
25. $f(x) = 2.02016e^x + 3.0144e^{-x}$

27. a. El polinomio de grado dos es:

$$P_2(x) = 0.996558 - 0.465262x^2$$

y el valor de la integral es: $I = 1.84 \times 10^{-5}$

- b. Las gráficas de $P(x)$ y $f(x)$, se muestran a continuación:



C.5 Capítulo 5

1. a.

i	1	2	3	4	5	6	7	8	9	10
t_i	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
v_i	6.9036	5.8370	4.7950	3.7732	2.7673	1.7734	0.7877	-0.1934	-1.1735	-2.1560

b.

i	1	2	3	4	5
t_i	0.2	0.4	0.6	0.8	1
v_i	5.8711	3.8302	1.8449	-0.1142	-2.0751

c.

i	1	2	3	4	5
t_i	0.2	0.4	0.6	0.8	1
v_i	5.8623	3.8139	1.8215	-0.1445	-2.1126

d.

$v_1^{(0)}$	$v_1^{(1)}$	$v_1^{(2)}$	$v_1^{(3)}$	$v_1^{(4)}$	$v_1^{(5)}$
6.903636	6.918491	6.918304	6.918306	6.918306397	6.918306397

7.

i	0	1	2	3	4	5
x_i	0	0.1	0.2	0.3	0.4	0.5
$y_1(x_i)$	-0.5	-0.483999	-0.432092	-0.338765	-0.199289	-0.010071
$y_2(x_i)$	0	0.329850	0.717462	1.156932	1.638532	2.149426

9.

$y_1^{(0)}_{(0.01)}$	$y_2^{(0)}_{(0.01)}$	$y_1^{(1)}_{(0.01)}$	$y_2^{(1)}_{(0.01)}$	$y_1^{(2)}_{(0.01)}$	$y_2^{(2)}_{(0.01)}$
0	0.01	5×10^{-5}	0.01	5×10^{-5}	0.01

Bibliografía

- [1] Apostol, T., *Calculus (Vol. I y II)*, Reverté, 1973.
- [2] Bakhvalov, N., *Métodos Numéricos*, Paraninfo, 1980.
- [3] Burden, R. y Douglas, J., *Análisis Numérico*, International Thomson Editores, 1998.
- [4] Carnahan, B., *Applied Numerical Methods*, John Wiley & Sons, 1964.
- [5] Conte-Carl de Boor, *Análisis Numérico*, McGraw-Hill, 1972.
- [6] Chapra, S. C. y Canale, R. P., *Métodos Numéricos para Ingenieros*, McGraw-Hill, 1987.
- [7] Demidovich-Marón, *Cálculo Numérico Fundamental*, Paraninfo, 1977.
- [8] García, F. y Nevot, A., *Análisis Numérico*, Paraninfo, 1992.
- [9] Gerald, C., *Análisis Numérico*, Alfaomega, 1991.
- [10] Henrici, P., *Elementos de Análisis Numérico*, John Wiley & Sons, 1964.
- [11] Isaacson, E., y Keller, H., *Analysis of Numerical Methods*, John Wiley & Sons, 1966.
- [12] Kincaid, D. y Cheney, W., *Análisis Numérico*, Addison-Wesley Iberoamericana, 1994.

- [13] Mason, J., *Cálculo Numérico*, Anaya Multimedia, 1986
- [14] Maron, M. y López, R., *Análisis Numérico*, CECSA, 1995.
- [15] Nakamura, S., *Métodos Numéricos Aplicados con Software*, Prentice-Hall, 1992.
- [16] Nakamura, S., *Análisis Numérico y Visualización Gráfica con Matlab*, Prentice-Hall, 1997.
- [17] Nieves, A. y Domínguez, C., *Métodos Numéricos Aplicados a la Ingeniería*, CECSA, 1996.
- [18] Ortega, J. M., *Numerical Analysis*, Academic Press, 1973.
- [19] Press, W., Flannery, B., Teukolsky, S., y Vetterling, W., *Numerical Recipes in Pascal*, Cambridge University Press, 1989.
- [20] Ralston, A., *Introducción al Análisis Numérico*, McGraw-Hill, 1965.
- [21] Scheid, F. y Di Constanzo, R. E., *Teoría y Problemas de Análisis Numérico*, McGraw-Hill, 1991.
- [22] Smith, W. A., *Análisis Numérico*, Prentice-Hall Hispanoamericana, 1988.
- [23] Stoer, J., y Bulirsch, R., *Introduction to Numerical Analysis*, Springer-Verlag, 1991.
- [24] Swokowski, E., *Cálculo con Geometría Analítica*, Grupo Editorial Iberoamérica, 1989.

Indice

A

- Absoluto, error, 5
- Aitken, método de, 68
- Ajuste de datos, 263
- Algoritmo, definición, 3
 - de sustitución
 - hacia atrás, 119
- Aproximación, teoría de,
(ver Ajuste de
datos), 263
- Aritmética de computadora, 7
- Aritmética de dígitos finitos, 7

B

- Bairstow, método de, 89
- Bien condicionada, matriz, 183
- Bisección, algoritmo de, 34
- Bisección, método de, 33

C

- Característica de un número, 9
- Característico, polinomio, 172
- Característico, valor, 171
- Característico, vector, 171
- Cero,
 - multiplicidad del, 66
- Choleski, método de, 136
- Cifras significativas, 7
- Condición, número de, 182
- Continua,
 - aproximación de mínimos
cuadrados, 278

- Contracción de mapeos
teorema de, 77

Convergencia

- aceleración de la, 68
- lineal, 51
- orden de la, 51
- cuadrática, 52
- rapidez de, 22, 51

Convergente,

- sucesiones, 19
- matrices, 175

Crecimiento de error, 16

- Crout, algoritmo, 138
- Crout, método de, 136
- Cuadrática, convergencia, 52
- Cuadrática, fórmula, 11
- Cuadratura(ver Integración
Numérica), 250
- Cuadratura Gaussiana
usando polinomios de
Legendre, 283
- Cuadratura Numérica(ver
Integración Numérica),
250

D

- Derivación o diferenciación
numérica, 241
- inestabilidad de la, 248
- fórmulas de tres puntos,
243
- Diferencia dividida,
definición, 218
- Diferencia finita o progresiva,
definición, 70

Diferenciación (ver Derivación Numérica), 241

Diferenciales, ecuaciones, 301
de orden mayor, 323
sistemas de, 320

Euler, método de, 301

Runge-Kutta de orden 2, método de, 311

Runge-Kutta, de orden 4 método de, 312

Taylor, método de, 306

Predictor-Corrector método de, 319

problemas de valor inicial, 301

Dígitos significativos, 7

Discreta de mínimos cuadrados aproximación, 278

ecuaciones normales caso general, 265

Distancia entre vectores, 156

Dominante diagonalmente matriz estrictamente, 179

Doolittle, método de, 136

E

Ecuaciones normales, 265

Error,
absoluto, 5
creciente exponencial del, 17
en la aritmética de computadora, 7
crecimiento lineal del, 17
relativo, 5
de redondeo, 3
de truncamiento, 3
propagado, 13

Error asintótico, constante de, 51

Escalado, pivoteo, 151

Escalar, producto, 271

Estabilidad de los métodos de derivación, 248

Estrictamente dominante diagonalmente, matriz, 179

Euclideana, norma, 155

Euler, algoritmo de, 303

Euler, método de, 301

F

Factorización de una matriz (ver Matriz, factorización), 134

Factorización directa, 136

Falsa posición, método de, 101

Función de iteración, 41

G

Gauss-Jordan, método de, 130

Gaussiana, eliminación con sustitución hacia atrás, 119

Gauss-Seidel, algoritmo iterativo de, 168

Gauss-Seidel, método iterativo de, 166

Gaussiana, eliminación, 115

Gaussiana, eliminación con pivoteo parcial, algoritmo de, 150

Gaussiana, eliminación con pivoteo escalado, algoritmo de, 153

H

Hermite, polinomios de, 277
 Hilbert, matriz de, 184

I

Inestabilidad
 de los métodos de derivación, 248

Integración numérica, 250
 Simpson, regla compuesta de un tercio, 257
 Simpson, regla de tres octavos de, 257
 Trapezoidal, regla compuesta, 253

Interpolación,
 Lagrange, polinomios de, 209
 Newton con diferencias divididas, polinomio de, 217
 Newton con diferencias finitas, polinomio de, 229

Inversa de una matriz, 126

J

Jacobi, método iterativo de, 160

L

Lagrange, polinomios de, definición, 209
 error de truncamiento, 211

Legendre, polinomios de
 definición, 275
 raíces y coeficientes, 286
 para la cuadratura gaussiana, 283

Lineal, sistema,
 sustitución hacia atrás o reversiva, 117

M

Mal condicionada, matriz, 183

Mantisa, 9

Matriz,
 número de condición, 182
 convergente, 175
 algoritmo de Crout, 138
 algoritmo de factorización directa, 146
 factorización, 136
 método de Gauss-Jordan, 130
 Hilbert, 184
 mal condicionada, 183
 inversa de, 126
 norma (ver también Norma de una matriz), 156
 estrictamente dominante diagonalmente, 179
 bien condicionada, 183
 Método de posición falsa, 101
 Mínimos cuadrados
 continuos, 278
 discretos, 278
 general, 265
 por polinomios ortogonales, 278

Müller, método de, 226

Multiplicidad de una raíz, 66

N

Newton, método de, 53

dificultades con, 61

algoritmo para sistemas

de dos ecuaciones, 84

definición para sistemas

no lineales, 82

modificación para raíces

múltiples, 66

convergencia cuadrática

del, 52

Newton-Raphson, método de, 53

No lineal, sistemas de

ecuaciones, 76, 82

Norma de una matriz, 156

Norma de un vector, 155

Normales, ecuaciones, 265

Número de máquina, 7

O

Orden superior, ecuaciones

diferenciales de, 323

Ortogonales

funciones, 272

polinomios, 272

P

Peso, función de, 271

Pivote, elemento, 115

Pivoteo

estrategias, 147

parcial, 147

escalado de columna, 151

Polinomio

característico, 172

Hermite, 277

interpolante, 205

Legendre, de, 275

Newton, de, 217

ortogonal, 272

Taylor, de, 207

Potencia, método de, 184

Predictor-Corrector,

método, 319

Producto escalar de

funciones, 271

Progresiva, diferencia, 70

Punto Fijo, 40

en \mathcal{R} , 40

en \mathcal{R}^2 , 76

Punto Fijo, iteración, 40

convergencia, 42

cota de error, 44

Punto Fijo, algoritmo de, 44

Punto Fijo, teorema de, 42

Punto flotante, forma de, 8

R

Radio espectral, 174

Rapidez de convergencia, 22, 51

Redondeo, error de, 3

Regula Falsi, 101

Relativo, error, 5

Residual, vector, 181

Rolle, teorema de, 332

Runge-Kutta

cuarto orden, 312

orden dos, 311

sistemas de ecuaciones

diferenciales, 326

S

- Secante, método de, 104
- Series, Taylor, 4
- Significativas, cifras, 7
- Simpson
 - regla compuesta de un tercio, 257
 - regla de tres octavos de, 257
- Steffensen, método de, 73
- Sustitución reversiva, 117

T

- Taylor, métodos para problemas de valor inicial, 306
- Taylor, polinomio de, 207
 - error de truncamiento, 208
- Taylor, series de, 4
- Taylor,
 - teorema de una variable, 334
 - teorema de dos variables, 337

Trapezoidal,

- regla compuesta, de integración, 253
- Truncamiento, error de, 3

V

Valor

- característico, 171
- dominante, 174

Valor inicial

- Euler, método de, 301
- Runge-Kutta de cuarto orden, 312
- Runge-Kutta de orden dos, 311
- Runge-Kutta de cuarto orden para sistemas, 326
- Taylor, métodos de, 306

Valor intermedio,

- teorema del, 331

Valor medio, teorema del, 333

Vector

- norma de, 155
- unitario, 156

Vector característico, 171

Vectores, convergencia de, 156

*Elementos de Análisis Numérico
algoritmos y aplicaciones*

se terminó de imprimir en el mes de diciembre de 1999
en los talleres de Servicio Copias Gráficas S.A.
(RUC: 10069912) Telefax: 424-9693
Jr. Jorge Chávez 1059, Lima 5, Perú

Este libro ha sido escrito como resultado de la experiencia adquirida durante la enseñanza de cursos sobre métodos numéricos para estudiantes de Ciencias e Ingeniería.

El libro presenta los conceptos teóricos fundamentales de los temas del Análisis Numérico y va acompañado de los algoritmos y programas en Pascal correspondientes. Este libro podrá ser usado por un estudiante a nivel de pregrado, con conocimientos básicos de análisis matemático.