

MÉTODOS NUMÉRICOS BÁSICOS

(Prerrequisitos: Límites y continuidad de funciones de una variable. Derivadas de funciones de una variable. Integrales definidas)

Introducción

Resumimos aquí métodos sencillos y muy básicos para enfrentar algunos problemas que no tienen soluciones basadas en las técnicas clásicas del Álgebra o del Análisis Matemático, o que sus soluciones clásicas se obtienen por procedimientos demasiado laboriosos. Se llaman métodos numéricos porque hacen uso de valores numéricos obtenidos por computadoras o calculadoras científicas a través de los correspondientes programas, a partir de determinados datos numéricos del problema a resolver. Lo que hacen fundamentalmente los métodos es operar con esos datos numéricos del problema, obteniendo normalmente algunos valores intermedios, para llegar finalmente a la solución exacta o a una solución aproximada del problema, dentro de un determinado margen de error prefijado. Por supuesto, cualquiera de los métodos numéricos está basado en teoremas de las Matemáticas, a veces muy específicos, y los métodos más sofisticados requieren el uso de computadoras de gran memoria y velocidad.

El conjunto de los métodos numéricos constituyen una parte sumamente importante de las Matemáticas que se llama Cálculo Numérico o Análisis Numérico, donde aparecen cuestiones de Álgebra (como el cálculo rápido de determinantes, la solución eficiente de grandes sistemas de ecuaciones lineales o el cálculo de autovalores de matrices) y cuestiones de Análisis Matemático (como la resolución aproximada de ecuaciones, la interpolación y la extrapolación de funciones, el ajuste de curvas, la diferenciación numérica, la integración numérica y la resolución numérica de ecuaciones diferenciales). Muchos de los mayores progresos tecnológicos actuales, entre los cuales cabe destacar los éxitos espaciales, se han logrado utilizando métodos del Análisis Numérico.

En este pequeño resumen sólo veremos dos métodos de interpolación polinómica, dos métodos de resolución aproximada de ecuaciones y tres métodos para el cálculo aproximado de integrales definidas.

La interpolación polinómica consiste en hallar el único polinomio que pasa por un cierto número de puntos dados en el plano con abscisas diferentes. Este problema puede resolverse por el método clásico de “coeficientes indeterminados”, que conduce a un sistema de ecuaciones lineales compatible con solución única, el cual puede resolverse por el “método de Gauss” o el “método de Gauss-Jordan” (los cuales son programables y pueden considerarse también como métodos numéricos). Pero hay otros métodos clásicos más ágiles, entre los cuales están el “método de Lagrange” y el “método de diferencias progresivas de Newton-Gregory”, los cuales pueden programarse y veremos en el próximo apartado.

La resolución de ecuaciones se refiere a obtener en forma aproximada una cierta solución de una ecuación (muchas veces no polinómica), para la que no disponemos de una “fórmula” que nos la determine exactamente. Sabemos que las ecuaciones polinómicas de primer grado son todas resolubles exactamente, despejando la incógnita a través de operaciones fáciles. Sabemos también que las ecuaciones polinómicas de segundo grado son resolubles, después de llevarlas a una forma ordenada, mediante una conocida “fórmula” que nos da exactamente sus dos soluciones, tanto si son reales como si son imaginarias. Para ecuaciones de tercer grado ya no existe una “fórmula general”, como la anterior, que nos determine sus tres soluciones en cualquier caso; pero hay varias fórmulas que resuelven diferentes casos particulares de dichas ecuaciones. Lo mismo sucede con las ecuaciones de cuarto grado. Y ya para ecuaciones de grado cinco y superiores se sabe que no hay modo de resolverlas usando radicales, salvo casos triviales. Pero existen además ecuaciones que no son polinómicas, como $x - \cos x = 0$, $e^x + x = 0$, $x + \ln x = 0$, etc... para las cuales tampoco hay una “fórmula” que las resuelva exactamente. Veremos para todas estas ecuaciones dos métodos de resolución aproximada: El “método de bisección” y el “método de Newton-Raphson”.

MÉTODOS NUMÉRICOS BÁSICOS

También sabemos que hay integrales indefinidas que no pueden resolverse usando “funciones elementales”, mediante alguno de los “métodos generales de integración” o una combinación de ellos. Y hay otras integrales indefinidas cuya solución es larga o difícil. Con lo cual las correspondientes integrales definidas no pueden calcularse exactamente aplicando la Regla de Barrow, o pueden resolverse por este procedimiento pero el trabajo de obtención de una cierta primitiva es demasiado largo o incómodo. Sin embargo, sabemos que esas integrales definidas existen como números reales (por ejemplo, si el integrando es una función continua en el intervalo de integración). ¿Cómo calcularlas entonces de un modo rápido? En la Sección 4.2 vimos el “método de los rectángulos” con tres variantes y el “método de los trapecios”, para obtener valores aproximados de una integral definida. Pues bien, aquí recordaremos dichos métodos y daremos uno adicional, llamado “método de Simpson”.

Métodos de interpolación polinómica

En Álgebra se demuestra el siguiente importante teorema.

TEOREMA: Dados $n + 1$ puntos del plano con abscisas diferentes $(x_0, y_0), (x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, existe un único polinomio $P_n(x)$, de grado menor o igual que n , con la propiedad de que $P_n(x_0) = y_0, P_n(x_1) = y_1, P_n(x_2) = y_2, \dots, P_n(x_n) = y_n$. Este polinomio se llama “polinomio interpolador” correspondiente a los puntos dados.

O dicho de otro modo, hay una única función polinómica $y = P_n(x)$ cuya gráfica pasa por todos los puntos dados.

Cuando los $n + 1$ puntos dados pertenezcan a la gráfica de una función desconocida y queramos saber el valor de esa función desconocida en un punto nuevo de abscisa \bar{x} intermedia entre las que tenemos, una forma de obtener aproximadamente dicho valor nuevo es hallar el “polinomio interpolador” $P_n(x)$ correspondiente a los puntos dados y luego calcular el valor $P_n(\bar{x})$, que deberá parecerse al valor de la función desconocida en dicho punto \bar{x} (se parecerá en mayor medida cuando las abscisas de los puntos dados sean muy próximas entre sí).

Como dijimos, los $n + 1$ coeficientes del “polinomio interpolador” (por ser de grado menor o igual que n) pueden hallarse resolviendo un sistema lineal de $n + 1$ ecuaciones con esas $n + 1$ incógnitas, el cual es compatible con solución única, usando para ello cualquier método de solución, como la Regla de Cramer, el método de Gauss o el de Gauss-Jordan (ver Sección 9.2).

Por ejemplo, si tuviésemos los 4 puntos $(1, -2), (3, 0), (-2, 5)$ y $(2, 4)$, el polinomio interpolador correspondiente sería de grado 3 como máximo y lo podemos escribir como $P_3(x) = ax^3 + bx^2 + cx + d$. Con lo cual, las 4 ecuaciones lineales con las 4 incógnitas a, b, c y d resultarán de pedir que el polinomio pase por los 4 puntos dados: Para que pase por $(1, -2)$ tendrá que ser $1^3a + 1^2b + 1c + d = -2$. Para que pase por $(3, 0)$ tendrá que ser $3^3a + 3^2b + 3c + d = 0$. Y las dos ecuaciones restantes serán

$$(-2)^3a + (-2)^2b + (-2)c + d = 5 \quad \text{y} \quad 2^3a + 2^2b + 2c + d = 4.$$

Pero hay otros métodos programables más rápidos para determinar $P_n(x)$, que veremos a continuación.

MÉTODO DE LAGRANGE (Joseph Louis Lagrange, matemático francés del siglo XVIII, nacido en Italia):

Consiste en expresar $P_n(x)$ como una combinación lineal de “polinomios de Lagrange de grado n ” (o sea, una suma de esos polinomios multiplicados previamente por ciertos números reales), con lo cual el polinomio resultante será con seguridad de grado menor o igual que n . Dichos “polinomios de Lagrange de grado n ” se construyen usando exclusivamente los valores de las $n + 1$

MÉTODOS NUMÉRICOS BÁSICOS

abscisas diferentes $x_0, x_1, x_2, \dots, x_n$ y tienen la propiedad de tomar el valor 1 en una sola de las abscisas dadas y tomar el valor cero en todas las demás.

El polinomio 0 de Lagrange, de grado n es:

$$L_0(x) = \frac{(x-x_1) \cdot (x-x_2) \cdot (x-x_3) \cdot \dots \cdot (x-x_n)}{(x_0-x_1) \cdot (x_0-x_2) \cdot (x_0-x_3) \cdot \dots \cdot (x_0-x_n)}$$

de modo que $L_0(x_0) = 1$ y todos sus demás valores en las abscisas dadas son cero. Obsérvese que x_0 es la única abscisa ausente en el numerador, el cual incluye n factores de primer grado, luego $L_0(x)$ es de grado n . Además x_0 aparece en todos los n factores del denominador, los cuales son diferentes de cero porque todas las abscisas son distintas.

El polinomio 1 de Lagrange, de grado n es:

$$L_1(x) = \frac{(x-x_0) \cdot (x-x_2) \cdot (x-x_3) \cdot \dots \cdot (x-x_n)}{(x_1-x_0) \cdot (x_1-x_2) \cdot (x_1-x_3) \cdot \dots \cdot (x_1-x_n)}$$

de modo que $L_1(x_1) = 1$ y todos sus demás valores en las abscisas dadas son cero. Obsérvese que x_1 es la única abscisa ausente en el numerador, el cual también incluye n factores de primer grado, luego $L_1(x)$ es de grado n . Además x_1 aparece en todos los n factores del denominador.

El polinomio 2 de Lagrange, de grado n es:

$$L_2(x) = \frac{(x-x_0) \cdot (x-x_1) \cdot (x-x_3) \cdot \dots \cdot (x-x_n)}{(x_2-x_0) \cdot (x_2-x_1) \cdot (x_2-x_3) \cdot \dots \cdot (x_2-x_n)}$$

de modo que $L_2(x_2) = 1$ y todos sus demás valores en las abscisas dadas son cero.

Etcétera... Y **el polinomio n de Lagrange**, de grado n es:

$$L_n(x) = \frac{(x-x_0) \cdot (x-x_1) \cdot (x-x_2) \cdot \dots \cdot (x-x_{n-1})}{(x_n-x_0) \cdot (x_n-x_1) \cdot (x_n-x_2) \cdot \dots \cdot (x_n-x_{n-1})}$$

de modo que $L_n(x_n) = 1$ y todos los demás valores en las abscisas dadas son cero. Obsérvese que x_n es la única abscisa ausente en el numerador, el cual vuelve a incluir n factores de primer grado, luego $L_n(x)$ es de grado n . Además x_n aparece en todos los n factores del denominador.

Pues bien, **la expresión del “polinomio interpolador”**, basada en los anteriores polinomios de Lagrange, es la siguiente combinación lineal de $n + 1$ términos:

$$P_n(x) = y_0 \cdot L_0(x) + y_1 \cdot L_1(x) + \dots + y_n \cdot L_n(x)$$

Pues de esta manera será $P_n(x_0) = y_0$, ya que $L_0(x_0) = 1, L_1(x_0) = 0, \dots, L_n(x_0) = 0$. Será $P_n(x_1) = y_1$, ya que $L_0(x_1) = 0, L_1(x_1) = 1, L_2(x_1) = 0, \dots, L_n(x_1) = 0$. Y así sucesivamente. Hasta llegar a $P_n(x_n) = y_n$, ya que $L_0(x_n) = 0, L_1(x_n) = 0, \dots, L_{n-1}(x_n) = 0, L_n(x_n) = 1$.

Nota: Los polinomios de Lagrange no deben ser escritos en forma ordenada según potencias de x , lo cual ocasionaría muchas operaciones innecesarias. **Deben introducirse en el ordenador con la forma presentada anteriormente**, sustituyendo los valores de las abscisas dadas en los factores de cada numerador y de cada denominador. Así, cuando nos den una abscisa nueva intermedia \bar{x} , obtendremos directamente

$$P_n(\bar{x}) = y_0 \cdot L_0(\bar{x}) + y_1 \cdot L_1(\bar{x}) + \dots + y_n \cdot L_n(\bar{x})$$

de modo que el ordenador calculará directamente los valores $L_0(\bar{x}), L_1(\bar{x}), \dots, L_n(\bar{x})$ y los sustituirá en la expresión anterior.

MÉTODOS NUMÉRICOS BÁSICOS

Ejemplo: Determinar, por el “método de Lagrange”, el polinomio interpolador correspondiente a los 4 puntos (0'4, -1'05), (0'5, 2'82), (0'8, 1'23) y (1, 0'46).

Los cuatro polinomios de Lagrange, de grado 3 para este caso, son:

$$L_0(x) = \frac{(x-0'5) \cdot (x-0'8) \cdot (x-1)}{(0'4-0'5) \cdot (0'4-0'8) \cdot (0'4-1)} \quad L_1(x) = \frac{(x-0'4) \cdot (x-0'8) \cdot (x-1)}{(0'5-0'4) \cdot (0'5-0'8) \cdot (0'5-1)}$$

$$L_2(x) = \frac{(x-0'4) \cdot (x-0'5) \cdot (x-1)}{(0'8-0'4) \cdot (0'8-0'5) \cdot (0'8-1)} \quad L_3(x) = \frac{(x-0'4) \cdot (x-0'5) \cdot (x-0'8)}{(1-0'4) \cdot (1-0'5) \cdot (1-0'8)}$$

Y el polinomio interpolador será entonces:

$$P_3(x) = -1'05 \cdot L_0(x) + 2'82 \cdot L_1(x) + 1'23 \cdot L_2(x) + 0'46 \cdot L_3(x)$$

MÉTODO DE DIFERENCIAS PROGRESIVAS DE NEWTON-GREGORY (Isaac Newton, matemático y físico inglés del siglo XVII) (David Gregory, matemático escocés del siglo XVII):

Es un método muy rápido, pero sólo puede aplicarse cuando las abscisas de los puntos dados formen una progresión aritmética, o sea que, en orden de menor a mayor, se tenga: $x_1 = x_0 + h$; $x_2 = x_1 + h$; $x_3 = x_2 + h$;; $x_n = x_{n-1} + h$ (con $h > 0$ fijo).

Entendemos a partir de ahora que los puntos dados están ya ordenados según los valores crecientes de sus abscisas: O sea que los puntos serán (x_0, y_0) , (x_1, y_1) , (x_2, y_2) , , (x_n, y_n) , cumpliéndose $x_0 < x_1 < x_2 < \dots < x_n$ y el paso de una abscisa a la siguiente se obtiene sumando la cantidad positiva constante h .

Consideramos ahora el siguiente **cuadro de “diferencias progresivas”**:

valores de y	diferencias Δy (de orden 1)	diferencias $\Delta^2 y$ (de orden 2)	dif. $\Delta^{n-1} y$ (de orden $n-1$)	dif. $\Delta^n y$ (de orden n)
y_0	$\Delta y_0 = y_1 - y_0$	$\Delta^2 y_0 = \Delta y_1 - \Delta y_0$	$\Delta^{n-1} y_0$	$\Delta^n y_0$
y_1	$\Delta y_1 = y_2 - y_1$	$\Delta^2 y_1 = \Delta y_2 - \Delta y_1$	$\Delta^{n-1} y_1$	
y_2			
....			
y_{n-2}	$\Delta y_{n-2} =$ $= y_{n-1} - y_{n-2}$	$\Delta^2 y_{n-2} =$ $= \Delta y_{n-1} - \Delta y_{n-2}$			
y_{n-1}	$\Delta y_{n-1} = y_n - y_{n-1}$				
y_n					

En la primera columna vemos las ordenadas de los puntos dados (son $n + 1$). La segunda columna incluye las diferencias progresivas de primer orden (son n). La tercera columna incluye las diferencias progresivas de orden 2 (son $n - 1$). Etcétera... Cada nueva columna incluirá las diferencias progresivas de orden una unidad mayor que las que aparecen en la columna precedente (hay una diferencia menos cada vez). O sea que en la cuarta columna (no escrita) estarán las $n - 2$ diferencias progresivas de orden 3, en la siguiente estarán las $n - 3$ diferencias de orden 4, etc... Siguiendo así llegaremos a la penúltima columna con solamente dos diferencias progresivas de orden $n - 1$ y en la última columna habrá una sola diferencia de orden n (esta diferencia $\Delta^n y_0$ es igual a la diferencia de las dos anteriores $\Delta^{n-1} y_1 - \Delta^{n-1} y_0$). Las diferencias se llaman “progresivas” porque siempre se resta así: “el valor siguiente menos el valor que corresponde”.

MÉTODOS NUMÉRICOS BÁSICOS

Pues bien, se demuestra que el “polinomio interpolador” correspondiente a los puntos dados es:

$$P_n(x) = y_0 + \frac{\Delta y_0}{1! \cdot h} \cdot (x - x_0) + \frac{\Delta^2 y_0}{2! \cdot h^2} \cdot (x - x_0) \cdot (x - x_1) + \frac{\Delta^3 y_0}{3! \cdot h^3} \cdot (x - x_0) \cdot (x - x_1) \cdot (x - x_2) + \dots + \frac{\Delta^n y_0}{n! \cdot h^n} \cdot (x - x_0) \cdot (x - x_1) \cdot \dots \cdot (x - x_{n-1})$$

Obsérvese que los numeradores de todas las fracciones que aparecen en la expresión anterior son las diferencias progresivas $\Delta y_0, \Delta^2 y_0, \Delta^3 y_0, \dots, \Delta^n y_0$ que aparecen en la primera fila del cuadro de diferencias obtenido (la correspondiente a y_0). Y los denominadores son los factoriales sucesivos multiplicados por las potencias sucesivas del paso h .

Se ve claramente que al sustituir en el polinomio x por x_0 se anulan todos los términos menos el primero, con lo cual se tiene $P_n(x_0) = y_0$. De modo análogo, al sustituir en el polinomio x por x_1 se anulan todos los términos menos los dos primeros, quedando

$$P_n(x_1) = y_0 + \frac{\Delta y_0}{1! \cdot h} \cdot (x_1 - x_0) = y_0 + \frac{\Delta y_0}{h} \cdot h = y_0 + \Delta y_0 = y_0 + (y_1 - y_0) = y_1$$

Es decir, se comprobó que $P_n(x_1) = y_1$. Del mismo modo, al sustituir en el polinomio x por x_2 se anulan todos los términos menos los tres primeros, llegando después de operar a $P_n(x_2) = y_2$. Y de modo análogo se puede comprobar que $P_n(x_3) = y_3, P_n(x_4) = y_4, \dots, P_n(x_n) = y_n$.

Ejemplo: Determinar, por el método de Newton-Gregory de diferencias progresivas, el polinomio interpolador correspondiente a los puntos $(2^3, 5), (3^1, -2), (2^7, 0), (2^9, 2)$ y $(2^5, -3)$.

Los puntos no están ordenados por sus abscisas crecientes. Ordenándolos, de la menor abscisa que es 2^3 hasta la mayor que es 3^1 , tenemos: $(x_0, y_0) = (2^3, 5); (x_1, y_1) = (2^5, -3); (x_2, y_2) = (2^7, 0); (x_3, y_3) = (2^9, 2); (x_4, y_4) = (3^1, -2)$. Ahora podemos aplicar el método, porque observamos que las abscisas están en progresión aritmética de paso $h = 0^2$. El polinomio interpolador será de grado menor o igual que 4, pues tenemos 5 puntos.

Hacemos el correspondiente cuadro de diferencias progresivas:

valores y	diferencias Δy	diferencias $\Delta^2 y$	diferencias $\Delta^3 y$	diferencia $\Delta^4 y$
5	$-3 - 5 = -8$	$3 - (-8) = 11$	$-1 - 11 = -12$	$-5 - (-12) = 7$
-3	$0 - (-3) = 3$	$2 - 3 = -1$	$-6 - (-1) = -5$	
0	$2 - 0 = 2$	$-4 - 2 = -6$		
2	$-2 - 2 = -4$			
-2				

Los valores que intervienen en la expresión del polinomio son los de la primera fila:

$$y_0 = 5; \Delta y_0 = -8; \Delta^2 y_0 = 11; \Delta^3 y_0 = -12; \Delta^4 y_0 = 7$$

Por tanto, el polinomio interpolador será:

$$P_4(x) = 5 + \frac{-8}{1! \cdot 0^2} \cdot (x - 2^3) + \frac{11}{2! \cdot 0^2^2} \cdot (x - 2^3) \cdot (x - 2^5) + \frac{-12}{3! \cdot 0^2^3} \cdot (x - 2^3) \cdot (x - 2^5) \cdot (x - 2^7) + \frac{7}{4! \cdot 0^2^4} \cdot (x - 2^3) \cdot (x - 2^5) \cdot (x - 2^7) \cdot (x - 2^9)$$

Como dijimos en el ejemplo anterior, no conviene perder tiempo haciendo operaciones, por lo cual introduciríamos así el polinomio en la calculadora o el ordenador. Lo que nos interesa normalmente es usar esta expresión para obtener el valor que toma el polinomio interpolador en cada abscisa que queramos (en este caso intermedia entre 2^3 y 3^1).

MÉTODOS NUMÉRICOS BÁSICOS

Como recomendación final sobre interpolación polinómica, diremos que es preferible usar siempre el método de Newton-Gregory al método de Lagrange, por ser el primero más rápido. Pero si los puntos dados no tienen sus abscisas en progresión aritmética, sólo es aplicable el método de Lagrange.

Métodos de resolución de ecuaciones

Recordamos que una función $f(x)$ se dice continua en un intervalo cerrado y acotado $[a, b]$ si es continua en todos sus puntos interiores, continua por la derecha en el extremo a y continua por la izquierda en el extremo b . En Análisis Matemático se demuestra el siguiente importante teorema visto en la Sección 2.4:

TEOREMA DE BOLZANO: Si $f(x)$ es una función continua en un intervalo $[a, b]$ y además dicha función toma valores de signos contrarios en los extremos de dicho intervalo, existe al menos un punto c intermedio del intervalo donde la función vale cero, el cual es entonces una solución de la ecuación $f(x) = 0$.

MÉTODO DE BISECCIÓN:

Busca obtener una aproximación suficientemente buena de alguna de las soluciones de una ecuación $f(x) = 0$, situadas en un cierto intervalo $[a, b]$ (el cual conviene que sea lo más pequeño posible, para que el procedimiento que vamos a aplicar no sea muy largo). Pero debemos estar seguros de que exista al menos una solución de la ecuación en el mencionado intervalo, para lo cual tendremos que comprobar que la función $y = f(x)$ cumple las hipótesis del Teorema de Bolzano en $[a, b]$. Por tanto, la función tendrá que ser continua en ese intervalo y además tiene que tomar valores de signos contrarios en sus extremos.

El “método de bisección” lo que hace es dividir el intervalo inicial en dos mitades (de ahí su nombre) y ver en cuál de ellas se siguen cumpliendo las hipótesis del Teorema de Bolzano, para lo cual se utiliza el punto central del intervalo que es $(a + b)/2$. Entonces, puede suceder que la función $f(x)$ se haga cero en ese punto central, con lo cual ya tendríamos el valor exacto de la solución buscada; pero normalmente la función será positiva o será negativa en dicho punto central, con lo cual en una sola de las dos mitades del intervalo se vuelve a dar el cambio de signo de la función (la cual sigue siendo continua en esa mitad). Entonces, **desecharemos el intervalo inicial, nos quedaremos como nuevo intervalo de partida con la mitad donde haya cambio de signo de la función y repetiremos el procedimiento.** Y en este nuevo paso obtendremos la solución exacta o determinaremos otro intervalo más pequeño (de longitud la cuarta parte del inicial) donde sigue habiendo con seguridad alguna solución de la ecuación. Y así sucesivamente... Es lo que se llama un “método de aproximaciones sucesivas”, que habrá que detener en algún momento.

Pues bien, el método se detendrá cuando encontremos la solución exacta o cuando se llegue a un intervalo de longitud menor o igual que un $\varepsilon > 0$ dado previamente, en cuyo caso podremos concluir que el extremo izquierdo de ese último intervalo será una aproximación por defecto de la solución c buscada, con error menor que ε ; el extremo derecho será una aproximación por exceso de dicha solución, también con error menor que ε , y el nuevo punto medio de ese último intervalo (si lo queremos obtener) también servirá como aproximación de la solución c buscada, con error menor que $\varepsilon/2$, suponiendo que la función no se anule en dicho valor medio (en cuyo caso habríamos encontrado en este último paso la solución c en forma exacta).

Nota: El Teorema de Bolzano garantiza la existencia de al menos una solución en el intervalo dado, pero no especifica cuántas hay. Además, las hipótesis del teorema son condiciones suficientes, pero no necesarias, para la existencia de alguna solución en el intervalo. Así, podría haber alguna solución en el intervalo sin que la función tuviese signos contrarios en sus extremos o sin

MÉTODOS NUMÉRICOS BÁSICOS

que la función fuese continua en dicho intervalo (o ambas cosas a la vez). Por tanto, al desechar en el método de bisección la mitad del intervalo en cuyos extremos la función toma valores con igual signo, podríamos estar perdiendo soluciones. Y al elegir la otra mitad, no es seguro que en ese nuevo intervalo haya una sola solución.

En cambio, si agregamos una tercera condición pidiendo que la derivada de la función exista en (a, b) y sea siempre diferente de cero, entonces **habrá una única solución c de la ecuación dada en dicho intervalo** (se demuestra por reducción al absurdo aplicando el Teorema de Rolle, visto en la Sección 3.1).

Ejemplo: Hallar valores aproximados por defecto y por exceso, con errores menores que 0.1 , de la solución de la ecuación cúbica $x^3 - 3x + 5 = 0$ que se encuentra en el intervalo $(-3, -2)$.

Aquí es $f(x) = x^3 - 3x + 5$, el intervalo inicial que debemos tomar será $[-3, -2]$ y el valor de ε es 0.1 . Se sabe que hay al menos una solución de la ecuación dada en el intervalo $[-3, -2]$, porque la función f es continua en todo \mathbb{R} (luego lo será en el intervalo dado) y además $f(-3) = -13$ y $f(-2) = 3$ (la función toma valores de signos contrarios en los extremos del intervalo). Y como la solución que buscamos estará en el intervalo cerrado, pero vemos que no está en -3 ni está en -2 , estará en el intervalo $(-3, -2)$.

Pero además $f'(x) = 3x^2 - 3$, con lo cual la derivada solamente se anula cuando $x^2 = 1$, o sea en los puntos $x = -1$ y $x = 1$ (ninguno perteneciente al intervalo $[-3, -2]$), con lo cual sabemos que dicha derivada será diferente de cero en todos los puntos de $(-3, -2)$. En consecuencia, en este caso hay una única solución de la ecuación dada en el intervalo señalado.

Empecemos el “método de bisección” con los datos que tenemos. Podemos ordenar los resultados que vamos obteniendo en un cuadro. Así tenemos:

paso	valores a	valores b	$b - a$	$(a + b)/2$	$f(a)$	$f(\frac{a+b}{2})$	$f(b)$
1	-3	-2	$1 > \varepsilon$	-2.5	-13	-3.125	3
2	-2.5	-2	$0.5 > \varepsilon$	-2.25	-3.125	0.359...	3
3	-2.5	-2.25	$0.25 > \varepsilon$	-2.375	-3.125	-1.271...	0.359...
4	-2.375	-2.25	$0.125 > \varepsilon$	-2.3125	-1.271...	-0.428...	0.359...
5	-2.3125	-2.25	$0.0625 < \varepsilon$	PARAMOS			

Conclusiones:

Aproximación por defecto de la solución exacta c , con error menor que 0.1 : $\boxed{-2.3125}$

Aproximación por exceso de la solución exacta c , con error menor que 0.1 : $\boxed{-2.25}$

Otra aproximación, con error en valor absoluto menor que 0.05 : $\boxed{-2.28125}$ (último punto medio)

Explicaciones: En el paso 1 el cambio de signo está entre el primer punto medio y b (ver tres últimos valores de la fila 1 del cuadro), luego el segundo intervalo (el del paso 2) ha cambiado el extremo a por el punto medio obtenido antes (-2.5) y deja b igual. Ahora, en el paso 2, lo primero que hallamos es la longitud $b - a$ del nuevo intervalo, que es 0.5 (muy lejos todavía de 0.1), luego seguimos los cálculos: Hallamos el nuevo punto medio (-2.25) y evaluamos con la calculadora la función en a , en $\frac{a+b}{2}$ y en b (tres últimos valores de la fila 2, dos de ellos ya obtenidos). Observamos que el cambio de signo está ahora entre a y el nuevo punto medio, luego el tercer intervalo (el del paso 3) cambia el extremo b por el punto medio del paso 2 y deja a igual. En el paso 3 la nueva longitud $b - a$ es 0.25 , luego seguimos los cálculos: Hallamos el nuevo punto medio (que es -2.375) y evaluamos con la calculadora la función otra vez en a , en $\frac{a+b}{2}$ y en b (tres últimos valores de la fila 3, dos de ellos ya obtenidos). Observamos ahora que el cambio de signo se produce nuevamente en la mitad derecha del intervalo, luego el cuarto intervalo (el del paso 4) cambia a por el último punto medio y deja b igual. En el paso 4 la nueva longitud es 0.125 que está cerca de $\varepsilon = 0.1$ pero todavía es mayor, luego seguimos los cálculos: Hallamos el nuevo punto medio

MÉTODOS NUMÉRICOS BÁSICOS

(-2'3125) y evaluamos la función otra vez en a , en $\frac{a+b}{2}$ y en b (tres últimos valores de la fila 4, dos de ellos ya obtenidos). Observamos cambio de signo otra vez en la mitad derecha del último intervalo, luego el quinto intervalo (el del paso 5) cambia a por el último punto medio y deja b igual. En el paso 5 la nueva longitud es 0'0625 que ya es menor que ε , luego paramos el proceso.

Nota 1: En la práctica tendremos la ecuación $f(x) = 0$ que queremos resolver, pero no siempre se tiene el intervalo $[a, b]$ donde hay que buscar una solución de la misma. Entonces lo primero que debe hacerse es buscar ese intervalo, lo más pequeño que se pueda, donde se cumplan las hipótesis del Teorema de Bolzano: Se puede empezar evaluando la función f en valores de x de su dominio que sean números enteros consecutivos, hasta encontrar cambio de signo entre dos seguidos. Entonces, si la función dada es continua en el intervalo correspondiente, ya podemos aplicar el “método de bisección” y estaremos comenzando con un intervalo de longitud 1, como en el ejemplo anterior.

Nota 2: En el caso de ecuaciones polinómicas, como la del ejemplo anterior, lo mejor es hacer divisiones del polinomio dado entre binomios de la forma $x - k$, mediante la regla de Ruffini, dándole a k valores enteros sucesivos hasta encontrar dos restos de signos contrarios (que son dos valores del polinomio).

Por ejemplo, en el caso de la ecuación cúbica que hemos resuelto, vemos que el valor de $x^3 - 3x + 5$ es 5 si hacemos $x = 0$, pero cuando x sea negativo de valor absoluto cada vez mayor, la función tendrá que ser negativa, ya que $\lim_{x \rightarrow -\infty} f(x) = -\infty$. Por tanto, probamos por ejemplo el valor $x = -4$ y al dividir por Ruffini el polinomio de grado 3 entre $x + 4$ se obtiene resto -47 (o sea $f(-4) = -47$). Esto nos indica que hay alguna solución de la ecuación entre -4 y 0 , pero este es un intervalo muy grande (habría que dar muchos más pasos en el “método de bisección” para llegar a una aproximación aceptable como la que obtuvimos). Entonces volvemos a dividir entre $x + 3$ y se obtiene resto -13 . Y finalmente, al dividir entre $x + 2$ se obtiene resto 3. Conclusión: Hay alguna solución c entre -3 y -2 (este es un buen intervalo de partida pues su longitud es 1 y fue el que usamos anteriormente).

Nota 3: En la práctica se suele usar una variante de este “método de bisección”, partiendo siempre de un intervalo de extremos enteros consecutivos (como hicimos antes) donde se cumplan las hipótesis del Teorema de Bolzano. Así a y b serán aproximaciones enteras por defecto y por exceso de la solución buscada, siendo a exactamente la parte entera de la solución c buscada. Entonces, en vez de dividir en dos el intervalo $[a, b]$, se divide en diez usando los puntos intermedios $a + 0'1, a + 0'2, \dots$ y $a + 0'9$. Se evalúa ahora la función en dichos puntos hasta encontrar cambio de signo en dos consecutivos, con lo cual tendremos ya aproximaciones por defecto y por exceso de la solución c con una cifra decimal (donde la menor de ellas nos determinará exactamente la cifra de las décimas de la solución). Si repetimos el procedimiento dividiendo en diez el nuevo intervalo obtenido y evaluando la función en los puntos intermedios hasta encontrar nuevo cambio de signo, tendremos aproximaciones de c con dos cifras decimales y la menor nos determinará la segunda cifra decimal exacta de la solución. Esto puede programarse en un ordenador, que nos dará rápidamente una aproximación de la solución con tantas cifras decimales exactas como queramos (las calculadoras científicas que permitan usar el método, ya incluirán el correspondiente programa).

Otro ejemplo: Resolvamos la ecuación no polinómica $x - \cos x = 0$ por el método de bisección mejorado citado en la Nota 3 anterior.

Si escribimos la ecuación como $x = \cos x$, podemos interpretar que la solución $x = c$ que buscamos es la abscisa del punto de corte entre las gráficas de las funciones $y = x$ e $y = \cos x$ (y si dibujamos dichas gráficas vemos que hay un único punto de corte, cuya abscisa está situada entre 0 y $\pi/2$). Luego se trata de hallar la solución única de la ecuación dada, situada en el intervalo

MÉTODOS NUMÉRICOS BÁSICOS

$[0, \pi/2]$. Pero para dividir en 10 subintervalos un intervalo inicial de longitud 1, debemos averiguar si la función $f(x) = x - \cos x$ es positiva o negativa en $x = 1$ (en $x = 0$ la función toma valor $0 - \cos 0 = -1$, luego si $f(1)$ fuese positivo, podríamos trabajar con el intervalo que va de 0 a 1, y si $f(1)$ fuese negativo habrá que trabajar con el intervalo que va de 1 a 2 (en el cual está $\pi/2$, donde sabemos que la función ya es positiva). Usamos la calculadora y obtenemos $f(1) = 0.4596... > 0$. Por tanto, se trata de buscar la solución c en el intervalo $[0, 1]$, donde $f(x)$ es continua (lo es en todo \mathbb{R}) y en cuyos extremos toma valores de signos contrarios (condiciones de la hipótesis del Teorema de Bolzano, lo cual garantiza su existencia). Pero además $f'(x)$ es la función $1 + \sin x$, que es diferente de cero en todo el intervalo $(0, 1)$ porque los valores de $\sin x$ son positivos o cero, luego la solución c es única (cosa que ya sabíamos, porque vimos que hay un solo punto de corte entre las gráficas mencionadas).

Falta evaluar $f(x) = x - \cos x$ en los puntos de abscisas $0.1, 0.2, 0.3$, etc... y detectar cuáles son dos consecutivos donde aparece cambio de signo, y así tendremos un nuevo intervalo cerrado de longitud 0.1 (de extremos los dos valores encontrados) donde vuelven a darse las condiciones del teorema de Bolzano. Hecho esto con una calculadora (muy fácilmente si la misma acepta la expresión $x - \cos x$ que iremos evaluando sucesivas veces), descubriremos que los valores en 0.1 hasta 0.7 son negativos y el valor en 0.8 es ya positivo ($f(0.7) = -0.064...$ y $f(0.8) = 0.103...$). Por tanto, la solución buscada está en el interior del intervalo $[0.7, 0.8]$ y podemos ya escribir ya $c = 0.7...$ con la cifra de las décimas exacta. Ahora dividimos el intervalo anterior en 10 subintervalos, evaluando la función en las nuevas abscisas $0.71, 0.72, 0.73$, etc... hasta encontrar nuevamente un cambio de signo. Así obtenemos valores negativos en $0.71, 0.72$ y 0.73 , viendo que el valor de la función en 0.74 es positivo ($f(0.73) = -0.015...$ y $f(0.74) = 0.001...$), con lo cual concluimos que la solución c está entre 0.73 y 0.74 , puediendo ya escribir $c = 0.73...$, con todas las dos cifras decimales exactas. Y siguiendo así iremos obteniendo nuevas cifras decimales exactas de la solución c de la ecuación dada (que debe estar más cerca de 0.74 que de 0.73 , por los valores dados anteriormente).

MÉTODO DE NEWTON-RAPHSON (Joseph Raphson, matemático inglés que vivió entre los siglos XVII y XVIII):

Este método es más rápido que el anterior. Se trata otra vez de buscar una solución de la ecuación $f(x) = 0$ en un intervalo $[a, b]$, de modo que se cumplan las hipótesis del Teorema de Bolzano. Pero, para garantizar el buen funcionamiento de este nuevo método, se exigen dos condiciones más: 1) Que la derivada de la función f no se anule en el intervalo cerrado $[a, b]$. Y 2) Que la derivada segunda de f mantenga un signo fijo en todo $[a, b]$.

El método consiste en aplicar una “fórmula recurrente” (o sea, una fórmula que se utiliza sucesivas veces) a partir de un punto inicial x_0 del intervalo, el cual representa una primera aproximación de la solución c que buscamos (la cual sabemos que es única en el intervalo dado, pues hemos supuesto que la derivada primera de la función no se hace cero en el mismo).

La fórmula recurrente es: $x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$, para todo entero positivo n .

A partir del punto inicial x_0 permite obtener otro punto x_1 (usándola para $n = 1$) y a partir de x_1 nos da otro punto x_2 (usándola para $n = 2$), etcétera... Pues bien, se demuestra que, eligiendo convenientemente el punto inicial x_0 , los puntos sucesivos x_1, x_2, \dots que se van obteniendo quedan siempre en el intervalo $[a, b]$ y constituyen una “sucesión de números reales” que converge hacia la solución que buscamos. O sea, $\lim_{n \rightarrow \infty} x_n = c$. (Ver Sección 3.8). Obsérvese que, al haber supuesto que la derivada primera de la función no se anula en ese intervalo, ningún denominador de la “fórmula recurrente” valdrá cero, con lo cual podremos aplicar dicha fórmula una vez tras otra sin problemas.

Este método es también “de aproximaciones sucesivas” (como el “Método de bisección”) y también se dice que es “un método iterativo” por la aplicación reiterada (o “iterada”) de una misma fórmula.

MÉTODOS NUMÉRICOS BÁSICOS

Falta determinar la aproximación inicial x_0 conveniente y un criterio para parar su aplicación.

Para determinar la aproximación inicial x_0 conveniente, se procede así: Cuando $f(a)$ tenga el mismo signo que $f''(x)$, debemos tomar esa aproximación inicial en el punto a (o sea, $x_0 = a$). Y cuando $f(b)$ tenga el mismo signo que $f''(x)$, tomaremos dicha aproximación inicial en b (o sea, $x_0 = b$). Y esto funciona siempre, pues hemos supuesto que $f''(x)$ mantiene un signo fijo en todo el intervalo y como $f(a)$ y $f(b)$ tienen signos contrarios, el signo de f'' coincidirá necesariamente con el de $f(a)$ o con el de $f(b)$.

Como dijimos, se demuestra que de este modo las aproximaciones x_0, x_1, x_2, \dots obtenidas se acercan cada vez más al valor desconocido de la solución exacta c de la ecuación $f(x) = 0$ en el intervalo dado, con lo cual x_n será una buena aproximación de c si tomamos n suficientemente grande. Además, cuando se tome x_0 en a , dichas aproximaciones crecen hacia la solución, con lo cual serán siempre “aproximaciones por defecto” de c ; mientras que cuando se tome x_0 en b , lo harán decreciendo hacia la solución y entonces serán todas “aproximaciones por exceso” de c .

Pero, además, se demuestra que el error cometido al tomar el valor de x_n como aproximación de c , es en valor absoluto menor o igual que $\frac{|f(x_n)|}{m_1}$, donde m_1 es el mínimo absoluto de $|f'(x)|$ en $[a, b]$ (o un número positivo lo mayor posible que se mantenga menor que todos los valores de $|f'(x)|$ en el intervalo). Así tenemos un criterio para saber cuándo debemos parar las aplicaciones sucesivas de la “fórmula recurrente”: Cuando la expresión mencionada anteriormente sea **menor** que el $\varepsilon > 0$ dado (o elegido), tomaremos la última aproximación hallada (x_n) como solución aproximada de la ecuación en $[a, b]$ (sabiendo que el valor absoluto del error cometido de este modo será menor que ε).

Para el cálculo de m_1 (mínimo de $|f'(x)|$ en $[a, b]$) debemos tener en cuenta los signos de $f'(x)$ y de $f''(x)$ en el intervalo. En general hay cuatro posibilidades y habrá que ver en cuál nos encontramos al resolver cada problema concreto:

1) $f'(x)$ positiva y $f''(x)$ positiva: Entonces será $|f'(x)| = f'(x)$ y, al ser $f''(x)$ positiva, será $f'(x)$ estrictamente creciente. Conclusión: $m_1 = f'(a)$.

2) $f'(x)$ positiva y $f''(x)$ negativa: Volverá a ser $|f'(x)| = f'(x)$ y, como $f''(x)$ es negativa, será $f'(x)$ estrictamente decreciente. Conclusión: $m_1 = f'(b)$.

3) $f'(x)$ negativa y $f''(x)$ positiva: Entonces $|f'(x)| = -f'(x)$ y, como $f''(x)$ es positiva, será $f'(x)$ estrictamente creciente, con lo cual $-f'(x)$ será estrictamente decreciente. Conclusión: $m_1 = -f'(b)$.

4) $f'(x)$ negativa y $f''(x)$ negativa: Volverá a ser $|f'(x)| = -f'(x)$ y, como $f''(x)$ es negativa, será $f'(x)$ estrictamente decreciente, con lo cual $-f'(x)$ será estrictamente creciente. Conclusión: $m_1 = -f'(a)$.

Nota: Sabemos que $f'(x)$ no se hace cero en el intervalo $[a, b]$ (condición exigida desde el principio), con lo cual necesariamente mantendrá $f'(x)$ un mismo signo en todo el intervalo (o sea, todos sus valores serán positivos o todos sus valores serán negativos). Porque si hubiese dos puntos en el intervalo donde la derivada tuviese signos contrarios, siendo continua (por existir f''), según el Teorema de Bolzano tendría que dar valor cero en algún punto intermedio (y hemos supuesto que esto no ocurre).

En cada caso concreto, una vez hallados los valores x_0 y m_1 que correspondan, se sustituyen las expresiones de $f(x_{n-1})$ y $f'(x_{n-1})$ en la fórmula recurrente y podremos comenzar el proceso de cálculo de las aproximaciones sucesivas de acuerdo a este posible esquema:

MÉTODOS NUMÉRICOS BÁSICOS

paso	x_{n-1}	$f(x_{n-1})$	$\frac{ f(x_{n-1}) }{m_1}$	$f'(x_{n-1})$	$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$
1 ($n = 1$)	x_0	$f(x_0)$	$\frac{ f(x_0) }{m_1}$	$f'(x_0)$	$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$
2 ($n = 2$)	x_1	$f(x_1)$	$\frac{ f(x_1) }{m_1}$	$f'(x_1)$	$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$
3 ($n = 3$)	x_2	$f(x_2)$	$\frac{ f(x_2) }{m_1}$	$f'(x_2)$	$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}$
...

Explicación: Empezamos con x_0 y calculamos $f(x_0)$. Entonces, si $|f(x_0)|/m_1$ no es menor que ε , calculamos x_1 aplicando por primera vez la “fórmula recurrente”. Y si $|f(x_0)|/m_1$ es menor que ε , paramos el proceso y la solución aproximada será x_0 . Ahora, si hemos calculado x_1 , obtenemos $f(x_1)$ y si $|f(x_1)|/m_1$ no es menor que ε , calcularemos x_2 aplicando la “fórmula recurrente” por segunda vez, y en caso contrario paramos el proceso tomando x_1 como solución aproximada. Etcétera... Y cuando lleguemos a $\frac{|f(x_n)|}{m_1} < \varepsilon$ pararemos el proceso en ese paso (que será el $n + 1$) sin aplicar de nuevo la “fórmula recurrente” (en cada problema el número de pasos a dar será diferente, en general). Y entonces se toma el valor x_n obtenido en el paso anterior como solución aproximada de la ecuación $f(x) = 0$ en el intervalo $[a, b]$, teniendo un error en valor absoluto menor que ε . Además, x_n será “una solución aproximada por defecto” si hemos partido de $x_0 = a$ y será “una solución aproximada por exceso” si hemos tomado $x_0 = b$.

Nota importante: Este método también se llama “**método de las tangentes**” pues para pasar de cada aproximación x_{n-1} a la siguiente x_n , lo que se hace geoméricamente es hallar la recta tangente a la gráfica de la función $y = f(x)$ en el punto de abscisa x_{n-1} (ya conocida) y obtener su corte con el eje OX (que nos dará la abscisa x_n).

En efecto, la ecuación de dicha recta tangente es $y - f(x_{n-1}) = f'(x_{n-1}) \cdot (x - x_{n-1})$, y para obtener su punto de corte con el eje OX habrá que hacer $y = 0$ despejando la variable x . Así se obtiene $x = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$, que es la “fórmula recurrente” usada para hallar x_n .

Se recomienda al lector ponerse ejemplos gráficos en diferentes situaciones (combinando función estrictamente creciente y función estrictamente decreciente con función cóncava hacia arriba y función cóncava hacia abajo), eligiendo en cada caso el punto x_0 adecuado y comprobando gráficamente en todos los ejemplos que los puntos de corte con OX de las sucesivas rectas tangentes dibujadas se aproximan muy rápidamente al corte de la gráfica con dicho eje (el punto $x = c$).

Ejemplo: Resolvamos nuevamente la ecuación cúbica $x^3 - 3x + 5 = 0$, en el mismo intervalo $[-3, -2]$, con error en valor absoluto menor que $0,1$, aplicando el “método de las tangentes” o “de Newton-Raphson” (anteriormente habíamos resuelto este problema por el “método de bisección” en 5 pasos, incluido el de parada; ver pág. 7).

Entonces ya sabemos que la función $f(x) = x^3 - 3x + 5$ cumple las hipótesis del Teorema de Bolzano en el intervalo $[-3, -2]$. Y ya habíamos visto que se cumple la primera de las dos condiciones adicionales exigidas en este nuevo método: $f'(x) = 3x^2 - 3$ no se anula en ningún punto del intervalo $[-3, -2]$. Y en cuanto a la segunda condición, vemos que $f''(x) = 6x$ es siempre negativa en el intervalo. Con lo cual, podemos comprobar que $f'(x)$ se mantiene positiva en dicho intervalo hallando solamente uno de sus valores en el mismo (por ejemplo, $f'(-2) = 9 > 0$).

Podemos entonces aplicar el método de Newton-Raphson: Tenemos $x_0 = -3$ porque recordemos que el valor de la función en ese extremo del intervalo es -13 (mismo signo que el de la derivada

MÉTODOS NUMÉRICOS BÁSICOS

segunda en el intervalo). Y además, teniendo en cuenta que $f'(x)$ es positiva y que $f''(x)$ es negativa, será $m_1 = f'(b) = f'(-2) = 9$ (ver la pág. 10).

La “fórmula recurrente” en este caso es: $x_n = x_{n-1} - \frac{x_{n-1}^3 - 3x_{n-1} + 5}{3x_{n-1}^2 - 3}$ (habría que introducir esta fórmula en la calculadora científica, para usarla sucesivas veces).

Pues bien, teniendo ya la fórmula recurrente, la aproximación inicial x_0 , el número m_1 y la cota superior del error en valor absoluto ($\varepsilon = 0'1$), podemos iniciar los cálculos (que hacemos, por ejemplo, con 7 cifras decimales y redondeo para la presentación que sigue, aunque mejor es usar todas las cifras decimales que dé la calculadora):

paso	x_{n-1}	$f(x_{n-1})$	$ f(x_{n-1}) /m_1$	$f'(x_{n-1})$	x_n
$n = 1$	$x_0 = -3$	$f(x_0) = -13$	$ f(x_0) /9 = 1'4444444 > \varepsilon$	$f'(x_0) = 24$	$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = -2'4583333$
$n = 2$	$x_1 = -2'4583333$	$f(x_1) = -2'4816980$	$ f(x_1) /9 = 0'2757442 > \varepsilon$	$f'(x_1) = 24'38082$	$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = -2'3565444$
$n = 3$	$x_2 = -2'3565444$	$f(x_2) = -1'0169684$	$ f(x_2) /9 = 0'1129964 > \varepsilon$	$f'(x_2) = 13'659905$	$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = -2'2820952$
$n = 4$	$x_3 = -2'2820952$	$f(x_3) = -0'0387715$	$ f(x_3) /9 = 0'0043079 < \varepsilon$	PARAMOS	

Obsérvese que con sólo tres aplicaciones de la fórmula recurrente (4 pasos incluyendo el de parada) finalizó el proceso, puesto que la última aproximación calculada, $x_3 = -2'2820952$, es ya suficientemente buena para el ε que hemos fijado (además sabemos que se trata de una aproximación por defecto, pues hemos partido del extremo izquierdo del intervalo). En efecto, el error cometido con esa aproximación (respecto al valor exacto, desconocido, de la solución c que buscamos) es en valor absoluto menor o igual que $0'0043079$, que es mucho menor que $0'1$. Sin embargo, por el “método de bisección” tuvimos que dar 5 pasos incluyendo el de parada y llegamos a la aproximación $-2'28125$ con un error en valor absoluto menor que $0'03125$ (sin saber si esta aproximación es por defecto o es por exceso).

Conclusión: $-2'2820952$ es una “aproximación por defecto” de la solución exacta de la ecuación $x^3 - 3x + 5 = 0$ en el intervalo $[-3, -2]$, con error menor que $0'1$ (en realidad, con error menor que $0'0043079$). Luego $-2'2820952 < c < -2'2820952 + 0'0043079$, siendo esta última suma $-2'2777873$, con lo cual podemos tomar $c \cong -2'28$ sin tantas cifras decimales y siendo exacta la de las décimas; en cambio, por el método de bisección teníamos $-2'28125 - 0'03125 < c < -2'28125 + 0'03125$, que operando es $-2'3125 < c < -2'25$, con lo cual no podemos asegurar ni siquiera la primera cifra decimal de c).

Nota: En la práctica, se trabaja normalmente con valores de ε bastante menores que $0'1$, con lo cual habrá que usar más veces la “fórmula recurrente”, pero con este método obtendremos suficientes cifras decimales exactas del valor desconocido c con pocas aplicaciones adicionales de la misma. Por ejemplo, si tomamos $\varepsilon = 0'001$ para obtener al menos dos cifras decimales exactas de la solución, y aplicamos la fórmula recurrente sólo una vez más, se llega a la aproximación $x_4 = -2'2790239$ que tiene un error menor que $0'0000071$. Es decir, que con una aplicación más de la fórmula recurrente se ha obtenido una aproximación que ya nos da 4 cifras decimales exactas del valor de c , puesto que $-2'2790239 < c < -2'2790239 + 0'0000071$, o sea que c está entre $-2'2790239$ y $-2'2790168$, lo cual nos dice que podemos dar definitivamente el valor de la solución como $c \cong -2'2790$ con todas sus cifras exactas.

MÉTODOS NUMÉRICOS BÁSICOS

Como recomendación final sobre resolución de ecuaciones, diremos que es preferible usar siempre el método de Newton-Raphson al de bisección, por ser mucho más rápido. En cualquiera de ellos hay que verificar las hipótesis del Teorema de Bolzano, pero para usar Newton-Raphson la función que esté igualada a cero debe cumplir, en el intervalo donde busquemos la solución, las dos condiciones extra mencionadas para tener garantía del buen funcionamiento del método. A veces esto se consigue acortando el intervalo. Cuando se aplique el método de Newton-Raphson cumpliéndose solamente las hipótesis del T. de Bolzano puede haber fallos (el método se puede quedar bloqueado porque algún denominador valga cero; puede dar las mismas aproximaciones varias veces, sin poder avanzar, o bien puede dar aproximaciones que terminen saliéndose del intervalo dado). Por lo tanto, en esos casos lo seguro es aplicar el método de bisección.

Métodos numéricos para calcular integrales definidas

La definición de una integral definida como límite de un sumatorio, nos permite obtener dos métodos de cálculo numérico para llegar a un resultado aproximado del valor de la integral: “El método de los rectángulos”, con tres variantes, y “el método de los trapecios”. (Ver al respecto la Sección 4.2). Los recordamos y añadimos un tercer método más eficiente: “El método de Simpson”.

Si queremos definir $\int_a^b f(x)dx$, con $f(x)$ continua en el intervalo $[a, b]$, tomaremos “una partición” P_n del intervalo de integración, con $n + 1$ puntos en total, que determinen n subintervalos de igual longitud h , de modo que los puntos de la partición queden en progresión aritmética de paso h ; serán $x_0 = a, x_1 = x_0 + h, x_2 = x_1 + h, x_3 = x_2 + h, \dots, x_n = x_{n-1} + h$, resultando $x_n = b$. El valor de h es igual a la longitud total del intervalo $[a, b]$ dividida entre el número de subintervalos de la partición, o sea $h = (b - a)/n$.

De este modo se tiene la definición de la integral definida:

$$\int_a^b f(x)dx = \lim_{n \rightarrow +\infty} \sum_{i=1}^n f(t_i) \cdot h = \lim_{n \rightarrow +\infty} \frac{b-a}{n} \cdot [f(t_1) + f(t_2) + \dots + f(t_n)]$$

donde t_1 es algún punto del subintervalo $[x_0, x_1]$, t_2 es algún punto del subintervalo $[x_1, x_2]$, \dots y t_n es algún punto del subintervalo $[x_{n-1}, x_n]$.

Se demuestra que el límite anterior no depende de la elección que se haga de los puntos t_1, t_2, \dots, t_n . En efecto, al crecer n , crece el número de sumandos y disminuye h (la longitud de cualquier subintervalo); entonces, los diferentes puntos t_i que podamos elegir en el subintervalo $[x_{i-1}, x_i]$ estarán cada vez más próximos entre sí, con lo cual los valores $f(t_i)$ serán cada vez más parecidos, por ser $f(x)$ continua en $[a, b]$ (al ser la función continua en el intervalo, pequeñas variaciones de x producen pequeñas variaciones de los valores de la función a lo largo de todo el intervalo; esto no ocurre en cualquier intervalo, pero sí ocurre cuando el intervalo es cerrado y acotado). Y lo anterior vale para $i = 1, 2, \dots, n$.

Lo dicho anteriormente nos permite escribir la definición de la integral de muchas formas diferentes, según la elección que hagamos de los puntos t_i en los subintervalos.

1) Eligiendo los puntos t_i en los extremos izquierdos de los subintervalos $[x_{i-1}, x_i]$, se tiene:

$$\int_a^b f(x)dx = \lim_{n \rightarrow +\infty} \frac{b-a}{n} \cdot [f(x_0) + f(x_1) + \dots + f(x_{n-1})]$$

(intervienen x_0, x_1, \dots, x_{n-1} y falta x_n).

2) Eligiendo los puntos t_i en el extremos derechos de los subintervalos $[x_{i-1}, x_i]$, se tiene:

MÉTODOS NUMÉRICOS BÁSICOS

$$\int_a^b f(x)dx = \lim_{n \rightarrow +\infty} \frac{b-a}{n} \cdot [f(x_1) + f(x_2) + \dots + f(x_n)]$$

(intervienen x_1, x_2, \dots, x_n y falta x_0).

3) Eligiendo los puntos t_i en los puntos centrales de los subintervalos $[x_{i-1}, x_i]$, se tiene:

$$\int_a^b f(x)dx = \lim_{n \rightarrow +\infty} \frac{b-a}{n} \cdot [f(\frac{x_0+x_1}{2}) + f(\frac{x_1+x_2}{2}) + \dots + f(\frac{x_{n-1}+x_n}{2})]$$

MÉTODO DE LOS RECTÁNGULOS:

En cualquiera de las formas anteriores de la definición de la integral, podemos escribir que un valor aproximado de la misma es el dado directamente por la expresión de la derecha, a la cual habría que calcular el límite para obtener el valor de la integral, con tal que tomemos en dicha expresión un valor de n suficientemente grande.

Tenemos así las siguientes tres “fórmulas de aproximación”, que fundamentan “el método de los rectángulos” (en sus tres variantes), donde hemos llamado $y_0 = f(x_0)$, $y_1 = f(x_1)$, $y_2 = f(x_2)$, \dots , $y_{n-1} = f(x_{n-1})$ e $y_n = f(x_n)$:

$$\int_a^b f(x)dx \cong h \cdot (y_0 + y_1 + \dots + y_{n-1}) \quad (1^{\text{a}} \text{ fórmula de los rectángulos})$$

$$\int_a^b f(x)dx \cong h \cdot (y_1 + y_2 + \dots + y_n) \quad (2^{\text{a}} \text{ fórmula de los rectángulos})$$

$$\int_a^b f(x)dx \cong h \cdot [f(\frac{x_0+x_1}{2}) + f(\frac{x_1+x_2}{2}) + \dots + f(\frac{x_{n-1}+x_n}{2})] \quad (3^{\text{a}} \text{ fórmula de los rectángulos})$$

El inconveniente de esta última fórmula es que necesita más cálculos que las anteriores, pues la función integrando no se evalúa directamente en los puntos de la partición (la cual debe haberse obtenido previamente), sino que se evalúa en los puntos medios de los correspondientes subintervalos (los cuales hay que calcular adicionalmente). Además, hay ocasiones en que no se dispone de la expresión analítica de la función integrando, sino sólo tenemos una tabla de valores de la misma en los puntos equidistantes de una partición del intervalo de integración, con lo cual pueden aplicarse las dos primeras fórmulas pero no la tercera.

Ejemplo: Hagamos el cálculo aproximado de la integral $\int_0^1 \frac{dx}{x^2+1}$, cuyo valor exacto es $\pi/4$ (obtenido aplicando la Regla de Barrow), mediante las tres “fórmulas de los rectángulos”, usando para ello una partición de 11 puntos equidistantes del intervalo de integración ($n = 10$).

El paso de un punto a otro de la partición es $h = (1 - 0)/10 = 0.1$. Por tanto, la partición será $P_{10} = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$. Pues bien, los valores de la función integrando, $f(x) = 1/(x^2 + 1)$, en los puntos anteriores son (limitándonos a seis cifras decimales con redondeo para abreviar, aunque mejor sería trabajar con todas las cifras que nos dé la calculadora):

$$y_0 = 1 ; y_1 = 0.990099 ; y_2 = 0.961538 ; y_3 = 0.917431 ; y_4 = 0.862069 ; y_5 = 0.8$$

$$y_6 = 0.735294 ; y_7 = 0.671141 ; y_8 = 0.609756 ; y_9 = 0.552486 ; y_{10} = 0.5$$

Entonces, el cálculo aproximado de la integral usando la 1ª fórmula de los rectángulos es:

$$I \cong 0.1 \cdot (y_0 + y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 + y_9) = \boxed{0.8099814}$$

MÉTODOS NUMÉRICOS BÁSICOS

El cálculo aproximado de la integral usando la 2ª fórmula de los rectángulos es:

$$I \cong 0'1 \cdot (y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 + y_9 + y_{10}) = \boxed{0'7599814}$$

Y para aplicar la 3ª fórmula de los rectángulos hay que calcular los valores de $f(x)$ en los puntos medios de los 10 subintervalos, que son:

$$\begin{aligned} f(0'05) &= 0'997506 ; f(0'15) = 0'977995 ; f(0'25) = 0'941176 \\ f(0'35) &= 0'890869 ; f(0'45) = 0'831601 ; f(0'55) = 0'767754 \\ f(0'65) &= 0'702988 ; f(0'75) = 0'64 ; f(0'85) = 0'580552 ; f(0'95) = 0'525624 \end{aligned}$$

Entonces, el cálculo aproximado de la integral usando la 3ª fórmula de los rectángulos es:

$$I \cong 0'1 \cdot [f(0'05) + f(0'15) + f(0'25) + f(0'35) + f(0'45) + f(0'55) + f(0'65) + f(0'75) + f(0'85) + f(0'95)] = \boxed{0'7856065}$$

El valor exacto de la integral es $\pi/4 = 0'78539816\dots$, con lo cual vemos que la 1ª fórmula dio una “aproximación por exceso”, la 2ª fórmula dio una “aproximación por defecto” y la 3ª fórmula aproximó “por exceso”, pero dio un resultado bastante mejor que las dos primeras. Los errores en valor absoluto de los correspondientes resultados son: 0'024... para el primero; 0'025... para el segundo, y 0'0002... para el tercero.

En general podemos decir que si la función del integrando es “estrictamente creciente” en el intervalo de integración, la 1ª aproximación será siempre “por defecto” y la 2ª aproximación será siempre “por exceso”. Y si la función es “estrictamente decreciente” en el intervalo, ocurrirá lo contrario. Así, en el ejemplo anterior, la función tiene derivada $f'(x) = -2x/(x^2 + 1)^2$, que es negativa en el intervalo $(0, 1)$, siendo f continua en $[0, 1]$, luego podemos concluir que la función es “estrictamente decreciente” en el intervalo cerrado $[0, 1]$ (como se nota en los valores obtenidos), por lo cual la 1ª aproximación ha resultado “por exceso” y la 2ª aproximación ha resultado “por defecto”, como hemos visto.

Pero, cuando la función integrando tenga tramos en crecimiento y tramos en decrecimiento en el intervalo de integración, no podrá saberse previamente el sentido de las aproximaciones dadas por las fórmulas 1ª y 2ª de los rectángulos.

Tampoco es fácil pronosticar en general el sentido de la aproximación dada por la 3ª fórmula de los rectángulos. Pero si el integrando es estrictamente creciente o estrictamente decreciente, podemos asegurar que su resultado estará entre los resultados de las otras dos fórmulas, como puede comprobarse en el ejemplo anterior.

Lo que siempre podremos asegurar es que las aproximaciones obtenidas, por cualquiera de las fórmulas, se acercan más al resultado exacto de la integral cuando n se tome cada vez mayor, o sea cuando se aumenta el número de puntos de la partición utilizada (obvio, pues el valor exacto de la integral es el límite de cualquiera de las fórmulas cuando $n \rightarrow \infty$).

Nota importante: Este método se llama “de los rectángulos” porque la fórmula utilizada en cualquiera de sus variantes puede escribirse

$$I \cong h \cdot f(t_1) + h \cdot f(t_2) + \dots + h \cdot f(t_n)$$

la cual representa la suma de las “áreas orientadas” de n rectángulos (todos de base h y cuyas alturas son los valores $f(t_1), f(t_2), \dots, f(t_n)$, de modo que cuando los valores anteriores sean todos positivos corresponderán a rectángulos que estarán por encima del eje OX y sus “áreas orientadas” se tomarán positivas, ocurriendo que la suma de esas áreas es claramente una aproximación del área de la **única región** limitada por la gráfica de $y = f(x)$, el eje OX y las rectas

MÉTODOS NUMÉRICOS BÁSICOS

verticales $x = a$ y $x = b$, la cual es justamente el valor positivo de la integral definida de $f(x)$ en el intervalo $[a, b]$.

Y cuando los valores anteriores de la función sean todos negativos, corresponderán a rectángulos que estarán por debajo del eje OX y sus correspondientes “áreas orientadas” se tomarán negativas, ocurriendo también que la suma de dichos valores negativos es una clara aproximación del valor de la integral definida, el cual será el valor opuesto (y por tanto, negativo), del área limitada por la única región limitada por la gráfica de la función, el eje OX y las rectas verticales por los extremos del intervalo. (Región que estará en este caso por debajo de OX; ver Sección 4.2).

Y cuando haya valores positivos y valores negativos entre los que toma $f(x)$ en los puntos t_i de los n subintervalos de la partición, habrá unos rectángulos por encima de OX (con “área orientada” positiva) y habrá otros rectángulos por debajo de dicho eje (con “área orientada” negativa), siendo también la suma de todas esas “áreas orientadas” una aproximación del valor de la integral (que en estos casos es la suma de las “áreas orientadas” de las varias regiones limitadas por la gráfica, el eje OX y las rectas verticales $x = a$ y $x = b$, tomando con signo positivo las que queden por encima de OX y tomando con signo negativo las que queden por debajo de OX, con lo cual la integral tendrá finalmente valor positivo, negativo o cero, según sea el reparto de áreas por encima y por debajo de OX). Téngase en cuenta que la separación entre unas y otras de estas regiones que se tomarán con signos distintos, serán puntos de corte que tenga la gráfica con el eje.

MÉTODO DE LOS TRAPECIOS:

Consiste en obtener la media aritmética de los resultados de las fórmulas 1ª y 2ª de los rectángulos, con lo cual se compensan en parte los errores que ocasionan estas fórmulas.

Considerando de nuevo una determinada partición P_n del intervalo de integración, con $n + 1$ puntos equidistantes (o sea, en progresión aritmética de paso $h = (b - a)/n$) y hallando igualmente los valores de la función integrando en todos esos puntos (como hicimos para aplicar las dos primeras variantes de la “fórmula de los rectángulos”), donde ahora llamamos y_i al valor $f(x_i)$ para $i = 0, 1, 2, \dots, n$, podremos aplicar directamente la siguiente “fórmula de los trapecios”:

$$\int_a^b f(x) dx \cong \frac{h}{2} \cdot (E + 2R)$$

que nos dará también una aproximación del valor de la integral definida de la función $f(x)$ en el intervalo $[a, b]$, siendo $E = y_0 + y_n$ (suma de ordenadas extremas) y $R = y_1 + y_2 + \dots + y_{n-1}$ (suma de las ordenadas restantes).

Pues bien, la media aritmética de las dos primeras “fórmulas de los rectángulos” da exactamente el mismo valor anterior:

$$\begin{aligned} & \frac{1}{2} \cdot [h \cdot (y_0 + y_1 + \dots + y_{n-1}) + h \cdot (y_1 + \dots + y_{n-1} + y_n)] = \\ & \frac{h}{2} \cdot (y_0 + 2y_1 + \dots + 2y_{n-1} + y_n) = \frac{h}{2} \cdot (E + 2R) \end{aligned}$$

Entonces, si ya hubiésemos obtenido las aproximaciones de la integral dada utilizando las dos primeras “fórmulas de los rectángulos”, bastará hallar la media aritmética de esos dos resultados para tener la aproximación que da la “fórmula de los trapecios”. Pero en caso contrario, aplicaremos en forma directa la nueva fórmula de aproximación dada anteriormente.

NOTA IMPORTANTE: Este método se llama “de los trapecios” porque la fórmula utilizada puede escribirse

$$I \cong \frac{y_0 + y_1}{2} \cdot h + \frac{y_1 + y_2}{2} \cdot h + \frac{y_2 + y_3}{2} \cdot h + \dots + \frac{y_{n-2} + y_{n-1}}{2} \cdot h + \frac{y_{n-1} + y_n}{2} \cdot h$$

la cual representa la suma de las “áreas orientadas” de n trapecios colocados verticalmente (todos de altura h y cuyas bases son las dos ordenadas que aparecen en los numeradores; recuérdese que

MÉTODOS NUMÉRICOS BÁSICOS

el área de un trapecio es la semisuma de sus bases multiplicada por la altura). En efecto, al sacar factor común $h/2$ en la expresión dada anteriormente, queda:

$$I \cong \frac{h}{2} \cdot (y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n) = \frac{h}{2} \cdot (E + 2R)$$

Se recomienda dibujar algún caso de gráfica sencilla de una función $y = f(x)$, continua en un cierto intervalo $[a, b]$, que tenga por ejemplo valores positivos en dicho intervalo, eligiendo además una partición del mismo con subintervalos que tengan igual longitud. Y a partir de estos datos, dibujar los trapecios sucesivos que correspondan, viendo cómo la suma de sus áreas darán una aproximación del valor de la correspondiente integral definida. (Igual sucedería si los valores de la función integrando fuesen todos negativos).

Nota: Cuando la gráfica de la función integrando pase de un lado a otro del eje OX, por tener esa función valores positivos y negativos en el intervalo, sabemos que habrá al menos un punto de corte $(c, 0)$ de dicha gráfica con OX (por el Teorema de Bolzano). Pues bien, habrá la posibilidad de que el punto c coincida con uno de los puntos de la partición P_n que estemos tomando, pero para otras particiones ya no ocurrirá esta casualidad.

Entonces, si c fuese punto interior del subintervalo $[x_{i-1}, x_i]$ de la partición que hayamos elegido, la expresión $\frac{y_{i-1} + y_i}{2} \cdot h$ no corresponderá al área de un trapecio sino a la suma de las “áreas orientadas” de dos rectángulos a ambos lados de c (uno de base $h/2$ y altura y_{i-1} , cuyo “área orientada” será $\frac{h \cdot y_{i-1}}{2}$, y el otro de base $h/2$ y altura y_i , cuyo “área orientada” será $\frac{h \cdot y_i}{2}$). Y como sus alturas son de signos contrarios, podremos considerar que el “área orientada” de uno de esos rectángulos quede agregada a la suma de “áreas orientadas” de trapecios que tengan su mismo signo y el “área orientada” del otro rectángulo quede agregada también a la suma de “áreas orientadas” de trapecios de su mismo signo.

De todos modos, esas dos “áreas orientadas” de rectángulos que correspondan a puntos de corte de la función integrando con el eje OX, tienden a neutralizarse cuando h se tome muy pequeño, pues las bases de dichos rectángulos coinciden y sus alturas son valores consecutivos de la función $f(x)$ a ambos lados de un corte c con OX, con lo cual dichas alturas serán de signos contrarios y muy parecidas en valor absoluto (ambas muy cercanas a cero).

Pues bien, se tiene lo siguiente:

1) Cuando la función integrando sea positiva y “cóncava hacia arriba” en todo el intervalo de integración (por tener derivada segunda positiva en todo el intervalo), la aproximación dada por la “fórmula de los trapecios” será siempre “por exceso” (la gráfica queda en esos casos un poco por debajo de los lados superiores de los trapecios). Y cuando la función sea negativa ocurre lo contrario (la aproximación será “por defecto”).

2) Y cuando la función integrando sea positiva y “cóncava hacia abajo” en todo el intervalo (por ser su derivada segunda negativa), la aproximación que obtengamos por la “fórmula de los trapecios” será siempre “por defecto” (la gráfica queda en esos casos un poco por encima de los lados superiores de los trapecios). Y cuando la función sea negativa la aproximación será “por exceso”.

3) Pero cuando haya tramos con diferentes concavidades de la función integrando en el intervalo de integración, no podrá predecirse el sentido de la aproximación que se obtenga por este método.

Además, se demuestra que el valor absoluto del error de la aproximación obtenida por esta “fórmula de los trapecios” es menor o igual que $\frac{(b-a) \cdot h^2 \cdot M_2}{12}$, siendo M_2 el máximo del valor absoluto de la derivada segunda en el intervalo $[a, b]$ de integración (o un valor positivo no superado por el valor absoluto de la derivada segunda, que debe buscarse lo más pequeño posible). La expresión anterior se llama “una cota superior del error cometido en valor absoluto”.

MÉTODOS NUMÉRICOS BÁSICOS

Nótese que la exactitud del método de los trapecios, mejora mucho cuando tomamos valores de h suficientemente pequeños, ya que “la cota superior del error cometido” (antes mencionada) varía con el cuadrado de h (conviene entonces tomar siempre $h < 1$, con lo cual será $h^2 < h$). Esto nos indica que debemos trabajar siempre con particiones del intervalo que tengan un número de subintervalos mayor que la longitud del intervalo de integración, pues así $h = (b - a)/n$ será menor que uno.

Ejemplo: Vamos a obtener una aproximación de la misma integral del ejemplo de la pág. 14, utilizando ahora la “fórmula de los trapecios”.

Se trata entonces de calcular una aproximación de la integral $\int_0^1 \frac{dx}{x^2+1}$, usando la misma partición P_{10} de aquel ejemplo. El valor de $h = 0.1$ es bueno, aunque podría ser mejor poniendo más puntos en la partición, a costa de alargar los cálculos (haciéndolos a mano resulta tedioso, pero las máquinas lo hacen en muy poco tiempo). Por tanto, usaremos los mismos valores de la función integrando, los cuales habían sido obtenidos con seis cifras decimales y redondeo.

Y se tiene entonces $E = y_0 + y_{10} = 1.5$ y $R = y_1 + y_2 + \dots + y_9 = 7.099814$.

Por tanto:
$$I \cong \frac{0.1}{2} \cdot (1.5 + 2 \cdot 7.099814) = \boxed{0.7849814}$$

Y si hacemos la media aritmética de los valores obtenidos en la páginas 14 y 15, cuando aplicamos las dos primeras “fórmulas de los rectángulos”, resulta el mismo valor:

$$\frac{0.8099814 + 0.7599814}{2} = \boxed{0.7849814}$$

Ahora, como la función integrando es cóncava hacia abajo en el intervalo $(0, \sqrt{3}/3)$ y es cóncava hacia arriba en el intervalo $(\sqrt{3}/3, 1)$, no podemos asegurar que la aproximación obtenida sea por defecto o sea por exceso. En efecto, se tiene $f''(x) = (6x^2 - 2)/(x^2 + 1)^3$ con lo cual el punto $\sqrt{3}/3 = 0.577\dots$ es punto de inflexión de $f(x)$, resultando $f''(x)$ negativa desde 0 hasta dicho punto y $f''(x)$ positiva de ahí en adelante.

Sin embargo, como sabemos el valor exacto de la integral en este caso, que es $0.78539816\dots$ (dato que en muchas otras ocasiones no se tiene), vemos que la aproximación es por defecto, con las dos primeras cifras decimales exactas.

En cuanto a la acotación del error cometido por la “fórmula de los trapecios” en el ejemplo anterior, podemos tomar $\overline{M_2} = 4$, pues $|f''(x)| = \frac{|6x^2 - 2|}{(x^2 + 1)^3} \leq |6x^2 - 2|$ (ya que el denominador de la derivada segunda es siempre mayor o igual que 1). Y en efecto, $|6x^2 - 2| = 2 \cdot |3x^2 - 1|$ y si representamos la gráfica de la función $y = |3x^2 - 1|$ vemos claramente que su máximo absoluto en el intervalo $[0, 1]$ es el valor alcanzado en $x = 1$ que es 2 (ver Secciones 2.1 y 3.5). En conclusión, podemos decir que $|f''(x)| \leq 2 \cdot 2 = 4$ en el intervalo $[0, 1]$.

Por tanto, tenemos: $|error| \leq (1 - 0) \cdot 0.1^2 \cdot 4/12 = 0.01/3 = 0.0033333\dots$

Comprobación: El valor absoluto del verdadero error cometido es la diferencia (en valor absoluto) entre el valor exacto de la integral ($\pi/4 \cong 0.78539816\dots$) y el valor aproximado calculado por la “fórmula de los trapecios” que fue 0.7849814 . O sea, $|error| = 0.00041676\dots$ que es efectivamente menor que la cota obtenida $0.0033333\dots$ la cual nos asegura la posible exactitud de las dos primeras cifras decimales de la aproximación (si tomamos las tres primeras cifras decimales con redondeo de la aproximación, resulta $c \cong 0.785$, que coincide con el valor exacto redondeado también en la tercera cifra decimal, lo cual es un resultado muy bueno si tomamos en cuenta que hemos trabajado con $h = 0.1$ que es relativamente grande).

En la práctica normalmente no se dispone del valor exacto de la integral (justamente por ese motivo se estará aplicando un método aproximado), con lo cual no podremos saber el valor del verdadero error cometido al usarlo, como sí hemos podido en este caso. Entonces, lo único que podremos hacer muchas veces al respecto es obtener una cota superior del valor absoluto de dicho error, como hicimos en el ejemplo anterior aplicando la fórmula dada, lo cual nos indicará las cifras decimales de la aproximación obtenida que sean seguras.

MÉTODOS NUMÉRICOS BÁSICOS

Así, en el caso del ejemplo anterior obtuvimos $I \cong 0'7849814$, con error en valor absoluto menor o igual a $0'0033333 \dots < 0'0033334$. Lo cual significa que el valor exacto quedará comprendido entre $0'7849814 - 0'0033334 = 0'781648$ y $0'7849814 + 0'0033334 = 0'7883148$. Por tanto, lo que puede asegurarse es que $I \cong 0'78$ (con sus cifras decimales exactas; las demás no sirven).

MÉTODO DE SIMPSON (Thomas Simpson, matemático inglés del siglo XVIII):

Es mejor que los métodos anteriores. **Pero necesita una cierta partición del intervalo de integración, con un número impar de puntos equidistantes**, en cantidad suficiente para que el paso h sea lo más pequeño posible (como mínimo menor que 1). Se utilizan otra vez los valores de la función integrando en los puntos de la partición, como en los métodos anteriores. Obsérvese que si $n + 1$ es impar, n será par, luego **usaremos siempre un número par de subintervalos**.

Este método tiene la particularidad de dar valores exactos de la integral si la función integrando es polinómica de grado no mayor que 3, pero estos casos se resuelven fácilmente por la Regla de Barrow, con lo cual no se necesita ningún método numérico. Sin embargo, para integrales más complicadas, donde el cálculo de una primitiva sea muy difícil o muy largo, y también cuando no exista ninguna primitiva calculable del integrando (porque la misma no sea una “función elemental”) el método es muy útil.

Consiste en aplicar la llamada “fórmula de Simpson”:
$$\int_a^b f(x) dx \cong \frac{h}{3} \cdot (E + 4I + 2P)$$

siendo $E = y_0 + y_n$ (suma de las ordenadas extremas), siendo $I = y_1 + y_3 + \dots + y_{n-1}$ (suma de las ordenadas de índices Impares) y siendo $P = y_2 + y_4 + \dots + y_{n-2}$ (suma de las ordenadas de índices Pares, salvo las extremas).

No damos la demostración detallada de esta fórmula, pero se basa en aproximar el valor de la integral mediante la suma de las áreas orientadas de regiones parecidas a los trapecios que usábamos anteriormente, pero ahora esas regiones limitadas por segmentos rectilíneos verticales y **por arcos de parábolas** (estos arcos de parábola se basan en tres valores consecutivos de la función integrando en tres puntos seguidos de la partición que corresponden a dos subintervalos consecutivos; en efecto, por tres puntos dados pasa una única parábola o pasa una recta si están alineados). A saber:

La primera región corresponde a los tres primeros puntos x_0, x_1, x_2 de la partición, la cual queda limitada verticalmente a la izquierda por la ordenada y_0 , limitada verticalmente a la derecha por la ordenada y_2 , limitada por un lado horizontal sobre el eje OX que corresponde a la unión de los subintervalos $[x_0, x_1]$ y $[x_1, x_2]$ (de longitud $2h$) y limitada superiormente (o inferiormente, si los valores de la función son negativos) por el arco de la única parábola vertical que va del punto (x_0, y_0) al punto (x_2, y_2) pasando por el punto intermedio (x_1, y_1) ; pues bien, se demuestra que el área orientada de esa primera región vale $\frac{h}{3} \cdot (y_0 + 4y_1 + y_2)$.

El área orientada de la segunda región, similar a la anterior y correspondiente a los tres puntos x_2, x_3, x_4 de la partición, vale $\frac{h}{3} \cdot (y_2 + 4y_3 + y_4)$. Etcétera...

Y el área orientada de la última región, correspondiente a los tres últimos puntos x_{n-2}, x_{n-1}, x_n de la partición, vale $\frac{h}{3} \cdot (y_{n-2} + 4y_{n-1} + y_n)$.

Así, al sumar todas estas áreas orientadas, resulta

$$\frac{h}{3} \cdot (y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \dots + 2y_{n-2} + 4y_{n-1} + y_n)$$

Obsérvese que todas las ordenadas de índices impares quedan multiplicadas por 4 y todas las ordenadas de índices pares quedan multiplicadas por 2, salvo las dos extremas y_0 e y_n que quedan solas, al principio y al final del paréntesis. Por tanto, ha quedado la expresión $\frac{h}{3} (E + 4I + 2P)$.

Nota: Si en algún caso estuviesen alineados los tres puntos consecutivos del plano correspondientes a la terna x_0, x_1, x_2 , a la terna x_2, x_3, x_4, \dots o a la última terna x_{n-2}, x_{n-1}, x_n , el arco

MÉTODOS NUMÉRICOS BÁSICOS

de parábola definido por dichos puntos no será tal sino que será un segmento rectilíneo, con lo cual la región plana correspondiente será un trapecio. Pero el cálculo no se altera en absoluto por este motivo.

Además, se demuestra que “una cota superior” del valor absoluto del verdadero error cometido cuando se utiliza este método es $\frac{(b-a) \cdot h^4 \cdot M_4}{180}$, donde M_4 es el máximo absoluto de $|f^{(4)}(x)|$ en el intervalo $[a, b]$. (M_4 se puede sustituir por un valor cualquiera no superado por dicha derivada cuarta en valor absoluto, el cual debe buscarse lo más pequeño posible).

Volvemos a notar la importancia de trabajar con valores pequeños de h , pues la exactitud de esta fórmula depende de la potencia cuarta del paso h de la partición.

Esta acotación nos dice además que si el integrando es un polinomio de grado no mayor que 3, cuya derivada cuarta será entonces la función cero, no se cometerá error al aplicar la “fórmula de Simpson” con cualquier partición, pues se tiene $M_4 = 0$.

Ejemplo: Calculemos la aproximación que da la fórmula de Simpson para la misma integral que hemos considerado en los ejemplos anteriores, utilizando para ello la misma partición P_{10} de 11 puntos equidistantes del intervalo $[0, 1]$ (tiene un número impar de puntos, luego podemos aplicar este método), con $h = 0.1$ (paso menor que 1). La integral es $\int_0^1 \frac{dx}{x^2+1}$, de valor exacto $\pi/4$, y los valores de la función integrando son los mismos calculados cuando aplicamos las dos primeras “fórmulas de los rectángulos” (ver pág 14).

Tenemos entonces: $E = y_0 + y_{10} = 1.5$; $I = y_1 + y_3 + y_5 + y_7 + y_9 = 3.931157$;

$$P = y_2 + y_4 + y_6 + y_8 = 3.168657$$

Y entonces la aproximación dada por la “fórmula de Simpson” será:

$$\int_0^1 \frac{dx}{x^2+1} \cong \frac{0.1}{3} \cdot (1.5 + 4 \cdot 3.931157 + 2 \cdot 3.168657) = \boxed{0.785398}$$

resultado casi igual al valor exacto de la integral $\pi/4 = 0.78539816\dots$

Se nota la mayor exactitud del resultado de esta fórmula respecto a los obtenidos anteriormente, pues en todos los casos se ha usado la misma partición.

Veamos, por último, cuál puede ser “una cota superior del valor absoluto del error cometido” en este caso. Para ello necesitamos un valor para M_4 y esto nos obliga a calcular la derivada cuarta de la función integrando. Se tiene $f^{(4)}(x) = \frac{120x^4 - 240x^2 + 24}{(x^2+1)^5}$, con lo cual

$$|f^{(4)}(x)| \leq |120x^4 - 240x^2 + 24| = 24 \cdot |5x^4 - 10x^2 + 1|$$

porque el denominador $(x^2 + 1)^5$ es siempre mayor o igual que 1.

Ahora bien, $y = 5x^4 - 10x^2 + 1$ tiene derivada $y' = 20x^3 - 20x = 20x \cdot (x^2 - 1) = 20x \cdot (x + 1) \cdot (x - 1)$ y haciendo el cuadro de signos de y' resulta que es negativa en el intervalo $(0, 1)$, lo cual nos indica, junto con la continuidad en $[0, 1]$, que la función de grado 4 (sin valor absoluto) es estrictamente decreciente en el intervalo de integración $[0, 1]$. Así, el máximo absoluto de esta función en dicho intervalo se alcanzará en $x = 0$ (y vale 1) y su mínimo absoluto se alcanzará en $x = 1$ (y vale -4). Entonces pasa necesariamente por el valor cero en un punto interior del intervalo (por el T. de Bolzano), y como consecuencia, la función de grado 4 con valor absoluto, $y = |5x^4 - 10x^2 + 1|$, tendrá valor máximo absoluto 4 y valor mínimo absoluto cero en ese mismo intervalo $[0, 1]$. (Ver Sección 2.1).

Conclusión: $|f^{(4)}(x)| \leq 24 \cdot |5x^4 - 10x^2 + 1| \leq 24 \cdot 4 = 96$

luego podemos tomar $M_4 = 96$ para la acotación del valor absoluto del error.

MÉTODOS NUMÉRICOS BÁSICOS

Así, “una cota superior del error en valor absoluto” correspondiente a la aproximación que nos ha dado la fórmula de Simpson para esta integral es $(1 - 0) \cdot 0.1^4 \cdot 96/180 = 0.00005333\dots$ de lo cual deducimos que dicha aproximación deberá tener como mínimo 3 cifras decimales exactas (y ello es cierto pues hemos visto que realmente tiene 6 cifras decimales exactas).

Nota: En la práctica no siempre es fácil obtener el valor de M_4 por métodos analíticos, como acabamos de hacer. Pero se puede utilizar el método gráfico de representar el valor absoluto de la función derivada cuarta en el intervalo de integración (usando una calculadora gráfica), con lo cual se verá fácilmente el valor de su máximo en ese intervalo. También podemos tomar para M_4 cualquier valor real que no sea superado por el valor absoluto de la derivada cuarta en $[a, b]$, procurando que sea lo más pequeño posible para no agrandar innecesariamente “la cota superior” buscada.

Notas finales importantes:

- 1) Como ya dijimos, en la práctica no aplicaremos los métodos numéricos a integrales de las que se conozcan los valores exactos, como la considerada en los ejemplos que hemos puesto. Por tanto, al calcular un valor aproximado de alguna de ellas mediante uno de los métodos numéricos, al no conocerse el valor exacto de la integral tampoco conoceremos el verdadero error cometido, con lo cual cobra especial importancia el cálculo de “una cota superior” de dicho error en valor absoluto (lo más pequeña posible). Y ello nos permitirá tener una buena idea de la fiabilidad de las cifras obtenidas en el resultado aproximado (deben conservarse las cifras exactas y descartarse las demás).
- 2) Recuérdese que si $|x| < a$, con a positivo, será $-a < x < a$. Por tanto, si $|error| < C$, siendo C “una cota superior” obtenida, se tendrá $-C < error < C$. Y si sumamos a los tres miembros anteriores al valor A de la aproximación obtenida, será

$$A - C < A + error < A + C$$

Pero $A + error$ es el valor exacto I de la integral, luego se tendrá

$$\boxed{A - C < I < A + C}$$

Así vimos en la pág. 18 que al aplicar el método de los trapecios a la integral que tenemos, se obtenía el valor aproximado $A = 0.7849814$, siendo “una cota superior” del error cometido en valor absoluto $C = 0.0033333\dots$. Por lo cual, dijimos en la pág. 19 que el valor exacto I de la integral quedaba entre

$$0.7849814 - 0.0033334 = 0.7816480 \quad \text{y} \quad 0.7849814 + 0.0033334 = 0.7883148$$

de donde concluíamos que las cifras 0.78 eran exactas con seguridad (y el resultado aproximado de la integral por el método de los trapecios debía darse entonces así, con dos cifras decimales solamente, eliminando las demás cifras decimales obtenidas).

Y ahora, cuando aplicamos el método de Simpson a la misma integral, obtuvimos el valor aproximado $A = 0.785398$, siendo $C = 0.0000533\dots$ “una cota superior” del error cometido en valor absoluto. Así el valor exacto I de la integral estará entre

$$0.785398 - 0.000054 = 0.785344 \quad \text{y} \quad 0.785398 + 0.000054 = 0.785452$$

siendo ahora las cifras 0.785 exactas con seguridad (por tanto, el resultado aproximado de la integral por el método de Simpson deberá darse así, con tres cifras decimales solamente).

- 3) Si hubiésemos tomado particiones con muchos más puntos, los valores aproximados obtenidos en todos los casos habrían sido mejores, resultando cotas de error más pequeñas y obteniendo más cifras exactas en las aproximaciones. Téngase en cuenta que los resultados anteriores fueron obtenidos con una partición de solamente 11 puntos, lo cual nos dio

MÉTODOS NUMÉRICOS BÁSICOS

$h = 0'1$ que es válido pero no muy pequeño. Mejor sería haber aplicado los métodos con $h = 0'01$ por ejemplo, lo cual nos habría obligado a calcular los valores de la función integrando en los 101 puntos de la correspondiente partición del intervalo $[0, 1]$, que es $P_{100} = \{0, 0'01, 0'02, 0'03, \dots, 1\}$. Esto a mano es muy tedioso, pero para eso existen los ordenadores y las calculadoras programables.

- 4) Finalmente, cuando una calculadora tenga la opción de hallar valores de integrales definidas, seguro que utilizará para ello algún método numérico (elegido por el fabricante con su equipo técnico), pero si utilizan alguno de los métodos antes descritos lo harán con particiones de muchos puntos, de modo que los valores de h resultantes sean lo suficientemente pequeños como para poder presentar en la pantalla solamente cifras exactas del resultado aproximado obtenido, utilizando para ello todo el espacio disponible (y las pantallas disponen normalmente de espacio para escribir 8 o más cifras en total, incluyendo decimales y no decimales).
 - 5) En Cálculo Numérico se distinguen “los métodos directos” (donde se llega directamente a una solución exacta o bien permiten hallar en un solo paso del procedimiento una aproximación suficientemente buena) y “los métodos de aproximaciones sucesivas” (donde se repite la aplicación del método un cierto número de veces, obteniéndose cada vez una aproximación mejor). Como ejemplos de los primeros hemos visto los dos métodos de interpolación dados (Lagrange y Newton-Gregory, donde el polinomio obtenido es exacto pero su valor en el punto intermedio que interese es aproximado para la función desconocida) así como los tres métodos de integración aproximada (rectángulos, trapecios y Simpson, en todos los cuales obtenemos directamente una aproximación de la integral definida). Y como ejemplos de los segundos hemos visto los dos métodos de resolución de ecuaciones (bisección y Newton-Raphson, donde en el primero se van obteniendo en cada paso tres aproximaciones cada vez mejores de la solución exacta, y donde en el segundo método obtenemos sucesivas aproximaciones cada vez mejores por aplicación reiterada de la “fórmula recurrente”).
-