INTRODUCCIÓN. ERRORES
INTERPOLACIÓN Y APROXIMACIÓN DE FUNCIONES
DERIVACIÓN E INTEGRACIÓN NUMÉRICAS

Análisis Numérico

Volumen I, v0.5

Las Palmas, Diciembre 2000

Σ

AUTORES : JESÚS GARCÍA QUESADA

Edificio de Informática y Matemáticas
Campus Universitario de Tafira
35017 Las Palmas de Gran Canaria

ISBN-84-699-4380-4

Nº Registro: 762701

Índice General

Índice de Figuras

Capítulo 1

INTRODUCCIÓN

1.1 EVALUACIÓN DE POLINOMIOS Y SUS DERIVADAS

Sea el polinomio

$$p(x) = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + \cdots + a_n x^n$$

del cual queremos determinar su valor en un punto x_0, o sea, calcular el valor de $p(x_0)$.
El cálculo término a término en orden creciente de n puede resultar laborioso para valores elevados de n : si cada factor x^k se calcula haciendo $k-1$ multiplicaciones repetidas de x, el cálculo de $p(x_0)$ requiere $n(n+1)/2$ multiplicaciones y n sumas.
Por otro lado, si cada factor x^k se calcula por multiplicaciones sucesivas de $x.x^{k-1}$ entonces se necesitan $2n-1$ multiplicaciones y n sumas.
Sin embargo, si se usa la forma encajada de $p(x)$:

$$p(x) = ((\cdots(a_n x + a_{n-1})x + \cdots + a_1)x + a_0$$

sólo se necesitan n multiplicaciones y n sumas.
Por ejemplo, para $n = 4$:

$$p(x) = (((a_4 x + a_3)x + a_2)x + a_1)x + a_0$$

Este procedimiento fué descrito por W.G. Horner en el siglo XIX[1] aunque Newton había hecho uso de la misma idea 150 años antes (1669).
Ejemplo : Sea $p(x) = 3x^3 - 2x^2 + x - 2,\ x_0 = 2$

[1] Philosophical Transactions, Royal Society of London 109 (1819), 308–335

$$\begin{array}{c|cccc} & 3 & -2 & 1 & -2 \\ \hline 2 & 3 & 4 & 9 & \boxed{16} \end{array}$$

y por tanto es $p(2) = 16$. El algoritmo queda :

Algoritmo 1 Ruffini-Horner(a,x)

1. $i \longleftarrow n, \;\; v \longleftarrow 0$
2. $v \longleftarrow v * x + a_i, \;\; i \longleftarrow i - 1$
3. **si** $i \geq 0$ **entonces**
4. ir a paso 2
5. **fin si**
6. **escribir** v

y si llamamos t_1, t_2, t_{3-5} y t_6 a los tiempos que se requieren para la ejecución de los pasos anteriores, el tiempo total T necesario para la ejecución del algoritmo de Horner es :

$$T = t_1 + t_6 + (t_2 + t_{3-5})(n+1) = (t_1 + t_2 + t_{3-5} + t_6) + (t_2 + t_{3-5})n = m_1 + m_2 n = \mathcal{O}(n)$$

y por tanto utiliza "tiempo lineal"[PPB85]
Por otra parte, Ostrowski[2] ha demostrado que para $n \leqslant 4$ éste esquema es el que precisa de un número mínimo de operaciones. Para polinomios de grado superior, existen esquemas[Knu81] que requieren menos de $2n$ operaciones, sobre todo cuando $p(x)$ *se tiene que evaluar muchas veces.*
La secuencia de los valores calculados durante la evaluación de $p(x)$ para un determinado argumento $x = x_0$ está intimamente relacionada con la de eliminar el factor $(x - x_0)$ de $p(x)$ por medio de la **división sintética**.
Utilizando la notación anterior, sea :

$$\begin{cases} b_n = a_n, \\ b_i = b_{i+1}x_0 + a_i, \quad i = n-1, n-2, \cdots, 1, 0 \end{cases}$$

entonces la secuencia de los valores b_i es tal que se corresponde con los coeficientes del polinomio de grado $n - 1$, cociente de dividir $p(x)$ por $(x - x_0)$, y además $b_0 = p(x_0)$:

$$\frac{p(x)}{(x - x_0)} = q_{n-1}(x) + \frac{R_0}{(x - x_0)} \Longrightarrow p(x) = (x - x_0)q_{n-1}(x) + R_0$$

[2] Studies in Mathematics and Mechanics Presented to R. Von Mises, pp. 40–48, Academic Press, New York, 1954

siendo R_0 constante y realizando la división de forma directa o bien identificando los términos de igual exponente en x en ambos lados de la ecuación anterior se puede ver que:

$$q_{n-1}(x) = b_n x^{n-1} + b_{n-1}x^{n-2} + \cdots + b_2 x + b_1$$

y $R_0 = b_0 = p(x_0)$. Si ahora dividimos $q_{n-1}(x)$ por el factor $(x - x_0)$ de igual forma se obtiene un polinomio $q_{n-2}(x)$ de grado $n-2$ con un resto R_1 tal que :

$$q_{n-1}(x) = (x - x_0)q_{n-2}(x) + R_1$$

y escribiendo $q_{n-2}(x)$ de la forma:

$$q_{n-2}(x) = c_n x^{n-2} + c_{n-1}x^{n-3} + \cdots + c_3 x + c_2$$

por analogía podemos escribir :

$$\begin{cases} c_n = b_n, \\ c_i = c_{i+1}x_0 + b_i, \quad i = n-1, n-2, \cdots, 1 \end{cases}$$

donde $c_1 = R_1 = q_{n-1}(x_0)$. El polinomio original se convierte en

$$p(x) = p_n(x) = (x - x_0)^2 q_{n-2}(x) + (x - x_0)R_1 + R_0$$

y su derivada primera es :

$$p'(x) = 2(x - x_0)q_{n-2}(x) + (x - x_0)^2 q'_{n-2}(x) + R_1$$

de forma que $p'(x_0) = R_1 = c_1$.
Siguiendo con el proceso para $q_{n-2}(x)$, su división por $(x - x_0)$ dará origen a un nuevo polinomio $q_{n-3}(x)$ de resto R_2 y asi sucesivamente, de tal forma que el polinomio original puede escribirse en la forma :

$$p(x) = p_n(x) = (x - x_0)^n R_n + (x - x_0)^{n-1}R_{n-1} + \cdots + (x - x_0)R_1 + R_0$$

y por lo tanto:
$p(x_0) = R_0, \quad p'(x_0) = R_1, \quad p''(x_0) = 2R_2$, etc. y en general

$$R_k = \frac{p^{(k)}(x_0)}{k!}, \quad k = 0, 1, \ldots, n.$$

Ejemplo : Para calcular el valor del polinomio $p(x) = 3x^4 + 2x^3 - x^2 + 2x - 5$ y sus derivadas en $x_0 = 2$ construimos el esquema :

	3	2	-1	2	-5
2	3	8	15	32	**59**
	3	14	43	**118**	
	3	20	**83**		
	3	**26**			
	3				

Por tanto, como $\forall p(x) \in \mathbb{R}[x]$ es $p(x) \in \mathcal{C}^\infty(\mathbb{R})$, esto nos da una posibilidad de encontrar rápidamente el desarrollo de Taylor en potencias de $(x-x_0)$ para cualquier polinomio $p(x)$: Si es

$$p(x) = \sum_{k=0}^{\infty} \frac{p^{(k)}(x_0)}{k!}(x-x_0)^k$$

teniendo en cuenta que $p^{(k)}(x_0) = 0, \forall k > n \ \wedge \ \forall x_0$ y que $R_k = p^{(k)}(x_0)/k!$ tendríamos :

$$p(x) = \sum_{k=0}^{n} R_k(x-x_0)^k$$

donde los R_k vendrían dados por el algoritmo de Ruffini-Horner.
Ejemplo : Para el polinomio anterior tenemos :

$$\begin{aligned} p(x) &= 3x^4 + 2x^3 - x^2 + 2x - 5 = (\text{en potencias de } x-2) = \\ &= 3(x-2)^4 + 26(x-2)^3 + 83(x-2)^2 + 118(x-2) + 59 = \\ &= 3y^4 + 26y^3 + 83y^2 + 118y + 59 = p(y), \end{aligned} \tag{1.1}$$

haciendo el cambio de variable $y = x - 2$.
Análogamente, a partir de la última expresión de $p(x) = 3(x-2)^4 + 26(x-2)^3 + \cdots + 59$ podemos obtener también por Ruffini-Horner su desarrollo en potencias de $x = (x-0)$:

-2	3	26	83	118	59
	3	20	43	32	$-\mathbf{5}$
	3	14	15	**2**	
	3	8	$-\mathbf{1}$		
	3	**2**			
	3				

Por tanto, el algoritmo de Ruffini–Horner definitivo para la determinación del valor de un polinomio y sus derivadas en un punto x_0 queda de la siguiente forma :

Algoritmo 2 ALGORITMO DE RUFFINI–HORNER

para todo j tal que $0 \leq j \leq n$ **hacer**
  $b_j \longleftarrow a_j$
fin para
para $k = 0$ **hasta** $n - 1$ **hacer**
  para $j = n - 1$ **hasta** k **hacer**
    $b_j \longleftarrow b_j + x_0 * b_{j+1}$
  fin para
fin para

siendo los b_j los R_j anteriores, para $0 \leqslant j \leqslant n$.
Por tanto, el método de Horner requiere $(n^2 + n)/2$ multiplicaciones y $(n^2 + n)/2$ sumas.

1.2 ERRORES Y APROXIMACIONES

Veamos ahora algunos conceptos sencillos que nos ayudarán a desarrollar el temario de éste curso, que aunque son conocidos por todos, están carentes de una apropiada formulación.

1.2.1 Cifras significativas

Un *dígito significativo* de un número aproximado es cualquier dígito no nulo, en su representación decimal, o cualquier cero que no sea utilizado como indicador posicional del punto decimal.
Las cifras o dígitos significativos son aquellos que se pueden utilizar "con confianza".
Nos queda por definir el concepto de *dígito exacto*. Se dice que los n primeros dígitos decimales de un número aproximado son exactos si la diferencia, en valor absoluto, entre el número aproximado y exacto no excede de media unidad del dígito decimal situado en n-ésimo lugar, contando de izquierda a derecha.
Si la magnitud del error en $\tilde{x}$ no excede de $\frac{1}{2}10^{-t}$ entonces $\tilde{x}$ se dice que tiene **t decimales correctos**.
Los dígitos de $\tilde{x}$ que ocupan posiciones donde la unidad es mayor o igual que 10^{-t} se llaman *dígitos significativos* (no se cuentan ceros iniciales).

Ejemplo 1. 0.001234 ± 0.000004 *tiene cinco decimales correctos y tres dígitos significativos mientras que* 0.001234 ± 0.0000006 *tiene 4 decimales correctos y 2 dígitos significativos.*

Por lo tanto, si en lugar del número exacto A se toma como aproximación el número a se sabe que

$$\Delta = |A - a| \leqslant \frac{1}{2}10^{-n}$$

entonces, por definición, los n primeros dígitos decimales de dicho número son exactos.
Por ejemplo, con respecto al número exacto $A = 35.97$, el número $a = 36.00$ es una aproximación con un dígito decimal exacto, ya que $|A - a| = 0.03$ y $0.03 < 0.5 \times 10^1$
Si estamos trabajando con números aproximados, a veces es necesario redondear éste número, es decir, reeemplazarlo por un número que tenga $\mathbf{a_1}$ que tenga un menor número de dígitos significativos. Este nuevo número $\mathbf{a_1}$ tiene que ser elegido de tal forma que el error de redondeo $|\mathbf{a_1} - \mathbf{a}|$ sea mínimo.
Para efectuar el redondeo de un número a n dígitos significativos, han de eliminarse todos los dígitos a la derecha del dígito significativo de lugar n, o se reemplazan por ceros. Al realizar ésta operación, han de tenerse en cuenta :

1. Si el primer dígito despreciado es menor que 5, los dígitos restantes se dejan como están.

2. Si el primer dígito despreciado excede de 5, se añade una unidad al último dígito de los que quedan.

3. Si el primer dígito despreciado es exactamente 5 y hay dígitos no ceros entre los despreciados se añade una unidad al último dígito conservado.

   Sin embargo, si el primer dígito despreciado es exactamente 5, y el resto de los dígitos despreciados son ceros, el último dígito conservado se deja como está si es par, y se incrementa una unidad si es impar.

Ejemplo 2. *Se pide redondear el número* $\pi = 3.1415926535$ *a cinco, cuatro y tres cifras significativas.*

Para cinco cifras significativas y exactas :

$$3.14159;\ |\pi - 3.14159| \simeq 2.6 \times 10^{-6} < 0.5 \times 10^{-5}$$

Para cuatro cifras significativas y exactas : $3.1416;\ |\pi - 3.1416| \simeq 7.3 \times 10^{-6} < 0.5 \times 10^{-4}$
Para tres cifras significativas y exactas : $3.142;\ |\pi - 3.142| \simeq 4 \times 10^{-4} < 0.5 \times 10^{-3}$

1.2.2 Exactitud y precisión

Antes de entrar con las definiciones que existen del error, vamos a definir dos parámetros que nos permitirán caracterizarlo :

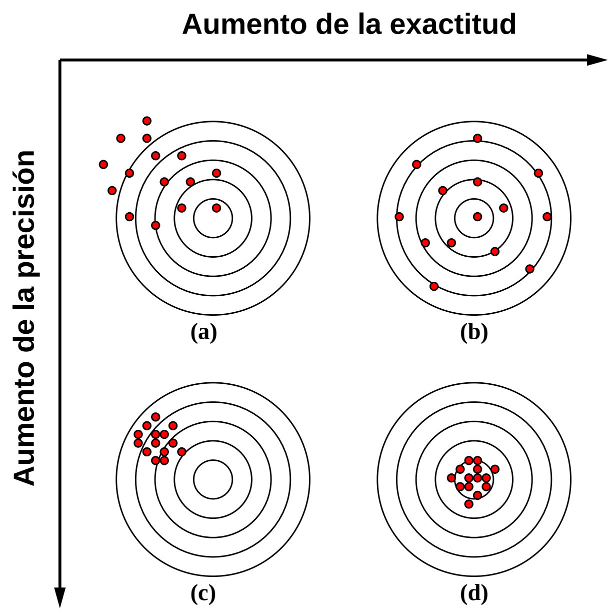

Figura 1.1: Exactitud y precisión

Exactitud Nos indica lo cerca que está un valor, ya sea calculado o medido, del verdadero valor.

Precisión Nos indica lo cerca que están unos valores de otros, ya sean calculados o medidos a través de la aplicación repetida de un método. De ésta forma la precisión nos indica lo separados que resultan los datos a la hora de hacer diferentes cálculos por el mismo método.

Cuando diseñemos un método numérico nos debemos asegurar que éste sea bastante preciso, ya que un método que aplicado repetidas veces vaya dando resultados alejados entre sí no es fiable. De aquí en adelante, en las consideraciones que hagamos sobre el error irán incluidos ambos aspectos, ya que nuestro objetivo será describir estos de una manera cuantitativa.

1.3 ERRORES. SU DEFINICIÓN

Los errores numéricos, como hemos visto en apartados anteriores, aparecen al usar aproximaciones para representar cantidades y operaciones matemáticas exactas. En las siguientes medidas cuantitativas irán incluidos tanto los errores de redondeo como los errores de truncamiento, que será definido más tarde. Para cualquier tipo de error que puede aparecer en una cifra aproximada se puede definir la siguiente relación :

$$\textbf{Valor verdadero} = \textbf{Aproximación} + \textbf{Error}$$

Teniendo en cuenta la expresión anterior, podemos llegar a la conclusión de que el error numérico es la diferencia entre el verdadero valor y la aproximación :

$$\mathbf{E_v} = \textbf{Valor verdadero} - \textbf{Aproximación}$$

En éste caso, el subíndice **v** quiere indicar que éste error es el verdadero.
El problema que plantea la definición anterior de error es que no tiene en cuenta la magnitud de las cantidades involucradas. Este hecho a veces resulta importante, ya que no es lo mismo cometer un error de un centímetro al medir un puente que a la hora de medir un tornillo. Para evitar esto, se normaliza el error cometido al verdader valor, obteniéndose el **error relativo** :

$$\mathbf{E_r} = \frac{\mathbf{E_v}}{\textbf{Valor verdadero}}$$

El error relativo se puede multiplicar por 100 y se obtiene entonces el **error relativo porcentual** :

$$\mathbf{e_r} = \frac{\mathbf{E_v}}{\textbf{Valor verdadero}} \times \mathbf{100} = \mathbf{E_r} \times \mathbf{100}$$

Ejemplo 3. *Supongamos que medimos un puente y un tornillo y resultan con 9999 cm y 9 cm respectivamente. Si los verdaderos valores son 10000 cm y 10 cm :*

tenemos que $E_a = |10000 - 9999| = 1$ cm $= |10 - 9|$ como error absoluto en los casos y sin embargo $E_r = \frac{1}{10000} \times 100\% = 0.01\%$ en el primer caso y $E_r = \frac{1}{10} \times 100\% = 10\%$ en el segundo de los casos.
En las definiciones anteriores el subíndice **v** indicaba que el error era calculado a partir del verdadero valor del número exacto. Sin embargo, raramente esa cantidad estará disponible. En los métodos numéricos, el verdadero valor sólo lo conoceremos cuando las funciones con las que tratamos puedan ser resueltas analíticamente. Esto ocurrirá en general cuando los sistemas sobre los cuales estamos trabajando sean simples, pero tratando con sistemas que nos podemos encontrar en el mundo real, el error no se podrá conocer con exactitud, ya que no se dispone del verdadero valor. En estos casos, una alternativa es normalizar el error usando la mejor aproximación posible al verdadero valor, que es la aproximación en sí misma :

$$\epsilon_{\mathbf{a}} = \frac{\textbf{Error aproximado}}{\textbf{Aproximación}} \times \mathbf{100}$$

donde ahora el subíndice **a** indica que el error es normalizado con un valor aproximado. El problema que nos planteamos ahora es cómo calcular el error, si no conocemos el valor del objeto de estudio. Este problema es uno de los desafíos con los que nos vamos a encontrar en el análisis numérico, ya que una buena estimación del error nos puede ayudar a decidir cómo de bueno resulta un método respecto a otros. Por ejemplo, algunos métodos numéricos usan un esquema iterativo para calcular una solución (en éste tipo de esquemas, la aproximación actual es calculada a partir de una que se ha calculado con anterioridad); éste proceso de cálculo sucesivo de aproximaciones se sigue para obtener aproximaciones cada vez mejores. En estos casos el error es estimado como la diferencia entre dos aproximaciones sucesivas, con lo que el error relativo porcentual se puede expresar como :

$$\epsilon_{\mathbf{a}} = \frac{\textbf{aproximación actual} - \textbf{aproximación anterior}}{\textbf{aproximación actual}} \times \mathbf{100}$$

Lo único que nos queda por discutir acerca del error es su signo. Lo que nos va a interesar en la mayoría de los casos no va a ser su signo, sino su magnitud, es decir su valor absoluto, ya que generalmente ésta medida se utilizará para compararla con una tolerancia que pondremos como límite de precisión, y que nos indicará qué es lo que vamos a considerar una buena aproximación.
Los siguientes teoremas relacionan el error relativo con el número de cifras significativas y han sido de amplio uso. Para las pruebas correspondientes véase [Sca66].

Teorema 1.3.1. *(Scarborough)*
Si el error relativo de cualquier número no es mayor que $\frac{1}{2\times 10^n} \Longrightarrow$ *el número es correcto a* n *cifras significativas*

Teorema 1.3.2. *(Scarborough)*
Si la tolerancia $t = (0.5 \times 10^{2-n})\%$ *y* $|\epsilon_a| < t \Longrightarrow$ *el resultado es correcto al menos a* n *cifras significativas*

Ejemplo 4. *Sabemos que* $e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$*. Queremos calcular* $e^{0.5}$*, con valor verdadero 1.648721271. Añadir términos hasta que el* ϵ_a *esté por debajo de una tolerancia* ϵ_s *establecida a priori que nos asegure tres cifras significativas.*

Para que sea correcto al menos a tres cifras significativas ha de ser $t = (0.5 \times 10^{2-3})\% = 0.05\%$

No. términos	Aproximación	$\epsilon_v = (val.verd. - aprox)/val.verd.\%$	$\epsilon_a\%$
1	1	39.3	
2	1.5	9.02	33.3
3	1.625	1.44	7.69
4	1.645833333	0.175	1.27
5	1.648437500	0.0172	0.158
6	1.648697917	0.0142	0.0158

Después de que se consideren seis términos, el error relativo aproximado porcentual es menor que 0.05%, y el cálculo termina. Nótese, sin embargo, que el resultado es exacto a cinco cifras significativas, en lugar de tres, que era lo solicitado. Esto ocurre debido a que las fórmulas que estiman el error cometido son conservadoras, lo que significa que aseguran que el resultado obtenido sea al menos tan bueno como lo especificado.

1.4 ERRORES DE REDONDEO

La principal causa de los errores de redondeo es debido al hecho de que los ordenadores retienen un número fijo de cifras significativas.
Para apreciar bien éste efecto entraremos en la representación de los números reales en el ordenador. Las cantidades fraccionales en el ordenador se presentan en el formato de coma flotante, en el que el número es representado por una parte racional, denominada mantisa y por otra entera, que es el exponente, como :

$$m * b^e$$

donde b es la base del sistema de numeración utilizado. Generalmente esto se suele almacenar en varias palabras de ordenador, donde se utilizará un bit para expresar el signo de la mantisa, los siguientes se utilizan para expresar el exponente, y el resto para la mantisa. Consideremos el siguiente ejemplo : supongamos una máquina hipotética que tiene un conjunto de números en coma flotante almacenados en una palabra de siete bits. Se empleará el primer bit como bit de signo, los siguientes tres bits son el signo del exponente (1 bit) y el propio exponente (2 bits), y los restantes corresponderán a la mantisa, ver figura 1.4. *Para evitar ambigüedades en la representación la mantisa es expresada en su forma normalizada, es decir, que el primer dígito que se encuentre a la derecha del punto decimal sea distinto de cero.*
El número más pequeño que podemos formar en ésta máquina imaginaria es :

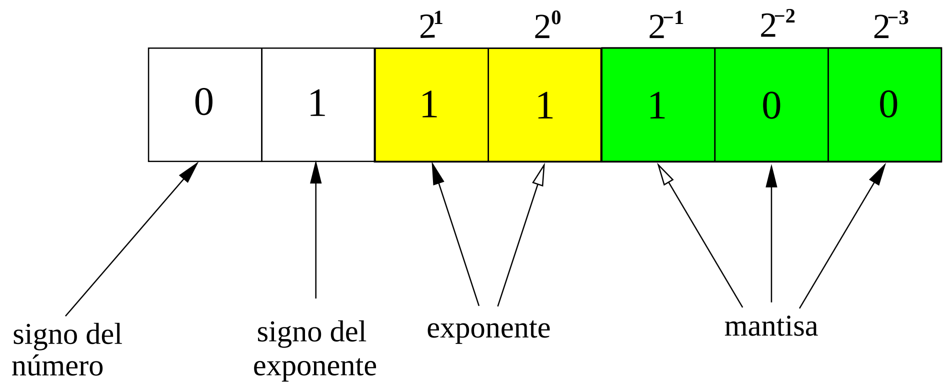

Figura 1.2: Representación en coma flotante

$$0111\,100 = 1 \times 2^{-1} * 2^{-3} = 0.0625$$

Evidentemente pueden existir otras mantisas más pequeñas : $000, 001, \ldots$ pero la mantisa 100 es la menor posible si tenemos en cuenta la normalización.
Los siguientes números que le siguen en magnitud son:

$$\begin{aligned} 0111\,101 &= (1 \times 2^{-1} + 1 \times 2^{-3}) * 2^{-3} = 0.078125 \\ 0111\,110 &= (1 \times 2^{-1} + 1 \times 2^{-2}) * 2^{-3} = 0.093750 \\ 0111\,111 &= (1 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3}) * 2^{-3} = 0.109375 \end{aligned} \tag{1.2}$$

Se puede observar que los números correspondientes en base 10 están espaciados en una cantidad 0.015625.
El siguiente número en la escala se consigue decrementando en 1 el exponente:

$$0110\,100 = (1 \times 2^{-1}) * 2^{-2} = 0.125$$

donde la distancia con el anterior permanece igual a la anterior, pero si incrementamos nuevamente la mantisa el paso entre números consecutivos es ahora de 0.03125, ya que cada incremento en base 2 pasa a ser un incremento de $2^{-3} * 2^{-2}$ (incremento entre números consecutivos multiplicado por el exponente actual).
Siguiendo éste esquema podemos llegar al número más grande que podemos representar, que es :

$$0011\,111 = (1 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3}) * 2^{3} = 7$$

Podemos entonces extraer las siguientes conclusiones :

1. **Sólo existe un rango limitado de cantidades que pueden ser representadas.** Con ésta representación en coma flotante, no sólo tenemos limitado el conjunto de números en su límite superior, sino que además tenemos un intervalo entre el cero y la primera cifra representable que no puede ser usada en ésta representación.

2. **Dentro de un determinado rango de representación, sólo existe un número limitado de cantidades que pueden ser representadas**. Este hecho nos indica que el grado de exactitud está limitado. Obviamente, los números irracionales no pueden ser representados en éste formato, y además, los racionales sólo lo pueden ser hasta un determinado número de dígitos. Los errores cometidos por aproximar estos dos casos a la representación en un ordenador se denominan errores por cuantización, y se pueden realizar de dos maneras:

   i) Representando sólo aquellos dígitos permitidos por la representación, *cortando* aquellos que no pueden ser representados (como veremos más adelante, esto puede ser considerado como un tipo de truncamiento). Si consideramos que el tamaño del intervalo m entre los números en ésta representación es de Δx, con éste tipo de representación siempre asociaremos al número exacto aquel número que ocupa el extremo inferior del intervalo, cometiendo de ésta manera un error máximo de Δx.

   ii) O bien, además de truncar, podemos producir un redondeo del número para asociarle aquel número, que dentro de la representación, sea más próximo al número exacto. Si tenemos en cuenta éste tipo de aproximación, el error máximo que podemos cometer es de $\Delta x/2$.

3. **El intervalo entre números Δx se va haciendo más grande a medida que va aumentando la magnitud de los números**.

   Esta característica permite a la representación en coma flotante mantener el mayor número de cifras significativas a costa de hacer el error de cuantización sea proporcional a la magnitud de los números representados. Este hecho se puede expresar de la siguiente manera :

$$\frac{|\Delta x|}{|x|} \leqslant \begin{cases} \epsilon & \text{en el caso de truncamiento} \\ \epsilon/2 & \text{en el caso de redondeo} \end{cases}$$

Donde ϵ es denominado épsilon–máquina y su valor es $\epsilon = b^{1-t}$, siendo b la base del sistema de numeración utilizado y t el número de dígitos significativos de la mantisa, y nos indica cúal es el máximo error que podemos cometer al cuantizar.

1.5 ERRORES DE TRUNCAMIENTO

Los errores de truncamiento son aquellos que aparecen cuando se utilizan aproximaciones en lugar de procedimientos matemáticos exactos.
Para profundizar más en éste tipo de error y conocer las expresiones que lo definen, entraremos en el estudio del método más utilizado de aproximación funcional, el desarrollo en serie de Taylor. El teorema de Taylor afirma que si una función f y sus primeras $n+1$ derivadas son continuas en un intervalo (x_i, x_{i+1}), el valor de la función en un punto x perteneciente a dicho intervalo es :

$$\boxed{f(x) = f(x_i) + f'(x_i)(x - x_i) + \frac{f''(x_i)}{2!}(x - x_i)^2 + \cdots + \frac{f^n(x - x_i)}{n!}(x - x_i)^n + R_n}$$

donde R_n es el resto, definido como :

$$R_n = \frac{1}{n!}\int_{x_i}^{x}(x-t)^n f^{n+1}(t)dt \qquad \text{o bien :} \qquad R_n = \frac{f^{n+1}(\xi)}{(n+1)!}(x - x_i)^{n+1}$$

En general, una serie de Taylor de orden n será suficiente para aproximar de manera exacta un polinomio de orden n, ya que para éste tipo de funciones las derivadas por encima del orden n son 0, y por tanto el término del error de truncamiento se anula. Sin embargo, existen otras funciones continuas y diferenciables para las cuales el desarrollo en serie por un número finito de términos no nos dan aproximaciones exactas, y por lo tanto, cada vez que se añada un nuevo término a la serie conseguiremos una aproximación más exacta. Aunque éste problema es bastante común,es verdad que para la mayoría de las funciones que aparecen, es suficiente considerar una buena aproximación donde aparezcan pocos términos. El problema radicará en saber escoger el número apropiado de términos para que la aproximación no sobrepase una determinada cota de error. Evidentemente éste problema se podría solucionar utilizando la fórmula de error que anteriormente describimos. Pero aparecen algunas cuestiones que hay que considerar :

1. No se sabe con exactitud cúal es el valor exacto de ξ

2. Se necesita calcular la derivada de orden $n+1$ de la función $f(x)$ y algunas veces ni siquiera se conoce la propia $f(x)$.

A pesar de los inconvenientes es útil de alguna manera una expresión que nos facilite una estimación del error que estamos cometiendo. Si denominamos h al tamaño del intervalo sobre el que estamos realizando la aproximación, el término del error puede ser expresado de la siguiente forma :

$$R_n = \frac{f^{n+1}(\xi)}{(n+1)!}h^{n+1}$$

es decir, proporcional al tamaño del intervalo y se representa como $O(h^{n+1})$. Esta notación es de gran utilidad, ya que nos permitirá comparar dos métodos basados en la aproximación de Taylor. Por ejemplo, un método que presente una estimación del error de $\mathcal{O}(h^2)$ será mejor que uno que presente un error de $\mathcal{O}(h)$, ya que no solo indicamos que para un mismo tamaño de intervalo el primer método es más exacto,sino que además, si reducimos el tamaño del intervalo para hacer el error más pequeño, éste se reducirá en mayor cantidad en el primer método. Por ejemplo, si reducimos el tamaño del intervalo en la mitad, el error en el primer método se reducirá en 1/4, mientras que en el segundo lo hará en 1/2.

1.6 PROPAGACIÓN DEL ERROR

Hasta ahora lo que hemos visto es el error que cometemos al realizar aproximaciones, es decir, errores que se cometen al intentar representar cantidades y cálculos matemáticos exactos por aproximaciones. Ahora lo que intentaremos analizar es el error que se comete cuando se intenta evaluar una función donde las variables que se utilizan para su evaluación están sujetas a error. Para ello supondremos que queremos evaluar la siguiente expresión:

$$N = f(u_1, u_2, \dots, u_n) \tag{1.3}$$

donde las variables independientes $u_1, u_2, \dots, u_n$ están sujetas a unos ciertos errores $\Delta u_1, \Delta u_2, \dots, \Delta u_n$ los cuales afectan al valor de N de la siguiente forma:

$$N + \Delta N = f(u_1 + \Delta u_1, u_2 + \Delta u_2, \dots, u_n + \Delta u_n) \tag{1.4}$$

Para encontrar ΔN desarrollaremos en serie de Taylor el segundo miembro de la expresión:

$$\begin{aligned} N + \Delta N = f(u_1 + \Delta u_1, u_2 + \Delta u_2, \ldots, u_n + \Delta u_n) = f(u_1, u_2, \ldots, u_n) + \\ + \Delta u_1 \frac{\partial f}{\partial u_1} + \Delta u_2 \frac{\partial f}{\partial u_2} + \cdots + \Delta u_n \frac{\partial f}{\partial u_n} + \\ + \frac{1}{2!}\left[(\Delta u_1)^2 \frac{\partial^2 f}{\partial u_1^2} + \cdots + (\Delta u_n)^2 \frac{\partial^2 f}{\partial u_n^2} + 2\Delta u_1 \Delta u_2 \frac{\partial^2 f}{\partial u_1 \partial u_2} + \cdots \right] + \cdots \end{aligned} \tag{1.5}$$

Si consideramos que $\Delta u_1, \Delta u_2, \ldots, \Delta u_n$ son pequeños, podemos despreciar los términos de segundo orden y superiores. Si ahora restamos la expresión (1.3) de la (1.5) obtenemos :

$$\boxed{\Delta N = \Delta u_1 \frac{\partial f}{\partial u_1} + \Delta u_2 \frac{\partial f}{\partial u_2} + \cdots + \Delta u_n \frac{\partial f}{\partial u_n}}$$

En la expresión anterior, los signos de cada uno de los sumandos pueden hacer que en algún caso el error se cancele. Puesto que estamos intentando encontrar una cota del mismo, nos conviene considerar la expresión anterior en sus errores absolutos :

$$\boxed{E_a = |\Delta N| \leqslant |\Delta u_1| \cdot \left|\frac{\partial f}{\partial u_1}\right| + |\Delta u_2| \cdot \left|\frac{\partial f}{\partial u_2}\right| + \cdots + |\Delta u_n| \cdot \left|\frac{\partial f}{\partial u_n}\right|}$$

Esta expresión del error nos define el error en valor absoluto que se comete cuando en las variables se producen los errores anteriormente indicados. Una medida mejor sería utilizar el error relativo :

$$\boxed{E_r = \left|\frac{\Delta N}{N}\right| \leqslant \left|\frac{\Delta u_1}{N}\right| \cdot \left|\frac{\partial f}{\partial u_1}\right| + \left|\frac{\Delta u_2}{N}\right| \cdot \left|\frac{\partial f}{\partial u_2}\right| + \cdots + \left|\frac{\Delta u_n}{N}\right| \cdot \left|\frac{\partial f}{\partial u_n}\right|}$$

Ejemplo 5. *Para el caso de una variable :* *Dado un valor de $\tilde{x} = 2.5$ con un error de $\Delta\tilde{x} = 0.01$, estimar el error resultante en la función $f(x) = x^3$.*

Es $N(x) = x^3$ y queda:

$$\Delta N(\tilde{x}) = \left(\frac{\partial N}{\partial x}\right)_{\tilde{x}} \Delta\tilde{x} = (3x^2 \Delta x)_{\tilde{x}} = 3 \times (2.5)^2 \times (0.01) = 0.1875$$

y como $f(2.5) = 15.625$, podemos predecir que $N(2.5) = 15.625 \pm 0.1875$, o sea, que el valor está entre 15.4375 y 15.8125.

(En efecto,si $x = 2.49$ la función valdría 15.4382 y si x fuese 2.51 sería 15.8132. Para éste caso, la aproximación del error de primer orden proporciona una estimación bastante próxima del verdadero error)

1.6.1 Aplicación de la fórmula del error a las operaciones fundamentales

ADICIÓN Sea $N = u_1 + u_2 + \cdots + u_n \Longrightarrow \Delta N = E_a = \Delta u_1 + \Delta u_2 + \cdots + \Delta u_n$
Por tanto, *el error absoluto de una suma de números aproximados es la suma algebraica de sus errores absolutos.*

RESTA Aquí $N = u_1 - u_2 \Longrightarrow \Delta N = E_a = \Delta u_1 - \Delta u_2$
Sin embargo, como los errores Δu_1 y Δu_2 pueden ser positivos o negativos, hemos de tomar la suma de los valores absolutos de los errores en orden a obtener el error máximo. Entonces :

$$\Delta N = |\Delta u_1| + |\Delta u_2|$$

MULTIPLICACIÓN En éste caso $N = u_1 u_2 \ldots u_n$ y por tanto :

$$E_r = \frac{\Delta N}{N} = \frac{\Delta u_1}{u_1} + \frac{\Delta u_2}{u_2} + \cdots + \frac{\Delta u_n}{u_n}$$

$\Longrightarrow$ El error relativo de un producto de n números aproximados es por lo tanto igual a la suma algebraica de los errores relativos de los números.
La exactitud de un producto se investigará siempre por medio del error relativo. El error absoluto, si se desea, se puede encontrar por la relación $E_a = E_r N$.

DIVISIÓN Aquí tenemos $N = u_1/u_2 \Longrightarrow E_r = \Delta u_1/u_1 + \Delta u_2/u_2 \Longrightarrow$ el error relativo de un cociente es igual a la suma algebraica de los errores relativos de dividendo y divisor, pero en orden a obtener el error máximo hemos de tomar la suma aritmética de los errores. Una fórmula simple para el error absoluto de un cociente se puede encontrar directamente, como sigue :
Sea ΔQ= error absoluto del cociente $u_1/u_2 \Longrightarrow$

$$\Longrightarrow \Delta Q = \frac{u_1 + \Delta u_1}{u_2 + \Delta u_2} - \frac{u_1}{u_2} = \frac{u_2 \Delta u_1 - u_1 \Delta u_2}{u_2(u_2 + \Delta u_2)} = \frac{u_1(\frac{\Delta u_1}{u_1} - \frac{\Delta u_2}{u_2})}{u_2 + \Delta u_2}$$

Sea w el mayor de los valores absolutos de $\Delta u_1/u_1$ y $\Delta u_2/u_2$ y tomemos los signos de Δu_1 y Δu_2 de forma que obtengamos el valor mayor de ΔQ.
Entonces, como $\Delta u_2/u_2 \leqslant \omega$ tenemos $\Delta u_2 \leqslant u_2\omega$ y por lo tanto si u_1 y u_2 están ambos sujetos a error del mismo orden de magnitud, tenemos :

$$\Delta Q \leqslant \frac{u_1(\omega + \omega)}{u_2 - \omega u_2} = \frac{2u_1\omega}{u - 2(1 - \omega)}$$

Si sólo u_1 ó u_2 están sujetos a error y los otros son libres de error en comparación con él, entonces :

$$\Delta Q \leqslant \frac{u_1\omega}{u_2 - \omega u_2} = \frac{2u_1\omega}{u_2(1-\omega)}$$

Finalmente, si ω es despreciable en comparación con 1, obtenemos $\Delta Q \leqslant 2(u_1/u_2)\omega$.
En el caso de que u_1 y u_2 son ambos sujetos a error del mismo orden de magnitud, y $\Delta Q \leqslant (u_1/u_2)\omega$ si y solo si u_1 ó u_2 están sujetos a error.
Como en el caso del producto, la exactitud de un cociente se investigará siempre por medio del error relativo.

POTENCIAS Y RAÍCES Aquí N tiene la forma $N = u^m \Longrightarrow E_r \leqslant m(\Delta u/u)$.
Para la p–ésima potencia de un número ponemos $m = p$ y tenemos :

$$E_r \leqslant p\left(\frac{\Delta u}{u}\right)$$

El error relativo de la p–ésima potencia de un número es así p veces el error relativo del número dado.
Para la raíz r–ésima de un número ponemos $m = 1/r$ y obtenemos:

$$E_r \leqslant \frac{1}{p}\frac{\Delta u}{u}$$

y por tanto el error relativo de la raíz r–ésima de un número aproximado es sólo $1/r$ veces el error relativo del número dado.

LOGARITMOS Aquí tenemos

$$N = \log_{10} u = 0.43429\ \ln u \Longrightarrow \Delta N = 0.43429\,\frac{\Delta u}{u} \Longrightarrow \Delta N < \frac{1}{2}\frac{\Delta u}{u}$$

El error absoluto en el logaritmo decimal de un número es así menos que la mitad del error relativo del número dado.
Sin embargo, un error en un logaritmo puede causar un error desastroso en el antilogaritmo o número correspondiente ya que :

$$\Delta u = \frac{u\Delta N}{0.43429} = 2.3026\,u\Delta N$$

El error en el antilogaritmo puede así ser varias veces el error en el logaritmo. Por ésta razón, es de suma importancia que el logaritmo de un resultado esté lo más libre de error posible.

1.7 CÁLCULOS CON COMA FLOTANTE

A continuación vienen algunos hechos que se dan en la matemática numérica y que conviene conocer.

1.7.1 Cancelación catastrófica

Supongamos que queremos calcular $e^{5.5}$ y no tenemos una función disponible para calcular e^x y lo haremos por tanto partiendo de :

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

Supongamos además que nuestro sistema de representación en coma flotante es en base 10 y con una precisión de 5 dígitos para la mantisa . Tendremos entonces :

$$\begin{aligned} e^{5.5} =& 1.0000 - 5.5000 + 15.125 - 27.730 + 38.129 - 41.942 \\ &+ 38.446 - 30.208 + 20.768 - 12.692 \\ &+ 6.9803 - 3.4902 + 1.5997 \cdots = +0.0026363 \end{aligned}$$

La suma termina después de 25 términos ya que los términos siguientes no la cambian.
Sin embargo, el valor real es $e^{5.5} = 0.00408677$ de forma que el resultado obtenido no tiene un solo dígito exacto.
¿Qué fué mal? Nótese que algunos de los términos, y consiguientemente sumas intermedias son varios órdenes de magnitud mayor que la respuesta final. En efecto, términos como 38.129 tienen un error de redondeo tan grande como el resultado final. Verdaderamente, los cuatro digitos más significativos de cada uno de los ocho términos que exceden de 10 en módulo se han perdido. Sería necesario que para estos ocho términos tuviesen diez dígitos significativos en la respuesta. Sin embargo, un onceavo dígito inicial sería necesario para lograr que el sexto dígito sea correcto en la suma.
Este fenómeno se llama *cancelación catastrófica*. Sin embargo, es importante darse cuenta que ésta gran cancelación no es la causa del error en la respuesta; solo magnifica el error previamente presente en los términos.
Aunque es posible conseguir más digitos significativos en un cálculo parar evitar la cancelación catastrófica, es siempre costosa en almacenamiento y tiempo de ejecución y a menudo requiere técnicas especiales de programación.
Para éste problema existe una cura mucho mejor, que es calcular la suma para $x = 5.5$ y entonces tomar el recíproco de la respuesta :

$$e^{-5.5} = \frac{1}{e^{5.5}} = \frac{1}{1 + 5.5 + 15.125 + \cdots} = 0.0040865$$

con nuestra aritmética de cinco digitos. (el error se reduce al 0.007%)
Uno de los puntos importantes a aprender de éste ejemplo es que un cálculo no tiene que ser muy largo para incurrir en serios errores de redondeo.

1.7.2 Inestabilidad de algunos algoritmos

El problema de calcular $e^{5.5}$ discutido anteriormente ilustra cómo un algoritmo mal concebido puede proporcionar una respuesta pobre a un problema prefectamente bien planteado. La dificultad fué corregida cambiando el algoritmo.
Para ciertos problemas, respuestas "buenas" no se pueden obtener por *ningún* algoritmo porque el problema es sensible a los pequeños errores cometidos en la representación de los datos y en la aritmética usada.
Es importante distinguir entre éstas dos clases de peligros porque existen algoritmos inestables y problemas sensibles en casi todas las ramas de la matemática numérica. Una vez se es consciente de estos síntomas, estos problemas son bastante fáciles de diagnosticar.
Veamos un ejemplo de algoritmo inestable. Supongamos que queremos calcular las integrales :

$$\boxed{E_n = \int_0^1 x^n e^{x-1} dx, \quad n = 1, 2 \dots}$$

por partes :

$$\int_0^1 x^n e^{x-1} dx = x^n e^{x-1}\Big|_0^1 - \int_0^1 n x^{n-1} e^{x-1} dx$$

o también

$$E_n = 1 - E_{n-1}, \quad n = 2, 3, \dots$$

resultando $E_1 = 1/e$ (calcular).
Usando base 10 y una precisión de 6 dígitos podemos usar ésta recurrencia para calcular aproximaciones a los primeros nueve valores de E_n :

$$\begin{array}{lll} E_1 \approx 0.367879 & E_4 \approx 0.170904 & E_7 \approx 0.110160 \\ E_2 \approx 0.264242 & E_5 \approx 0.145480 & E_8 \approx 0.118720 \\ E_3 \approx 0.207274 & E_6 \approx 0.127120 & \mathbf{E_9} \approx -\mathbf{0.0684800}\,!! \end{array}$$

Aunque el integrando $x^9 e^{x-1}$ es positivo en todo el intervalo $]0, 1[$ nuestro valor para E_9 es negativo ! ¿Qué ha causado el error?
Obsérvese que el único error de redondeo cometido en los cálculos anteriores fué en E_1, por el cual $1/e$ se redondeó a seis dígitos significativos.

Como la fórmula de recurrencia obtenida por integración es exacta para aritmética real, el error en E_9 es enteramente debido al error de redondeo en E_1.
Para ver cómo el error de aproximadamente 4.412×10^{-7} en E_1 llega a ser tan grande, nótese que es multiplicado por -2 en el cálculo de E_2, y entonces el error en E_2 es multiplicado por -3 calculando E_3 y así sucesivamente.
Asi, el error en E_9 es exactamente el error en E_1 multiplicado por $(-2)(-3)\cdots(-9) = 9! = 362880$ o sobre $9! \times 4.412 \times 10^{-7} \approx 0.1601$
Esta enorme magnificación del error en los datos del problema es resultado del algoritmo que hemos elegido.
El verdadero valor de E_9 (a tres cifras significativas) es $-0.06848 + 0.1601 = 0.0916$
¿Cómo podemos elegir un algoritmo diferente que evite la inestabilidad?
Si reescribimos la relación de recurrencia como :

$$E_{n-1} = \frac{1 - E_n}{n}, \quad n = \ldots, 3, 2$$

entonces, en cada etapa del cálculo, el error en E_n es disminuido por un factor de $1/n$. Asi, si empezamos con un valor para algún E_n con $n \gg 1$ y trabajamos hacia atrás, cualquier error inicial o error de redondeo que ocurran serán disminuidos en cada paso. Esto es llamado un **algoritmo estable**.
Para obtener un valor de comienzo notemos que :

$$E_n = \int_0^1 x^n e^{x-1} dx \leqslant \int_0^1 x^n dx = \left.\frac{x^{n+1}}{n+1}\right|_0^1 = \frac{1}{n+1}$$

Asi, $E_n \xrightarrow{n\to\infty} 0$. Por ejemplo, si aproximamos E_{20} por cero y lo usamos como valor de comienzo, estaremos cometiendo un valor inicial de como máximo $1/21$. Este error es multiplicado por $1/20$ al calcular E_{19} y así el error en E_{19} es a lo sumo $(1/20)(1/21) \approx 0.0024$. Para E_{15}, el error inicial ha quedado reducido a menos de 4×10^{-8} que es menor que el error de redondeo. Realizando los cálculos nos da :

$$\begin{array}{lll}
E_{20} \approx 0.0 & E_{16} \approx 0.0557190 & E_{12} \approx 0.0717733 \\
E_{19} \approx 0.05 & E_{15} \approx 0.0590176 & E_{11} \approx 0.0773523 \\
E_{18} \approx 0.05 & E_{14} \approx 0.0627322 & E_{10} \approx 0.0838771 \\
E_{17} \approx 0.0527778 & E_{13} \approx 0.0669477 & E_{9} \approx 0.0916123
\end{array}$$

En E_{15}, el error inicial en E_{20} ha sido completamentamente eliminado por la estabilidad del algoritmo, y los valores calculados de $E_{15}, \cdots, E_9$ son seguros a la precisión total de 6 dígitos, con un posible error de redondeo en el último dígito.

En conclusión, como un resultado intermedio de un cálculo numérico está afectado por un error de redondeo, éste error afectará a todos los resultados subsecuentes que dependen del resultado intermedio. O sea, el error de redondeo *se propaga*.
El efecto de cancelación catastrófica es un caso especial de éste fenómeno. El error de redondeo se propagaría incluso si todos los cálculos posteriores se realizaran con precisión infinita. En realidad, cada nuevo resultado intermedio introduce un nuevo error de redondeo, y todos estos errores afectarán al resultado final.
En situaciones simples (como la evaluación de una suma) sabemos que el error final es sencillamente la suma de todos los errores intermedios. Debido a efectos estadísticos, los errores intermedios acaso se cancelan unos con otros, al menos parcialmente.
En otros casos, (como iteración) los errores intermedios tendrán un efecto despreciable en el resultado final.
Algoritmos con esta propiedad se llaman estables. *La inestabilidad numérica se presenta entonces si los errores intermedios tienen una fuerte influencia en el resultado final.*

1.7.3 Condicionamiento de un problema

Ciertos problemas de cálculo son extremadamente sensibles a los datos. Este aspecto del análisis numérico es independiente de la representación en coma flotante o de los algoritmos usados.
Supongamos el polinomio cuadrático

$$(x-2)^2 = 10^{-6}$$

que tiene por raíces 2 ± 10^{-3}.
Sin embargo, un cambio de 10^{-6} en el término independiente puede producir un cambio de 10^{-3} en las raíces.
Pero no solo para polinomios con ceros cercanos se puede observar esta inestabilidad.
El siguiente ejemplo se debe a Wilkinson [Wil94]. Sea

$$p(x) = (x-1)(x-2)\dots(x-19)(x-20) = x^{20} - 210\,x^{19} + \cdots$$

que tiene los ceros $1, 2, \dots, 19, 20$ y están bien separados.
Wilkinson cambió el coeficiente de x^{19} del valor -210 al valor $-210 + 2^{-23}$ y estudió el efecto que tendría éste pequeño cambio en los ceros del polinomio. Calculó cuidadosamente las raíces en una máquina con una mantisa de 90 dígitos binarios.
Las soluciones, correctamente redondeadas al número de dígitos que tienen, son:

1.00000 0000	10.09526 6145 $\pm$ 0.64350 0904 i
2.00000 0000	11.79363 3881 $\pm$ 1.65232 9728 i
3.00000 0000	13.99235 8137 $\pm$ 2.51883 0070 i
4.00000 0000	16.73073 7466 $\pm$ 2.81262 4894 i
4.99999 9928	19.50243 9400 $\pm$ 1.94033 0347 i
6.00000 6944	
8.00726 7603	
8.91725 0249	
20.84690 8101	

Un pequeño cambio en el coeficiente -210 ha causado que diez de los ceros se conviertan en complejos.La tabla anterior se realizó por un cálculo muy preciso y no tiene errores de redondeo apreciables. La razón de que estos ceros se desplacen no es una cuestión de redondeo, ni siquiera una cuestión asociada al algoritmo usado para calcularlas sino una cuestión de sensibilidad del problema en sí mismo. Vamos a analizar lo que ocurrió. Podemos escribir el polinomio en la forma

$$p(x,\alpha) = x^{20} - \alpha\, x^{19} + \cdots$$

y entonces encontrar la derivada parcial de x respecto de α en cada una de las raíces de $p(x)$. Esto se hace diferenciando la ecuación $p(x,\alpha)$ con respecto a α :

$$\frac{\partial p(x,\alpha)}{\partial x}\frac{\partial x}{\partial \alpha} + \frac{\partial p(x,\alpha)}{\partial \alpha} = 0 \Longrightarrow \frac{\partial x}{\partial \alpha} = -\frac{\partial p/\partial \alpha}{\partial p/\partial x} = \frac{x^{19}}{\displaystyle\sum_{i=1}^{20}\prod_{\substack{j=1\\ j\neq i}}^{20}(x-j)}$$

y evaluando ésta expresión en cada raíz da:

$$\left.\frac{\partial x}{\partial \alpha}\right|_{x=i} = \frac{i^{19}}{\displaystyle\prod_{\substack{j=1\\ j\neq i}}^{20}(i-j)}, \quad i = 1, 2, \ldots, 20$$

Estos números dan una medida directa de la sensibilidad de cada una de las raíces al coeficiente α. Los valores que se obtienen son:

Raíz	$\partial x/\partial\alpha\Big\vert_{x=i}$	Raíz	$\partial x/\partial\alpha\Big\vert_{x=i}$	Raíz	$\partial x/\partial\alpha\Big\vert_{x=i}$	Raíz	$\partial x/\partial\alpha\Big\vert_{x=i}$
1	-8.2×10^{-18}	6	5.8×10^{1}	11	-4.6×10^{7}	16	2.4×10^{9}
2	8.2×10^{-11}	7	-2.5×10^{3}	12	2.0×10^{8}	17	-1.9×10^{9}
3	-1.6×10^{-6}	8	6.0×10^{4}	13	-6.1×10^{8}	18	1.0×10^{9}
4	2.2×10^{-3}	9	-8.3×10^{5}	14	1.3×10^{9}	19	-3.1×10^{8}
5	-6.1×10^{-1}	10	7.6×10^{6}	15	-2.1×10^{9}	20	4.3×10^{7}

1.7.4 Resolviendo una ecuación cuadrática

Existe un famoso algoritmo para resolver un polinomio de segundo grado, implícito en el siguiente teorema matemático :

Teorema 1.7.1. *Si a, b, c son reales y $a \neq 0$, entonces la ecuación $ax^2 + bx + c = 0$ se satisface por dos valores de x, que son*

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

Veamos como funcionan las fórmulas en un sistema de coma flotante F que usa una mantisa de 8 cifras significativas decimales y con un exponente en el rango de -50 a +50.

CASO 1: $a = 1, b = -10^5, c = 1$

las verdaderas raíces de la correspondiente ecuación cuadrática, correctamente redondeadas a 11 cifras significativas, son:

$$x_1 \approx 99999.999990, \quad x_2 \approx 0.000010000000001$$

Si usamos las expresiones del teorema, obtenemos:

$$x_1 \approx 100000.00, \quad x_2 \approx 0\,!!$$

De nuevo, al calcular x_2 hemos sido víctimas de la cancelación catastrófica que, como antes, revela el error que hemos cometido en la forma de calcular x_2. Existen varias alternativas para calcular las raíces de la ecuación cuadrática que no producen dicha cancelación.

Una sería utilizar el signo de b para determinar cúal de las fórmulas causa la menor cancelación y entonces usar esa fórmula para calcular una de las raíces :

$$x_1 = -\frac{b + signo(b)\sqrt{b^2 - 4ac}}{2a}$$

y entonces, como $ax^2 + bx + c = a(x - x_1)(x - x_2)$ esto implica que $ax_1x_2 = c$ y por tanto la otra raiz se puede calcular por

$$x_2 = \frac{c}{ax_1}$$

Para el caso expuesto,éste método nos da $x_1 = 100000.00, \quad x_2 = 1.0000000/100000.00 = 0.000010000000$, ambas aceptables.

CASO 2: $a = 6, b = 5, c = -4$

No hay dificultad en calcular

$$x_1 \approx 0.50000000, \quad x_2 \approx -1.3333333$$

o próximo a estos valores, según las fórmulas que se utilicen.

CASO 3: $a = 6 \times 10^{30}, b = 5 \times 10^{30}, c = -4 \times 10^{30}$

Como los coeficientes en éste caso son los del caso 2 multiplicados por 10^{30} las raíces no cambian. Sin embargo, la aplicación de cualesquiera de la fórmulas causa que overflow se produzca de inmediato, ya que $b^2 > 10^{50}$, fuera del rango de F. Probablemente éste tamaño uniformemente grande en $|a|, |b|$ y $|c|$ se podría detectar antes de entrar en el algoritmo, y los tres números se podrían dividir por el factor 10^{30} para reducir el problema al caso 2.

CASO 4: $a = 10^{-30}, b = -10^{30}, c = 10^{30}$

En este caso una de las raíces es próxima a 1, mientras que la otra es próxima a 10^{60}. Nuestro algoritmo debería determinar la primera, aunque la segunda caiga fuera del rango de F. Obviamente, cualquier intento de hacer que los coeficientes tengan magnitud aproximada simplemente dividiendo todos ellos por el mismo número está condenado al fracaso, ya que causaría un error de underflow o de overflow, según el caso. Esta ecuación, es un test severo para un algoritmo de resolución de ecuaciones cuadráticas e incluso para el sistema de representación que se use.

CASO 5: $a = 1.0000000, b = -4.0000000, c = 3.9999999$

Aquí las dos raíces son $x_1 \approx 1.999683772, \quad x_2 \approx 2.000316228$. Pero aplicando las fórmulas nos da

$$x_1 = x_2 = 2.0000000$$

con solo los cuatro primeros dígitos correctos.

Esta ecuación es la cuadrática $(x-2)^2 = e$ siendo $e = 0000001$ y tal como se discutió en un apartado previo, ésta es una forma de sensibilidad de la ecuación en sí y no del método para resolverla.

Esta ecuación se puede resolver aumentando la precisión a 16 dígitos para la mantisa, por ejemplo.

Bibliografía

[CC87] S.C. Chapra and R.P. Canale. *Métodos Numéricos para Ingenieros*. McGraw-Hill, México, 1987.

[CC89] S.C. Chapra and R.P. Canale. *Numerical Methods for Engineers*. McGraw-Hill International, New York, second edition, 1989.

[FMM77] G. E. Forsythe, M. A. Malcolm, and C. B. Moler. *Computer Methods for Mathematical Computations*. Prentice-Hall, Englewood Cliffs, N.J., 1977.

[For70] George E. Forsythe. Pitfalls in computation, or why a math book isn't enough. *American Mathematical Monthly*, 77:931–956, 1970.

[Gol91] David Goldberg. What every computer scientist should know about floating-point arithmetic. *ACM Computing Surveys*, 23(1):5–48, 1991.

[Hig96] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, 1996.

[IEE85] *IEEE Standard for Binary Floating-Point Arithmetic, ANSI/IEEE Standard 754-1985*. Institute of Electrical and Electronics Engineers, New York, 1985. Reprinted in SIGPLAN Notices, 22(2):9–25, 1987.

[Knu81] D. E. Knuth. *Seminumerical Algorithms*, volume 2 of *The Art of Computer Programming*. Addison-Wesley, Reading, Massachusetts, second edition, 1981. 2

[PPB85] Jr. P.W Purdom and C.A. Brown. *The Analysis of Algorithms*. Holt, Rinehart and Winston, New York, 1985. 1

[SA56] Irene A. Stegun and Milton Abramowitz. Pitfalls in computation. *J. Soc. Indust. Appl. Math.*, 4(4):207–219, 1956.

[Sca66] James B. Scarborough. *Numerical Mathematical Analysis*. Johns Hopkins University Press, Baltimore, MD, USA, second edition, 1966. 1.3

[Wil94] J. H. Wilkinson. *Rounding Errors in Algebraic Processes*. Dover Publications, New York, 1994. Unabridged and unaltered republication of the work first published by Prentice Hall, Englewood Cliffs, in 1963. 1.7.3

Capítulo 2

INTERPOLACIÓN

2.1 INTRODUCCIÓN.COEFICIENTES INDETERMINADOS

Supongamos un conjunto de puntos $(x_i, y_i), i = 0, 1, \ldots, n$ donde $y_i = f(x_i)$. Se dice que la función g interpola a f en los puntos (x_i, y_i) si

$$g(x_i) = f(x_i) = y_i, \text{ para cada } i = 0, 1, \cdots, n$$

En el caso de interpolación polinómica buscamos un polinomio

$$g(x) = p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

donde se quieren determinar los coeficientes $a_j,\ 0 \leqslant j \leqslant n$.
Por tanto se han de verificar las $n+1$ ecuaciones :

$$y_j = p(x_j) = a_0 + a_1 x_j + a_2 x_j^2 + \cdots + a_n x_j^n, \quad 0 \leqslant j \leqslant n$$

que expresado en forma matricial el sistema queda :

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}$$

o bien $V\vec{a} = \vec{y}$, donde V es una matriz conocida con el nombre de matriz de Vandermonde, que es no singular ya que si los puntos x_i son distintos su determinante vale:

$$|V| = (x_n - x_{n-1}) \cdots (x_n - x_0)(x_{n-1} - x_{n-2}) \cdots (x_1 - x_0) = \prod_{\substack{i,j=n \\ i>j}}^{0} (x_i - x_j)$$

que asegura que existe *solución única* si $x_i \neq x_j$, para $i \neq j$.
Este método de obtener los coeficientes mediante el planteamiento de un sistema lineal se llama **método de coeficientes indeterminados**.

Ejemplo 6. *Encontrar el polinomio de interpolación $p(x)$ de segundo grado tal que $p(0) = -1$, $p(1) = 2$, $p(2) = 7$.*

Buscamos el polinomio de grado dos $p(x) = a_0 + a_1x + a_2x^2$ que verifica las condiciones anteriormente expuestas. Por el método de coeficientes indeterminados con $x_0 = 0$, $x_1 = 1$, $x_2 = 2$ obtenemos :

$$\begin{aligned} a_0 + 0 + 0 &= -1 \\ a_0 + a_1 + a_2 &= 2 \\ a_0 + 2a_2 + 4a_2 &= 7 \end{aligned} \tag{2.1}$$

y resolviendo el sistema (p.e., por Gauss) obtenemos $p(x) = x^2 + 2x - 1$.
El método de los coeficientes indeterminados es un procedimiento con amplia aplicación a otros tipos de problemas de interpolación. Por ejemplo, si buscamos una función no polinómica en el sentido usual como en el siguiente ejemplo.

Ejemplo 7. *Queremos encontrar un polinomio trigonométrico $p(x)$ de la forma*

$$p(x) = c_0 + c_1 \cos x + c_2 \cos(2x) + c_3 \cos(3x)$$

tal que $p(x_j) = y_j$, $0 \leqslant j \leqslant 3$ siendo $x_0 = 0, x_1 = \pi/6, x_2 = \pi/4, x_3 = \pi/3$ con $y_0 = 0, y_1 = 2 - \sqrt{3}/2, y_2 = 1 + \sqrt{2}/2, y_3 = 3/2$.

El sistema lineal es :

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \sqrt{3}/2 & 1/2 & 0 \\ 1 & \sqrt{2}/2 & 0 & -\sqrt{2}/2 \\ 1 & 1/2 & -1/2 & -1 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 - \sqrt{3}/2 \\ 1 + \sqrt{2}/2 \\ 3/2 \end{pmatrix}$$

y la solución del sistema (por eliminación gaussiana, p.e.) produce $c_0 = 1, c_1 = -1, c_2 = 2, c_3 = -2$. Por tanto :

$$p(x) = 1 - \cos(x) + 2\cos(2x) - 2\cos(3x)$$

2.2 FÓRMULA DE LAGRANGE

Supongamos que $\{x_0, x_1, \cdots, x_n\} = \{x_i\}_{i=0}^n$ son $n+1$ puntos distintos del eje real y que $f : \mathbb{R} \to \mathbb{R}$ está definida sobre $I = [a, b]$ con $\{x_i\}_{i=0}^n \subseteq [a, b]$. Tenemos entonces :
Teorema 2.2.1. *Existe un único polinomio $p(x)$ de grado no mayor que n que interpola a f en los puntos $x_0, x_1, \cdots, x_n$:*

$$p(x_i) = f(x_i), \qquad i = 0, 1, \cdots, \ldots, \ldots, n$$

Unicidad. Sea $q(x)$ otro polinomio de grado menor o igual que n que interpola a f en $\{x_i\}_{i=0}^n$. Entonces :

$$h(x) = p(x) - q(x)$$

es un polinomio de grado menor o igual que n que cumple

$$h(x_i) = p(x_i) - q(x_i) = f(x_i) - f(x_i) = 0, \quad i = 0, 1, \ldots, n$$

o sea, $h(x)$ tiene al menos $n+1$ ceros distintos $\Longrightarrow h(x) = 0$ (idénticamente nulo) $\Longrightarrow p(x) = q(x), \quad \forall x.$ □

Existencia. Veamos ahora como se puede construir; escribiremos por brevedad

$$f(x_k) = y_k, \quad k = 0, 1, \ldots, n.$$

En primer lugar, construiremos un polinomio de grado n que sea nulo en todos los puntos x_i salvo en uno x_k en el cual valga 1. Tiene que ser de la forma :

$$L_k(x) = a \prod_{\substack{i=0 \\ i \neq k}}^{n} (x - x_i), \quad \text{siendo } a \in \mathbb{R}$$

y como su valor para $x = x_k$ debe ser 1 tenemos :

$$a = \frac{1}{\prod\limits_{\substack{i=0 \\ i \neq k}}^{n} (x_k - x_i)}$$

con lo que queda :

$$L_k(x) = \prod_{\substack{i=0 \\ i \neq k}}^{n} \frac{x - x_i}{x_k - x_i} = \frac{(x - x_0)(x - x_1) \cdots (x - x_{k-1})(x - x_{k+1}) \cdots (x - x_n)}{(x_k - x_0)(x_k - x_1) \cdots (x_k - x_{k-1})(x_k - x_{k+1}) \cdots (x_k - x_n)}$$

verificándose entonces que

$$L_k(x_i) = \delta_{ki}(\text{delta de Kronecker }) = \begin{cases} 1\,, & \text{si } i = k, \\ 0\,, & \text{si } i \neq k \end{cases}$$

y esto para $i = 0, 1, \ldots, n$, dentro de cada $k = 0, 1, \ldots, n$.
Por tanto, si se desea un polinomio de grado n que tome respectivamente los valores $y_0, y_1, \cdots, y_n$ en los puntos $x_0, x_1, \cdots, x_n$ basta tomar :

$$\begin{aligned} p(x) =\ & y_0\, L_0(x) + y_1\, L_(x) + \cdots + y_n\, L_n(x) = \\ & = \sum_{k=0}^{n} y_k\, L_k(x) = \sum_{k=0}^{n} y_k \prod_{\substack{i=0\\ i\neq k}}^{n} \frac{x - x_i}{x_k - x_i} = \\ & = y_0 \frac{(x-x_1)(x-x_2)\cdots(x-x_n)}{(x_0-x_1)(x_0-x_2)\cdots(x_0-x_n)} + y_1 \frac{(x-x_0)(x-x_2)\cdots(x-x_n)}{(x_1-x_0)(x_1-x_2)\cdots(x_1-x_n)} + \\ & + y_2 \frac{(x-x_0)(x-x_1)(x-x_3)\cdots(x-x_n)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)\cdots(x_2-x_n)} + \cdots + y_n \frac{(x-x_0)(x-x_1)\cdots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\cdots(x_n-x_{n-1})} \end{aligned}$$

que se denomina **fórmula de Lagrange** del polinomio de interpolación. □

Efectivamente ocurre que :

$$p(x_i) = \sum_{k=0}^{n} y_k\, L_k(x_i) = 0 + \underbrace{\cdots}_{\text{i}} + 0 + y_i\,.1 + 0 + \underbrace{\cdots}_{\text{n-i}} + 0 = y_i$$

y esto para cada $i = 0, 1, \ldots, n$, con lo cual verifica las condiciones a cumplir por el polinomio que interpola en los puntos $\{x_i\}_{i=0}^{n}$.

Ejemplo 8. *Encontrar el polinomio de interpolación* $p(x)$ *de segundo grado tal que* $p(0) = -1$, $p(1) = 2, p(2) = 7$.

Tomando las x_i e y_i en el orden dado: $x_0 = 0$, $x_1 = 1$, $x_2 = 2$; $y_0 = -1$, $y_1 = 2$, $y_2 = 7$ Por la fórmula de Lagrange tenemos :

$$\begin{aligned} L_0(x) &= \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} = \frac{(x-1)(x-2)}{2}, \\ L_1(x) &= \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} = -x(x-2), \\ L_2(x) &= \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} = \frac{x(x-1)}{2} \end{aligned}$$

Por tanto, $p(x)$ viene dado por la siguiente fórmula :

$$p(x) = y_0\, L_0(x) + y_1\, L_1(x) + y_2\, L_2(x) = -L_0(x) + 2\, L_1(x) + 7\, L_2(x) = x^2 + 2x - 1$$

En general no se nos pide la expresión explícita del polinomio de interpolación como en el ejemplo anterior, sino **el valor de ese polinomio** en uno o varios puntos en los que se quiere interpolar, como es el caso del siguiente ejemplo.

Ejemplo 9. *Obtener por interpolación el valor para* $x = 3$ *conocidos los valores* $x_0 = 0, y_0 = -1$; $x_1 = 1, y_1 = 0$; $x_2 = 2, y_2 = 7$; $x_3 = 4, y_3 = 63$.

Por la fórmula de Lagrange tenemos, sustituyendo ya el valor $x = 3$:

$$\begin{aligned}
L_0(x) &= \frac{(x-1)(x-2)(x-4)}{(0-1)(0-2)(0-4)} \Longrightarrow L_0(3) = \frac{(3-1)(3-2)(3-4)}{-8} = \frac{2.1.(-1)}{-8} = \frac{1}{4} \\
L_1(x) &= \frac{x(x-2)(x-4)}{(1-0)(1-2)(1-4)} \Longrightarrow L_1(3) = \frac{3.1.(-1)}{1.(-1).(-3)} = -1 \\
L_2(x) &= \frac{x(x-1)(x-4)}{(2-0)(2-1)(2-4)} \Longrightarrow L_2(3) = \frac{3.2.(-1)}{2.(1).(-2)} = \frac{3}{2} \\
L_3(x) &= \frac{x(x-1)(x-2)}{(4-0)(4-1)(4-2)} \Longrightarrow L_3(3) = \frac{3.2.1}{4.3.2} = \frac{1}{4}
\end{aligned}$$

Entonces

$$\begin{aligned}
p(3) &= y_0\, L_0(3) + y_1\, L_1(3) + y_2\, L_2(3) + y_3\, L_3(3) = (-1).\frac{1}{4} + 0.(-1) + 7.\frac{3}{2} + 63.\frac{1}{4} = \\
&= -\frac{1}{4} + \frac{21}{2} + \frac{63}{4} = 26
\end{aligned}$$

que es lo que tiene que dar ya que los valores dados son de la función $f(x) = x^3 - 1$. Obsérvese que podríamos habernos ahorrado el cálculo de $L_1(x)$ ya que $y_1 = 0$ y el resultado del sumando siempre será cero.

2.3 ERROR DEL POLINOMIO DE INTERPOLACIÓN

Teorema 2.3.1 (ERROR). *Sea* $f\colon I \to \mathbb{R}, \{x_i\}_{i=0}^{n} \subseteq I, x_i \neq x_j$ *para* $i \neq j$ *y supongamos que* f *es derivable* $n+1$ *veces en* I *con derivada continua* $\Longrightarrow \forall x \in I, \exists \xi_x \in$ *menor de los intervalos que contiene a los puntos* $x, x_0, x_1, \ldots, x_n$ *tal que :*

$$\boxed{E = f(x) - p(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}(x-x_0)(x-x_1)\ldots(x-x_n)}$$

donde $p(x)$ *es el polinomio que interpola a* f *en* $\{x_i\}_{i=0}^{n}$.

Demostración. Si x es uno de los puntso x_k no hay nada que probar ya que ambos miembros se anulan para cualquier ξ.
Si x es un valor fijo diferente de los x_k, consideramos la función auxiliar $F = F(t)$ definida por :

$$F(t) = f(t) - p(t) - cL(t), \qquad \text{donde } c = \frac{f(x) - p(x)}{L(x)} \tag{2.2}$$

y donde estamos llamando $L(x) = (x - x_0)(x - x_1)\ldots(x - x_n) = \prod_{i=0}^{n}(x - x_i)$.
Tenemos $F(x_k) = f(x_k) - p(x_k) - cL(x_k) = y_k - y_k - 0 = 0$ para $k = 0, 1, 2, \ldots, n$ y también $F(x) = f(x) - p(x) - cL(x) = 0$, por definición de c.
La función F tiene entonces al menos $n+2$ ceros distintos en el intervalo I. Por el teorema de Rolle, F' debe tener por lo menos $n+1$ ceros en el menor de los intervalos que contiene a x y los x_k, la segunda derivada F'' debe tener no menos de n ceros, $\cdots$, la $(n+1)$-ésima derivada debe tener por lo menos un cero. Sea ξ_x tal cero. Derivando $(n+1)$ veces la ecuación (2.2) y haciendo $t = \xi_x$:

$$0 = F^{(n+1)}(\xi_x) = f^{(n+1)}(\xi_x) - c(n+1)! \tag{2.3}$$

ya que la derivada $(n+1)$-ésima de $p(x)$ es cero. Por tanto, usando (2.3) tenemos :

$$cL(x) = f(x) - p(x) = \frac{1}{(n+1)!}f^{(n+1)}(\xi_x)L(x)$$

□

Ejemplo 10. *¿Cúal es el error máximo que puede presentarse con dos puntos de interpolación?*

Supongamos dos puntos de interpolación $(x_0, f(x_0)), (x_1, f(x_1))$. Entonces el polinomio es :

$$p(x) = f(x_0)\frac{x - x_1}{x_0 - x_1} + f(x_1)\frac{x - x_0}{x_1 - x_0} = \frac{(x_1 - x)f(x_0) + (x - x_0)f(x_1)}{x_1 - x_0}$$

y por otra parte :

$$E = f(x) - p(x) = \frac{(x - x_0)(x - x_1)}{2!}f''(\xi_x)$$

Supongamos que $|f''(x)| \leqslant M$, $\forall x \in [x_0, x_1]$. El máximo de la función $|\frac{1}{2}(x - x_0)(x - x_1)|$ entre x_0 y x_1 se presenta en $x = \frac{1}{2}(x_0 + x_1)$ con valor $\frac{1}{8}(x_1 - x_0)^2$. Por tanto:

$$|f(x) - p(x)| \leqslant \frac{(x_1 - x_0)^2}{8}M$$

Por ejemplo, si calculamos el valor de $\operatorname{sen} x$ por una tabla de senos con paso h usando la interpolación lineal, el error está acotado por el valor $h^2/8$ ya que $M = 1$ en éste caso.

Ejemplo 11. *¿Con qué grado de exactitud podemos calcular $\sqrt{115}$ mediante interpolación polinómica para la función $y = \sqrt{x}$ si elegimos los puntos $x_0 = 100, x_1 = 121, x_2 = 144$? ¿Y si se eligen $x_0 = 100, x_1 = 110, x_2 = 120$?*

Tenemos $y' = \frac{1}{2}x^{-\frac{1}{2}},\ y'' = -\frac{1}{4}x^{-\frac{3}{2}},\ y''' = \frac{3}{8}x^{-\frac{5}{2}}$. Entonces :

$$M = \max_{x\in[100,144]} |y'''| = \frac{3}{8}\frac{1}{\sqrt{100^5}} = \frac{3}{8}10^{-5} \text{ para } 100 \leqslant x \leqslant 144 \Longrightarrow$$

$$\Longrightarrow |E| \leqslant \frac{3}{8}\times 10^{-5}\times\frac{1}{3!}|(115-100)(115-121)(115-144)| = \frac{1}{16}\times 10^{-5}\times 15\times 6\times 29 \simeq 1.6\times 10^{-3}$$

Sin embargo, si elegimos $x_0 = 100, x_1 = 110, x_2 = 120$ obtendríamos :

$$|E| \leqslant \frac{3}{8}\times 10^{-5}\times\frac{1}{3!}\times 15\times 5\times 5 \simeq 2.3\times 10^{-4}$$

.

2.4 ELECCIÓN ÓPTIMA DE LOS PUNTOS

Sabemos que la fórmula del error para la interpolación polinómica es

$$E = f(x) - p(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}(x-x_0)(x-x_1)\ldots(x-x_n)$$

y nos interesa escoger los puntos de forma que se obtenga el mínimo error posible.
Para lograr esto utilizaremos los polinomios de Tchebychev.

2.4.1 Polinomios de Tchebychev

Definición 1. *Se definen recursivamente como :*

$$\boxed{\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_n(x) &= 2x\,T_{n-1}(x) - T_{n-2}(x), \quad \forall n \geqslant 2 \end{aligned}}$$

Ejemplo 12. *Los primeros polinomios de Tchebychev son :*

$$\begin{aligned} T_2(x) &= 2x\,T_1(x) - T_0(x) = 2x^2 - 1 \\ T_3(x) &= 2x\,T_2(x) - T_1(x) = 4x^3 - 3x \\ T_4(x) &= 2x\,T_3(x) - T_2(x) = 8x^4 - 8x^2 + 1, \quad \textit{etc.} \end{aligned}$$

Se puede probar también que se verifica la siguiente relación que nos será útil para obtener las raíces de $T_n(x)$:

$$T_n(x) = \cos(n \arccos x), \quad \text{si} \quad x \in [-1, 1]$$

y por tanto

$$\left|T_n(x)\right| \leqslant 1, \quad -1 \leqslant x \leqslant 1$$

Propiedad 1. *Se verifica que:*

$$\max \left| \prod_{i=0}^{n} (x - x_i) \right| \geqslant \frac{1}{2^n},$$

para cualquier elección posible de los x_i.

Propiedad 2. *Se verifica que :*

$$\max \left| \prod_{i=0}^{n} (x - x_i) \right| = \frac{1}{2^n},$$

si los puntos x_i son las raíces del polinomio de Tchebychev T_{n+1} de grado $n+1$.

Para calcular las raíces de $T_n(x)$, usamso la relación vista antes :

$$T_n(x) = \cos(n \arccos x) = 0$$

y recordando que $\cos x = 0 \iff x = \pi/2 + \pi k, k \in \mathbb{Z}$, tenemos que :

$$n \arccos x = \frac{\pi}{2} + k\pi \Longrightarrow \arccos x = \frac{\pi + 2\pi k}{2n} = \frac{2k+1}{2n}\,\pi$$

O sea :

$$\boxed{x_k = \cos\left(\frac{2k+1}{2n}\,\pi\right), \quad k = 0, 1, \ldots, n-1}$$

que son las n raíces del polinomio $T_n(x)$.

Ejemplo 13. *Para calcular $\sqrt{115}$ con tres puntos elegidos en el intervalo $[100, 200]$, si ahora elegimos los puntos x_i de forma que sean las raíces del polinomio de Tchebychev $T_3(x)$ tenemos que :*

$$x'_k = \cos\Big(\frac{2k+1}{6}\pi\Big), \quad k = 0, 1, 2$$

$x'_0 = \cos\frac{\pi}{6} = 0.86602541, \quad x'_1 = \cos\frac{\pi}{2} = 0, \quad x'_2 = \cos\frac{5\pi}{6} = -0.86602541$
Ahora es necesario pasar al intervalo de interpolación las raíces obtenidas, ya que las raíces del polinomio de Tchebychev de grado n caen dentro del intervalo [-1,1], siendo simétricas dentro del mismo.
Por tanto, necesitamos construir una aplicación (biyectiva) que nos transforme el intervalo [-1,1] en el intervalo [100,120] (en general al intervalo $[\min\{x_i\}, \max\{x_i\}]$)
Construimos $f\colon [-1,1] \to [100,120]$ de forma que $f(-1) = 100$ y $f(1) = 120$, pero esto no es sino un problema de interpolación que podemos resolver con la fórmula de Newton en diferencias divididas :

$$\begin{array}{lll} x_0 = -1 \longrightarrow & y_0 = 100 \searrow & \\ & & f[x_0, x_1] = \frac{120-100}{1+1} = 10 \\ x_1 = +1 \longrightarrow & y_1 = 120 \nearrow & \end{array}$$

con lo cual el polinomio de interpolación es $p(x) = 100 + 10(x+1) = 10x + 110$ y los valores correspondientes en el intervalo [100,120] de los x'_k obtenidos son :
$x_0 = p(x'_0) = 10x'_0 + 110 = 118.66025404, \quad x_1 = p(x'_1) = 110, \quad x_2 = p(x'_2) = 101.33974596$
y el error de interpolación en éste caso es :

$$|E| \leqslant \frac{3}{8} \times 10^{-5} \times \frac{1}{3!}\ |(115 - x_0)(115 - x_1)(115 - x_2)| \simeq 6.47208691 \times 10^{-5}$$

que sería el menor error que se podría cometer utilizando interpolación polinómica al calcular $\sqrt{115}$ en el intervalo [100,120] con tres puntos de interpolación.

2.5 FÓRMULA DE NEWTON EN DIFERENCIAS DIVIDIDAS

Sea $p_k(x)$ el polinomio de interpolación en los puntos $x_0, x_1, \ldots, x_k$ (grado máximo = k). Considerando $p_k(x), p_{k-1}(x)$ y su diferencia :

$$q_k(x) = p_k(x) - p_{k-1}(x)$$

vemos que para los puntos $x_0, x_1, \ldots, x_{k-1}$ tenemos que :

$$p_{k-1}(x_i) = y_i = p_k(x_i), \quad 0 \leqslant i \leqslant k-1$$

y también que para el siguiente punto x_k tenemos que $p_k(x_k) = y_k$, sin conocerse el valor a priori que pueda tener $p_{k-1}(x_k)$.
Por tanto, el polinomio $q_k(x)$ verifica :

$$q_k(x_i) = p_k(x_i) - p_{k-1}(x_i) = y_i - y_i = 0, \quad 0 \leqslant i \leqslant k-1$$

Ahora bien, $q_k(x)$ es un polinomio de grado máximo k ya que es la resta de dos polinomios, $p_k(x)$ de grado k y $p_{k-1}(x)$ de grado $k-1$ y según se acaba de ver se anula en los k puntos anteriores tiene con lo cual se puede expresar de la siguiente forma :

$$q_k(x) = a_k(x-x_0)(x-x_1)\cdots(x-x_{k-1}) = a_k\prod_{i=0}^{k-1}(x-x_i)$$

Por otra parte, en el punto x_k se cumple :

$$q_k(x_k) = p_k(x_k) - p_{k-1}(x_k) = a_k(x_k-x_0)(x_k-x_1)\cdots(x_k-x_{k-1})$$

y despejando entonces a_k de ésta última identidad tenemos :

$$a_k = \frac{y_k - p_{k-1}(x_k)}{(x-x_0)(x-x_1)\cdots(x-x_{k-1})}$$

con lo cual podemos poner :

$$p_k(x) = p_{k-1}(x) + q_k(x)$$

donde lo que parece complicado es calcular el a_k, que sería el coeficiente de x^k en el polinomio $p_k(x)$ pero para esto se puede utilizar las diferencias divididas:

Definición 2. *Dada la función f de la cual se conoce su valor en los puntos $x_0, x_1, \ldots, x_k$, se llama* **diferencia dividida de f en los puntos** $x_0, x_1, \ldots, x_k$ *al valor* $a_k = \mathbf{f}[\mathbf{x_0}, \mathbf{x_1}, \cdots, \mathbf{x_k}]$ *y se calcula recursivamente como sigue :*

$$\boxed{\begin{aligned} f[x_i] &= f(x_i) = y_i \\ f[x_i, x_{i+1}] &= \frac{f[x_{i+1}] - f[x_i]}{x_{i+1} - x_i} \\ f[x_i, x_{i+1}, \cdots, x_{i+k}] &= \frac{f[x_{i+1}, x_{i+2}, \cdots, x_{i+k}] - f[x_i, x_{i+1}, \cdots, x_{i+k-1}]}{x_{i+k} - x_i} \end{aligned}}$$

Lema 2.5.1.

$$f[x_i, x_{i+1}, \cdots, x_{i+k}] = \frac{f[x_{i+1}, x_{i+2}, \cdots, x_{i+k}] - f[x_i, x_{i+1}, \cdots, x_{i+k-1}]}{x_{i+k} - x_i}$$

Demostración. Sea $p_j(x)$ el polinomio de grado $\leqslant j$ que coincide con $f(x)$ en los puntos $x_i, x_{i+1}, \ldots, x_{i+j}$ y sea $q_{k-1}(x)$ el polinomio de grado $\leqslant k-1$ que coincide con $f(x)$ en los puntos $x_{i+1}, x_{i+2}, \ldots, x_{i+k}$. Entonces :

$$p(x) = \frac{x - x_i}{x_{i+k} - x_i}\, q_{k-1}(x) + \frac{x_{i+k} - x}{x_{i+k} - x_i}\, p_{k-1}(x)$$

es un polinomio de grado $\leqslant k$ que verifica :

$$p(x_j) = f(x_j), \quad \text{para } j = i, i+1, \ldots, i+k$$

ya que :

$$\begin{aligned}
&\underline{\text{Para } i} : p(x_i) = \frac{x_{i+k} - x_i}{x_{i+k} - x_i}\, p_{k-1}(x_i) = y_i = f(x_i) \\
&\underline{\text{Para } i+k} : p(x_{i+k}) = \frac{x_{i+k} - x_i}{x_{i+k} - x_i}\, q_{k-1}(x_{i+k}) = y_{i+k} = f(x_{i+k}) \\
&\underline{\text{ y para cada } j = i+1, \ldots, i+k-1} : \\
&p(x_j) = \frac{x_j - x_i}{x_{i+k} - x_i}\, q_{k-1}(x_j) + \frac{x_{i+k} - x_j}{x_{i+k} - x_i}\, p_{k-1}(x_j) = \\
&\left[\frac{x_j - x_i}{x_{i+k} - x_i} + \frac{x_{i+k} - x_j}{x_{i+k} - x_i}\right] y_j = \frac{x_{i+k} - x_i}{x_{i+k} - x_i}\, y_j = y_j
\end{aligned}$$

Por tanto, por la unicidad del polinomio de interpolación, tendremos que $p(x) = p_k(x)$ y entonces

$$\begin{aligned}
&f[x_i, x_{i+1}, \cdots, x_{i+k}] = \text{coeficiente término principal de } p_k(x) = \\
&= \frac{\text{coeficiente término principal de } q_{k-1}(x)}{x_{i+k} - x_i} - \frac{\text{coeficiente término principal de } p_{k-1}(x)}{x_{i+k} - x_i} = \\
&= \frac{f[x_{i+1}, x_{i+2}, \cdots, x_{i+k}] - f[x_i, x_{i+1}, \cdots, x_{i+k-1}]}{x_{i+k} - x_i}
\end{aligned}$$

□

¿Cómo organizar el cálculo de la tabla de diferencias divididas?

Ejemplo 14. *El cálculo de las diferencias divididas para cuatro puntos se ordenaría como sigue :*

$$
\begin{array}{lllll}
x_0 \longrightarrow & y_0 = \mathbf{f[x_0]} \searrow & & & \\
 & & \mathbf{f[x_0,x_1]} \searrow & & \\
x_1 \longrightarrow & y_1 = f[x_1] \nearrow & & \mathbf{f[x_0,x_1,x_2]} \searrow & \\
 & & f[x_1,x_2] \nearrow & & \mathbf{f[x_0,x_1,x_2,x_3]} \\
x_2 \longrightarrow & y_2 = f[x_2] & & f[x_1,x_2,x_3] \nearrow & \\
 & & f[x_2,x_3] & & \\
x_3 \longrightarrow & y_3 = f[x_3] & & &
\end{array}
$$

Podemos abordar entonces el cálculo del polinomio de interpolación en los puntos (x_0, y_0), (x_1, y_1),(x_2, y_2),$\ldots$, (x_n, y_n) de la siguiente forma :

$$
\begin{aligned}
p_0(x) &= a_0 = f[x_0] = f(x_0) = y_0 \\
p_1(x) &= p_0(x) + a_1(x - x_0) = f[x_0] + a_1(x - x_0) = f[x_0] + f[x_0, x_1](x - x_0) \\
p_2(x) &= p_1(x) + a_2(x - x_0)(x - x_1) = f[x_0] + a_1(x - x_0) + a_2(x - x_0)(x - x_1) = \\
&= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\
&\vdots \\
p_n(x) &= a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \cdots + a_n(x - x_0)(x - x_1)\ldots(x - x_{n-1}) = \\
p_n(x) &= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \cdots \\
&\quad + f[x_0, x_1, \cdots, x_n](x - x_0)(x - x_1)\ldots(x - x_{n-1})
\end{aligned}
$$

o también de forma más concisa :

$$
\boxed{p_n(x) = \sum_{i=0}^{n} f[x_0, x_1, \ldots, x_i] \prod_{j=0}^{i-1} (x - x_j)}
$$

que se denomina **fórmula de interpolación de Newton en diferencias divididas**.
Para la evaluación del polinomio de interpolación en su forma de Newton en diferencias divididas $p_n(x) = f[x_0]+f[x_0, x_1](x-x_0)+\cdots+f[x_0, x_1, \ldots, x_n](x-x_0)(x-x_1)\cdots(x-x_{n-1})$ usaremos el anidamiento del esquema de Ruffini–Horner :

$$
p_n(z) = \ (\cdots(a_n(z - x_{n-1}) + a_{n-1})(z - x_{n-2}) + \cdots + a_1)(z - x_0) + a_0
$$

para la evaluación en un punto z, y donde se ha puesto $a_k = f[x_0, \ldots, x_k]$.
Obsérvese que se necesitan n productos y 2n? sumas/restas.

Ejemplo 15. *Obtener una fórmula para la suma de los cuadrados de los primeros números naturales.*

Sabemos que $\sum\limits_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$ y como queremos obtenerla por interpolación construimos un conjunto de valores según los diferentes valores de n. Como el polinomio ha de ser el mismo para cualquier posible ordenación de los puntos, elegimos el siguiente orden:

n	**y**	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$	$f[x_i, \ldots, x_{i+3}]$	$f[x_i, \ldots, x_{i+4}]$
$3 \longrightarrow$	**14**				
		9			
$2 \longrightarrow$	5		**23/6**		
		50/3		**1/3**	
$5 \longrightarrow$	55		19/6		**0**
		54/4		1/3	
$1 \longrightarrow$	1		23/6		
		29/3			
$4 \longrightarrow$	30				

El polinomio es por tanto:

$$\begin{aligned} p(x) =& 14 + 9(x-3) + \frac{23}{6}(x-3)(x-2) + \frac{1}{3}(x-3)(x-2)(x-5) = \\ =& \frac{2x^3 + 3x^2 + 6x}{6} = \frac{x(x+1)(2x+1)}{6} \end{aligned} \tag{2.5}$$

como cabría esperar.

2.6 NEWTON EN DIFERENCIAS FINITAS

En el caso particular de que las abcisas de los nodos de interpolación *sean equidistantes* la expresión del polinomio de interpolación de Newton en diferencias divididas adopta otras formas que se han usado mucho, la *fórmula en diferencias progresivas* y la *fórmula en diferencias regresivas.* Antes de desarrollarlas necesitamos de algunas definiciones previas. Dado un conjunto de puntos $(x_i, y_i), 0 \leqslant i \leqslant n$ donde $y_i = f(x_i)$ se define **diferencia progresiva** de orden 1 en y_k y se denota por Δy_k a

$$\Delta y_k = f(x_k + h) - f(x_k) = f(x_{k+1} - f(x_k) = y_{k+1} - y_k = \Delta^1 y_k$$

y **diferencia regresiva** de orden 1 en y_k y se denota por ∇y_k a

$$\nabla y_k = y_k - y_{k-1} = \nabla^1 y_k$$

Análogamente, se define diferencia progresiva de orden 2 en y_k a

$$\begin{aligned}\Delta^2 y_k &= \Delta(\Delta y_k) = \Delta(y_{k+1} - y_k) = \Delta y_{k+1} - \Delta y_k = \\ &= y_{k+2} - y_{k+1} - (y_{k+1} - y_k) = y_{k+2} - 2y_{k+1} + y_k = \nabla^2 y_{k+2}\end{aligned}$$

En general se definen

$$\boxed{\begin{array}{lcl} \Delta^m y_k & \overset{(def)}{=} & \Delta(\Delta^{m-1} y_k) \\ \nabla^m y_k & \overset{(def)}{=} & \nabla(\nabla^{m-1} y_k) \end{array}}$$

y convenimos en que $\Delta^0 y_k = y_k, \nabla^0 y_k = y_k$ y ocurre además que $\Delta^m y_k = \nabla^m y_{k+m}$ y también:

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	Δ^4 y	$\Delta^5 y$	$\Delta^6 y$	$\Delta^7 y$	$\Delta^8 y$
x_0	y_0								
		Δy_0							
x_1	y_1		$\Delta^2 y_0$						
		Δy_1		$\Delta^3 y_0$					
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$				
		Δy_2		$\Delta^3 y_1$		$\Delta^5 y_0$			
x_3	y_3		$\Delta^2 y_2$		$\Delta^4 y_1$		$\Delta^6 y_0$		
		Δy_3		$\Delta^3 y_2$		$\Delta^5 y_1$		$\Delta^7 y_0$	
x_4	y_4		$\Delta^2 y_3$		$\Delta^4 y_2$		$\Delta^6 y_1$		$\Delta^8 y_0$
		Δy_4		$\Delta^3 y_3$		$\Delta^5 y_2$		$\Delta^7 y_1$	
x_5	y_5		$\Delta^2 y_4$		$\Delta^4 y_3$		$\Delta^6 y_2$		
		Δy_5		$\Delta^3 y_4$		$\Delta^5 y_3$			
x_6	y_6		$\Delta^2 y_5$		$\Delta^4 y_4$				
		Δy_6		$\Delta^3 y_5$					
x_7	y_7		$\Delta^2 y_6$						
		Δy_7							
x_8	y_8								

Tabla 2.1: Tabla de diferencias progresivas

$$\Delta^n y_k = \sum_{j=0}^{n} (-1)^j \binom{n}{j} y_{k+n-j}$$

En la tabla 2.1 aparece como se forman las diferencias de diferentes ordenes. Observe la regularidad de la **diagonal superior**.
De la misma forma, en la tabla 2.2 aparecen las mismas diferencias que se calcularon en la tabla de dif. progresivas, pero ahora con la notación propia de las diferencias regresivas. Observe la regularidad de la **diagonal inferior**.

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$	∇^4 y	$\nabla^5 y$	$\nabla^6 y$	$\nabla^7 y$	$\nabla^8 y$
x_0	y_0								
		∇y_1							
x_1	y_1		$\nabla^2 y_2$						
		∇y_2		$\nabla^3 y_3$					
x_2	y_2		$\nabla^2 y_3$		$\nabla^4 y_4$				
		∇y_3		$\nabla^3 y_4$		$\nabla^5 y_5$			
x_3	y_3		$\nabla^2 y_2$		$\nabla^4 y_1$		$\nabla^6 y_0$		
		∇y_4		$\nabla^3 y_5$		$\nabla^5 y_6$		$\nabla^7 y_7$	
x_4	y_4		$\nabla^2 y_5$		$\nabla^4 y_6$		$\nabla^6 y_7$		$\nabla^8 y_8$
		∇y_5		$\nabla^3 y_6$		$\nabla^5 y_7$		$\nabla^7 y_8$	
x_5	y_5		$\nabla^2 y_6$		$\nabla^4 y_7$		$\nabla^6 y_8$		
		∇y_6		$\nabla^3 y_7$		$\nabla^5 y_8$			
x_6	y_6		$\nabla^2 y_7$		$\nabla^4 y_8$				
		∇y_7		$\nabla^3 y_8$					
x_7	y_7		$\nabla^2 y_8$						
		∇y_8							
x_8	y_8								

Tabla 2.2: Tabla de diferencias regresivas

Necesitamos ahora establecer una relación entre diferencias divididas y diferencias finitas para poder reescribir la fórmula de Newton en términos de difrencias progresivas o regresivas. Dicha relación nos viene dada por el siguiente lema.

Lema 2.6.1.

$$\forall i \geqslant 0 : \qquad f[x_k, x_{k+1}, \ldots, x_{k+i}] = \frac{\Delta^i \, y_k}{i! \, h^i}$$

Demostración. Por inducción tenemos :

- Lo demostramos para $i = 0$: $f[x_k] = f(x_k) = y_k = \Delta^0 y_k$
- Lo suponemos cierto para $i = n \geqslant 0$. Y entonces
- Lo demostramos para $i = n + 1$. Tenemos

$$\begin{aligned} f[x_k, \cdots, x_{k+n+1}] &= \frac{f[x_k, \cdots, x_{k+n}] - f[x_{k+1}, \cdots, x_{k+n+1}]}{x_k - x_{k+n+1}} = \\ &= \frac{\Delta^n y_k / n! h^n - \Delta^n y_{k+1} / n! h^n}{-(n+1)h} = \frac{\Delta^n y_{k+1} - \Delta^n y_k}{(n+1)!\, h^{n+1}} = \\ &= \frac{\Delta^n (y_{k+1} - y_k)}{(n+1)!\, h^{n+1}} = \frac{\Delta^n (\Delta y_k)}{(n+1)!\, h^{n+1}} = \frac{\Delta^{n+1} y_k}{(n+1)!\, h^{n+1}} \end{aligned}$$

□

2.7 NEWTON EN DIFERENCIAS PROGRESIVAS

Utilizando éste lema podemos entonces obtener la fórmula de Newton en diferencias progresivas, que es la misma que en diferencias divididas pero expresada en diferencias finitas, que es posible si en los puntos de interpolación las **abcisas son equidistantes**, o sea, si

$$x_{i+1} - x_i = h, \quad \forall i = 0, 1, \ldots, n-1$$

siendo h la diferencia constante entre dos abcisas consecutivas.
Tenemos entonces los puntos de interpolación $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$, donde $y_i = f(x_i)$ y además $x_{i+1} = x_i + h, \forall i = 0, 1, \ldots, n-1$ y podemos escribir $x_1 = x_0 + h \Longrightarrow x_2 = x_1 + h = x_0 + 2h \Longrightarrow x_3 = x_0 + 3h \Longrightarrow \ldots$ $\boxed{x_i = x_0 + i\, h, \quad i = 0, 1, 2, \ldots, n}$
La expresión del polinomio de interpolación en diferencias divididas es

$$\begin{aligned} p_n(x) = f[x_0] &+ f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \cdots \\ &+ f[x_0, x_1, \cdots, x_n](x - x_0)(x - x_1) \ldots (x - x_{n-1}) \end{aligned} \tag{2.6}$$

Aplicando el lema anterior a ésta fórmula obtenemos :

$$\begin{aligned} p_n(x) = y_0 &+ \frac{\Delta^1 y_0}{1!\, h^1} (x - x_0) + \frac{\Delta^2 y_0}{2!\, h^2} (x - x_0)(x - x_1) + \cdots \\ &+ \frac{\Delta^n y_0}{n!\, h^n} (x - x_0)(x - x_1) \ldots (x - x_{n-1}) \end{aligned} \tag{2.7}$$

por otra parte, hacemos el cambio :

$$\frac{x-x_0}{h} = s \Longrightarrow \frac{x-x_i}{h} = \frac{x-(x_0+i\,h)}{h} = \frac{x-x_0-i\,h}{h} = \frac{x-x_0}{h} - \frac{i\,h}{h} = s-i \qquad (2.8)$$

y sustituyendo en (2.7) :

$$\begin{aligned} p_n(x) &= \overbrace{y_0}^{\Delta^0 y_0} + \Delta^1 y_0 \, \frac{(x-x_0)}{h} + \Delta^2 y_0 \, \frac{(x-x_0)(x-x_1)}{2!\,h^2} + \cdots \\ &\quad + \Delta^n y_0 \, \frac{(x-x_0)(x-x_1)\ldots(x-x_{n-1})}{n!\,h^n} = \\ &= y_0 + \Delta^1 y_0 \, \overbrace{s}^{\binom{s}{1}} + \Delta^2 y_0 \, \overbrace{\frac{s(s-1)}{2!}}^{\binom{s}{2}} + \cdots \\ &\quad + \Delta^n y_0 \, \overbrace{\frac{s(s-1)\ldots(s-n+1)}{n!}}^{\binom{s}{n}} = \\ &= \sum_{k=0}^{n} \Delta^k y_0 \, \frac{s(s-1)\ldots(s-k+1)}{k!} = \sum_{k=0}^{n} \Delta^k y_0 \binom{s}{k} \end{aligned} \qquad (2.9)$$

con lo cual tenemos :

$$\boxed{p_n(x) = \sum_{k=0}^{n} \binom{s}{k} \, \Delta^k y_0}$$

que se conoce con el nombre de **fórmula de Newton en diferencias finitas progresivas**.

Ejemplo 16. *Obtener una fórmula para la suma de los primeros números naturales.*

Sabemos que $\sum\limits_{k=1}^{n} k = \frac{n(n+1)}{2}$ y como queremos obtenerla por interpolación con abcisas equidistantes construimos un conjunto de valores según los diferentes valores de n:

n	$\sum$	**y**	Δ^1	Δ^2	Δ^3	Δ^4
$1 \longrightarrow$	1	**1**				
			2			
$2 \longrightarrow$	$1+2$	3		**1**		
			3		**0**	
$3 \longrightarrow$	$1+2+3$	6		1		**0**
			4		0	
$4 \longrightarrow$	$1+2+3+4=$	10		1		
			5			
$1 \longrightarrow$	$1+2+3+4+5=$	15				

Tenemos $p_n(x) = \sum_{k=0}^{n} \binom{s}{k} \Delta^k y_0$
con $s = (x - x_0)/h = x - 1$ con lo que

$$\begin{aligned} p(x) =& 1 + 2(x-1) + \frac{1}{2!}(x-1)(x-2) = \frac{x^2 - 3x + 2 + 4x - 2}{2} = \\ =& \frac{x^2 + x}{2} = \frac{x(x+1)}{2} \end{aligned} \tag{2.10}$$

que es lo que debíamos obtener.

2.8 NEWTON EN DIFERENCIAS REGRESIVAS

También en éste caso las **abcisas son equidistantes**, o sea

$$x_{i+1} - x_i = h, \quad \forall i = 0, 1, \ldots, n-1$$

siendo h la diferencia constante entre dos abcisas consecutivas.
Tenemos entonces los puntos de interpolación $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$, donde $y_i = f(x_i)$ y además $x_{i+1} = x_i + h, \forall i = 0, 1, \ldots, n-1$ y podemos escribir $x_1 = x_0 + h \Longrightarrow x_2 = x_1 + h = x_0 + 2h \Longrightarrow x_3 = x_0 + 3h \Longrightarrow \ldots$ $\boxed{x_i = x_0 + i\,h, \quad i = 0, 1, 2, \ldots, n}$
Consideramos ahora los puntos de interpolación en el orden $(x_n, y_n), (x_{n-1}, y_{n-1}), \ldots, (x_0, y_o)$ y para éste orden la expresión del polinomio de interpolación en diferencias divididas es

$$\begin{aligned} p_n(x) = f[x_n] + f[x_n, x_{n-1}](x - x_n) + f[x_n, x_{n-1}, x_{n-2}](x - x_n)(x - x_{n-1}) + \cdots \\ + f[x_n, x_{n-1}, \cdots, x_0](x - x_n)(x - x_{n-1}) \ldots (x - x_1) \end{aligned} \tag{2.11}$$

Aplicando el lema anterior 2.6.1 a ésta última fórmula y considerando la relación entre diferencias progresivas y regresivas obtenemos :

$$f[x_i, x_{i+1}, \cdots, x_{i+k}] = \frac{\Delta^k y_i}{k!\, h^k} \quad = \frac{\nabla^k y_{i+k}}{k!\, h^k}$$

Considerando además que la diferencia dividida es una función simétrica de sus argumentos, o sea, que

$$f[x_0, x_1, \cdots, x_k] = f[x_{i_0}, x_{i_1}, \cdots, x_{i_k}]$$

para cualquier permutación posible $(i_0, i_1, \ldots, i_k)$ de $(0, 1, \ldots, k)$ tendríamos por ejemplo que
$f[x_n, x_{n-1}] = f[x_{n-1}, x_n]$, $f[x_n, x_{n-1}, x_{n-2}] = f[x_{n-2}, x_{n-1}, x_n]$, etc.
Sustituyendo entonces en la ecuación anterior 2.11:

$$\begin{aligned} p_n(x) = y_n &+ \frac{\nabla^1 y_n}{1!\, h^1}(x - x_n) + \frac{\nabla^2 y_n}{2!\, h^2}(x - x_n)(x - x_{n-1}) + \cdots \\ &+ \frac{\nabla^n y_n}{n!\, h^n}(x - x_n)(x - x_{n-1}) \ldots (x - x_1) \end{aligned} \tag{2.12}$$

por otra parte, hacemos el cambio :

$$\frac{x - x_n}{h} = u \Longrightarrow \begin{array}{l} \frac{x-x_{n-1}}{h} = \frac{x-(x_n-h)}{h} = \frac{x-x_n+h}{h} = \frac{x-x_n}{h} + \frac{h}{h} = u+1 \\ \frac{x-x_{n-2}}{h} = \frac{x-(x_n-2h)}{h} = \frac{x-x_n+2\,h}{h} = \frac{x-x_n}{h} + \frac{2\,h}{h} = u+2 \\ \frac{x-x_i}{h} = \frac{x-(x_n-(n-i)\,h)}{h} = \frac{x-x_n+n-i\,h}{h} = \frac{x-x_n}{h} + \frac{n-i\,h}{h} = u+n-i \end{array}$$

y sustituyendo en (2.12) :

$$\begin{aligned} p_n(x) = \overbrace{y_n}^{\nabla^0 y_n} &+ \nabla^1 y_n \frac{(x - x_n)}{h} + \nabla^2 y_n \frac{(x - x_n)(x - x_{n-1})}{2!\, h^2} + \cdots \\ &+ \nabla^n y_n \frac{(x - x_n)(x - x_{n-1}) \ldots (x - x_1)}{n!\, h^n} = \\ = y_n + \nabla^1 y_n \overbrace{u}^{\binom{u}{1}} &+ \nabla^2 y_n \frac{\overbrace{u(u+1)}^{\binom{u+1}{2}}}{2!} + \cdots \\ &+ \nabla^n y_n \frac{\overbrace{u(u+1)\ldots(u+n-1)}^{\binom{u+n-1}{n}}}{n!} = \\ = \sum_{k=0}^{n} \nabla^k y_n \frac{u(u+1)\ldots(u+k-1)}{k!} &= \sum_{k=0}^{n} \nabla^k y_n \binom{u+k-1}{k} \end{aligned} \tag{2.13}$$

con lo cual tenemos :

$$\boxed{p_n(x) = \sum_{k=0}^{n} \binom{u+k-1}{k} \nabla^k y_n}$$

que se conoce con el nombre de **fórmula de Newton en diferencias regresivas**.

Ejemplo 17. *Obtener una fórmula para la suma de los primeros números naturales.*

La tabla *es exactamente la misma* que se construyó para diferencias progresivas, pero se toma ahora la diagonal inferior, que aparece en negrita

n	$\sum$	**y**	∇^1	∇^2	∇^3	∇^4
$1 \longrightarrow$	$1 =$	1				
			2			
$2 \longrightarrow$	$1+2=$	3		1		
			3		0	
$3 \longrightarrow$	$1+2+3=$	6		1		**0**
			4		**0**	
$4 \longrightarrow$	$1+2+3+4=$	10		**1**		
			5			
$5 \longrightarrow$	$1+2+3+4+5=$	**15**				

Ahora la ecuación es: $p_n(x) = \sum_{k=0}^{n} \binom{u+k-1}{k} \nabla^k y_n$
siendo

$$u = \frac{x - x_n}{h} = x - 5, \quad h = 1$$

$$\begin{aligned} p(x) &= y_n + u\nabla y_n + \frac{u(u+1)}{2!}\nabla^2 y_n = \\ &= 15 + 5(x-5) + \frac{1}{2!}(x-5)(x-4) = \frac{x^2 - 9x + 20 + 10x - 20}{2} = \\ &= \frac{x^2 + x}{2} = \frac{x(x+1)}{2} \end{aligned}$$

que es de nuevo lo que debíamos obtener.

$\longrightarrow$ LIMITACIONES DE LA INTERPOLACIÓN POLINÓMICA.

Ejemplo 18. *Si se toma la función* $f(x) = 1/x$ *en los puntos* $0.1, 0.2, 0.5, 1.0, 2.0, 5.0, 10.0$ *el cálculo de* $f(1.5) = 0.\widehat{6}$ *da sucesivamente :*

$$10, -60, 122, -60, -14.5, -9.95, -8.3575$$

considerando $1, 2, \ldots, 7$ *puntos de interpolación sucesivamente :*

La tabla de diferencias divididas es la siguiente

0.1	10						
		-50					
0.2	5		100				
		-10		-100			
0.5	2		10		50		
		-2		-5		-10	
1.0	1		1		1		1
		-0.5		-0.2		-0.1	
2.0	0.5		0.1		0.02		
		-0.1		-0.01			
5.0	0.2		0.01				
		-0.02					
10.0	0.1						

y los sucesivos valores obtenidos:

$$p_0(1.5) = 10$$
$$p_1(1.5) = 10 - 50(1.5 - 0.1) = -60$$
$$p_2(1.5) = -60 + 100(1.5 - 0.1)(1.5 - 0.2) = 122$$
$$p_3(1.5) = 122 - 100(1.5 - 0.1)(1.5 - 0.2)(1.5 - 0.5) = -60$$
$$p_4(1.5) = -60 + 50(1.5 - 0.1)(1.5 - 0.2)(1.5 - 0.5)(1.5 - 1.0) = -14.5$$
$$p_5(1.5) = -14.5 + (-10)(1.5 - 0.1)(1.5 - 0.2)(1.5 - 0.5)(1.5 - 1.0)(1.5 - 2.0) = -9.95$$
$$p_6(1.5) = -9.95 + 1(1.5 - 0.1)(1.5 - 0.2)(1.5 - 0.5)(1.5 - 1.0)(1.5 - 2.0)(1.5 - 5.0) = -8.3575$$

dejan claro que las diferentes estimaciones que se obtienen por interpolación polinómica son muy pobres.

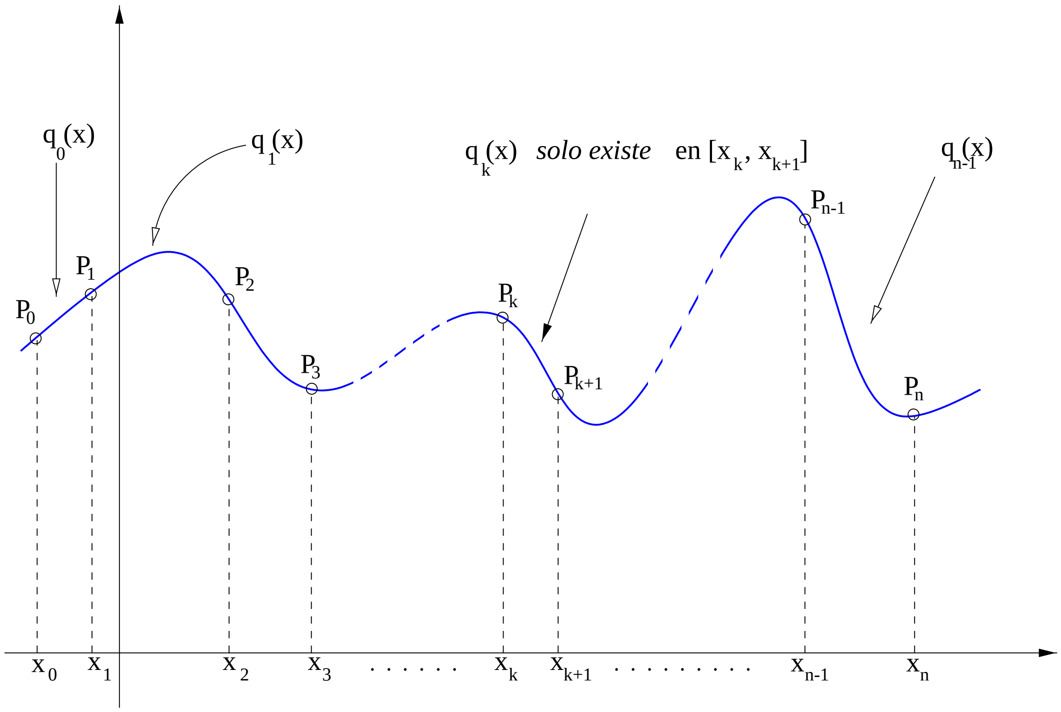

Figura 2.1: Interpolación polinómica a trozos.

2.9 INTERPOLACIÓN POR SPLINES CÚBICOS

Supongamos que tenemos los $n+1$ puntos:

$$P_k(x_k, y_k), \text{ donde } y_k = f(x_k), \quad k = 0, 1, \ldots, n$$

en los cuales se quiere interpolar la función f. Las abcisas no es necesario que sean equidistantes, pero se suponen *ordenados*, o sea,

$$x_0 < x_1 < x_2 < \cdots < x_n$$

En ésta sección trataremos de la interpolación polinómica a trozos. La idea es encontrar polinomios cúbicos $q_k(x)$ que interpolen la función f en el subintervalo $[x_k, x_{k+1}], k = 0, 1, \ldots, n-1$.

Definición 3. *La función $s(x)$ se llama cúbica a trozos en $[x_0, x_n]$ si existen polinomios cúbicos $q_0(x), q_1(x), \ldots, q_{n-1}(x)$ tales que :*

$$s(x) = q_k(x) \quad en \quad [x_k, x_{k+1}], \quad para\ k = 0, 1, \ldots, n-1$$

Para que $s(x)$ interpole en los puntos $P_0, P_1, \dots, P_n$ los $q_k(x)$ han de verificar :

$$\begin{cases} q_k(x_k) = y_k \\ q_k(x_{k+1}) = y_{k+1}, \quad k = 0, 1, \dots, n-1 \end{cases} \tag{2.14}$$

lo cual supone $2n$ condiciones. Llamaremos a $s(x)$ spline cúbico, o simplemente spline, si los polinomios $q_k(x)$ tienen la misma pendiente y la misma concavidad en los nodos que las unen, o sea :

$$\begin{cases} q'_{k-1}(x_k) = q'_k(x_k) \\ q''_{k-1}(x_k) = q''_k(x_k), \quad k = 1, 2, \dots, n-1 \end{cases} \tag{2.15}$$

lo cual supone $2(n-1)$ condiciones a cumplir. Al tener que verificar las condiciones (2.14) y (2.15) se asegura que $s(x)$ tiene su primera y segunda derivadas continuas en $[x_0, x_n]$. En éste caso se dice que $s(x)$ es un spline interpolador para $P_0, P_1, \dots, P_n$.
Si $s(x)$ es cúbica a trozos en el intervalo $[x_0, x_n]$, su derivada segunda $s''(x)$ es lineal en el mismo intervalo e interpola en los puntos $(x_k, s''(x_k))$ y $(x_{k+1}, s''(x_{k+1}))$ en $[x_k, x_{k+1}]$. Por tanto, $q_k(x)$ es un polinomio de grado uno que interpola en los puntos $(x_k, s''(x_k))$ y $(x_{k+1}, s''(x_{k+1}))$:

$$q''_k(x) = s''(x_k)\frac{x - x_{k+1}}{x_k - x_{k+1}} + s''(x_{k+1})\frac{x - x_k}{x_{k+1} - x_k}, \quad \text{para } k = 0, 1, \dots, n-1$$

Denotando con

$$h_k = x_{k+1} - x_k, \quad k = 0, 1, \dots, n-1$$

y

$$\sigma_k = s''(x_k), \quad k = 0, 1, \dots, n$$

tenemos :

$$q''_k(x) = \frac{\sigma_k}{h_k}(x_{k+1} - x) + \frac{\sigma_{k+1}}{h_k}(x - x_k), \quad \text{para } k = 0, 1, \dots, n-1 \tag{2.16}$$

donde h_k y σ_k son constantes (σ_k a determinar). Integrando dos veces :

$$q_k(x) = \frac{\sigma_k}{h_k}\frac{(x_{k+1} - x)^3}{6} + \frac{\sigma_{k+1}}{h_k}\frac{(x - x_k)^3}{6} + C_k + D_k x \tag{2.17}$$

donde el término lineal lo podemos escribir como :

$$C_k + D_k x = A_k(x - x_k) + B_k(x_{k+1} - x)$$

siendo A_k y B_k constantes arbitrarias, quedando entonces :

$$q_k(x) = \frac{\sigma_k}{h_k}\frac{(x_{k+1}-x)^3}{6} + \frac{\sigma_{k+1}}{h_k}\frac{(x-x_k)^3}{6} + A_k(x-x_k) + B_k(x_{k+1}-x) \tag{2.18}$$

Aplicando a (2.18) las condiciones (2.14) :

$$y_k = \frac{\sigma_k}{h_k}\frac{h_k^3}{6} + \frac{\sigma_{k+1}}{h_k}0 + A_k.0 + B_k h_k = \frac{\sigma_k}{6}h_k^2 + B_k h_k \tag{2.19}$$

$$y_{k+1} = \frac{\sigma_{k+1}}{h_k}h_k^3 + A_k h_k = \frac{\sigma_{k+1}}{6}h_k^2 + A_k h_k \tag{2.20}$$

y despejando de aquí A_k y B_k y sustituyendo en (2.18) resulta :

$$\boxed{\begin{aligned} q_k(x) = {} & \frac{\sigma_k}{6}\left[\frac{(x_{k+1}-x)^3}{h_k} - h_k(x_{k+1}-x)\right] \\ & + \frac{\sigma_{k+1}}{6}\left[\frac{(x-x_k)^3}{h_k} - h_k(x-x_k)\right] \\ & + y_k\left[\frac{x_{k+1}-x}{h_k}\right] + y_{k+1}\left[\frac{x-x_k}{h_k}\right], \qquad \text{para } k = 0,1,\ldots,n-1 \end{aligned}} \tag{2.21}$$

que *es la ecuación del spline* $q_k(x)$.

Nos falta aún conocer los valores $\sigma_0, \sigma_1, \ldots, \sigma_n$ (n+1 incógnitas) para lo cual usamos (2.15); derivando en (2.21) tenemos :

$$q_k'(x) = \frac{\sigma_k}{6}\left[\frac{-3(x_{k+1}-x)^2}{h_k} + h_k\right] + \frac{\sigma_{k+1}}{6}\left[\frac{3(x_k-x)^2}{h_k} - h_k\right] + \frac{y_{k+1}-y_k}{h_k}$$

Por tanto :

$$q_k'(x_k) = \frac{\sigma_k}{6}(-2h_k) + \frac{\sigma_{k+1}}{6}(-h_k) + \frac{y_{k+1}-y_k}{h_k} \tag{2.22}$$

$$q_k'(x_{k+1}) = \frac{\sigma_k}{6}(h_k) + \frac{\sigma_{k+1}}{6}(2h_k) + \frac{y_{k+1}-y_k}{h_k} \tag{2.23}$$

Reemplazando k por $k-1$ en (2.23) para obtener $q_{k-1}'(x_k)$ e igualando a (2.22) nos da :

$$\boxed{h_{k-1}\,\sigma_{k-1} + 2\,(h_{k-1}+h_k)\,\sigma_k + h_k\,\sigma_{k+1} = 6\left(\frac{y_{k+1}-y_k}{h_k} - \frac{y_k-y_{k-1}}{h_{k-1}}\right), \quad \text{para } k = 1,2,\ldots,n-1} \tag{2.24}$$

o también

$$\boxed{h_{k-1}\,\sigma_{k-1} + 2\,(h_{k-1}+h_k)\,\sigma_k + h_k\,\sigma_{k+1} = 6\left(\frac{\Delta\, y_k}{h_k} - \frac{\Delta\, y_{k-1}}{h_{k-1}}\right), \quad \text{para } k = 1, 2, \ldots, n-1} \tag{2.25}$$

o también

$$\boxed{h_{k-1}\,\sigma_{k-1} + 2\,(h_{k-1}+h_k)\,\sigma_k + h_k\,\sigma_{k+1} = 6\Big(f[x_k, x_{k+1}] - f[x_{k-1}, x_k]\Big), \quad \text{para } k = 1, 2, \ldots, n-1} \tag{2.26}$$

Como el índice k varía de 1 a $n-1$, se producen $n-1$ ecuaciones lineales con $n+1$ incógnitas $\sigma_0, \sigma_1, \ldots, \sigma_n$, lo cual produce un sistema *subdeterminado* que tiene infinitas soluciones. Existen varias estrategias para eliminar σ_0 de la primera ecuación y σ_n de la (n-1)-ésima produciendo un *sistema tridiagonal* de orden (n-1) en las variables $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}$.

ALTERNATIVA I Especificar el valor de $s''(x)$ en los puntos extremos : $\sigma_0 = s''(x_0)$ y $\sigma_n = s''(x_n)$. Si se pone $\sigma_0 = 0,\ \sigma_n = 0$ se denomina *spline cúbico natural*.

ALTERNATIVA II Suponer que $s''(x)$ es constante en los extremos : $\sigma_0 = \sigma_1$ y $\sigma_n = \sigma_{n-1}$

ALTERNATIVA III Suponer que $s''(x)$ es lineal cerca de los extremos : $\sigma_0 = \frac{1}{h_1}\big((h_0 + h_1)\sigma_1 - h_0\sigma_2\big)$ y

$\sigma_n = \frac{1}{h_{n-2}}\big(-h_{n-1})\sigma_{n-2} + (h_{n-2} + h_{n-1})\sigma_{n-1}\big)$

ALTERNATIVA IV Especificar el valor de $s'(x)$ en los puntos extremos :

$\sigma_0 = \frac{3}{h_0}\big[\Delta y_0 - s'(x_0)\big] - \frac{1}{2}\sigma_1$ y

$\sigma_n = \frac{3}{h_{n-1}}\big[s'(x_n) - \Delta y_{n-1}\big] - \frac{1}{2}\sigma_{n-1}$

Si hay que calcular muchas veces $s(z)$ entonces es preferible hacer la sustitución :

$$x_{k+1} - z = (x_{k+1} - x_k) - (z - x_k) = h_k - (z - x_k)$$

en $q_k(z)$ y entonces expresar éste en potencias de $z - x_k$ para obtener :

$$\begin{aligned} q_k(z) &= y_k + \alpha_1(z - x_k) + \alpha_2(z - x_k)^2 + \alpha_3(z - x_k)^3 \\ &= y_k + (z - x_k)(\alpha_1 + (z - x_k)(\alpha_2 + (z - x_k)\alpha_3)) \end{aligned}$$

evaluado con sólo 4 sumas/restas y 3 productos, donde

$$\alpha_1 = f[x_k, x_{k+1}] - \frac{h_k}{6}(\sigma_{k+1} + 2\sigma_k), \;\; \alpha_2 = \frac{\sigma_k}{2}, \;\; \alpha_3 = \frac{\sigma_{k+1} - \sigma_k}{6h_k}$$

En forma matricial, el sistema tridiagonal que resulta es (caso de spline cúbico natural):

$$\begin{bmatrix} 2\,(h_0 + h_1) & h_1 & \cdots & 0 \\ h_1 & 2\,(h_1 + h_2) & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & 2\,(h_{n-3} + h_{n-2}) & h_{n-2} \\ 0 & 0 & h_{n-2} & 2\,(h_{n-2} + h_{n-1}) \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \vdots \\ \sigma_{n-1} \end{bmatrix} =$$

$$= 6 \begin{bmatrix} f[x_1, x_2] - f[x_0, x_1] \\ f[x_2, x_3] - f[x_1, x_2] \\ f[x_3, x_4] - f[x_2, x_3] \\ \vdots \\ f[x_{n-1}, x_n] - f[x_{n-1}, x_{n-2}] \end{bmatrix}$$

o también

$$\begin{bmatrix} 2\,(h_0 + h_1) & h_1 & \cdots & 0 \\ h_1 & 2\,(h_1 + h_2) & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & 2\,(h_{n-3} + h_{n-2}) & h_{n-2} \\ 0 & 0 & h_{n-2} & 2\,(h_{n-2} + h_{n-1}) \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \vdots \\ \sigma_{n-1} \end{bmatrix} = 6 \begin{bmatrix} \frac{\Delta y_1}{h_1} - \frac{\Delta y_0}{h_0} \\ \frac{\Delta y_2}{h_2} - \frac{\Delta y_1}{h_1} \\ \frac{\Delta y_3}{h_3} - \frac{\Delta y_2}{h_2} \\ \vdots \\ \frac{\Delta y_{n-1}}{h_{n-1}} - \frac{\Delta y_{n-2}}{h_{n-2}} \end{bmatrix}$$

Ejemplo 19. *Interpolar por splines cúbicos la función* $f(x) = 1/x$ *en* $x = 1.5$ *tomando los puntos* $(0.1, 10.0), (0.2, 5.0), (0.5, 2.0), (1.0, 1.0), (2.0, 0.5), (5.0, 0.2), (10.0, 0.1)$.

Solución :

$$\begin{array}{ll} h_0 = 0.2 - 0.1 = 0.1 & h_3 = 2.0 - 1.0 = 1.0 \\ h_1 = 0.5 - 0.2 = 0.3 & h_4 = 5.0 - 2.0 = 3.0 \\ h_2 = 1.0 - 0.5 = 0.5 & h_5 = 10.0 - 5.0 = 5.0 \end{array}$$

El sistema que resulta es

$$0.1\,\sigma_0 + 2\,(0.1+0.3)\,\sigma_1 + 0.3\,\sigma_2 = 6\left(\frac{2-5}{0.5-0.2} - \frac{5-10}{0.2-0.1}\right)$$
$$0.3\,\sigma_1 + 2\,(0.3+0.5)\,\sigma_2 + 0.5\,\sigma_3 = 6\left(\frac{1-2}{1.0-0.5} - \frac{2-5}{0.5-0.2}\right)$$
$$0.5\,\sigma_2 + 2\,(0.5+1.0)\,\sigma_3 + 1.0\,\sigma_4 = 6\left(\frac{0.5-1.0}{2.0-1.0} - \frac{1-2}{1.0-0.5}\right)$$
$$1.0\,\sigma_3 + 2\,(1.0+3.0)\,\sigma_4 + 3.0\,\sigma_5 = 6\left(\frac{0.2-0.5}{5.0-2.0} - \frac{0.5-1.0}{2.0-1.0}\right)$$
$$3.0\,\sigma_4 + 2\,(3.0+5.0)\,\sigma_5 + 5.0\,\sigma_6 = 6\left(\frac{0.1-0.2}{1.0-5.0} - \frac{0.2-0.5}{5.0-2.0}\right)$$

Poniendo $\sigma_0 = \sigma_6 = 0$ tenemos

$$\begin{bmatrix} 0.8 & 0.3 & 0 & 0 & 0 \\ 0.3 & 1.6 & 0.5 & 0 & 0 \\ 0 & 0.5 & 3.0 & 1.0 & 0 \\ 0 & 0 & 1.0 & 8.0 & 3.0 \\ 0 & 0 & 0 & 3.0 & 16.0 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \end{bmatrix} = 6 \begin{bmatrix} 40 \\ 8 \\ 1.5 \\ 0.4 \\ 0.08 \end{bmatrix}$$

de donde se obtienen

$$\sigma_1 = 311.65398570643 \qquad \sigma_2 = -31.077295217152$$
$$\sigma_3 = 8.4549532710280 \qquad \sigma_4 = -.82621220450797$$
$$\sigma_5 = 0.18491478834524$$

Para $x = 1.5$ habrá que elegir $q_3(x)$

$$q_3(x) = \frac{\sigma_3}{6}\left[\frac{(2.0-x)^3}{1.0} - 1.0\,(2.0-x)\right] + \frac{\sigma_4}{6}\left[\frac{(x-1.0)^3}{1.0} - 1.0\,(x-1.0)\right] + 1.0(2.0-x) + 0.5(x-1.0)$$

y su valor $q_3(x) = 0.27320367097855$ es una mejor estimación que la obtenida por interpolación polinómica (ver figura 2.2).

Los diferentes splines que resultan son:

$$q_0(x) = 519.423309x^3 - 155.826993x^2 - 39.611534x + 15$$
$$q_1(x) = -190.406267x^3 + 270.070753x^2 - 124.791083x + 20.678637$$
$$q_2(x) = 13.177416x^3 - 35.304772x^2 + 27.896679x - 4.769324$$
$$q_3(x) = -1.546861x^3 + 8.868059x^2 - 16.276152x + 9.954953$$
$$q_4(x) = +0.0561737x^3 - 0.750148x^2 + 2.960264x - 2.869324$$
$$q_5(x) = -0.00616383x^3 + 0.1849148x^2 - 1.715052x + 4.922870$$

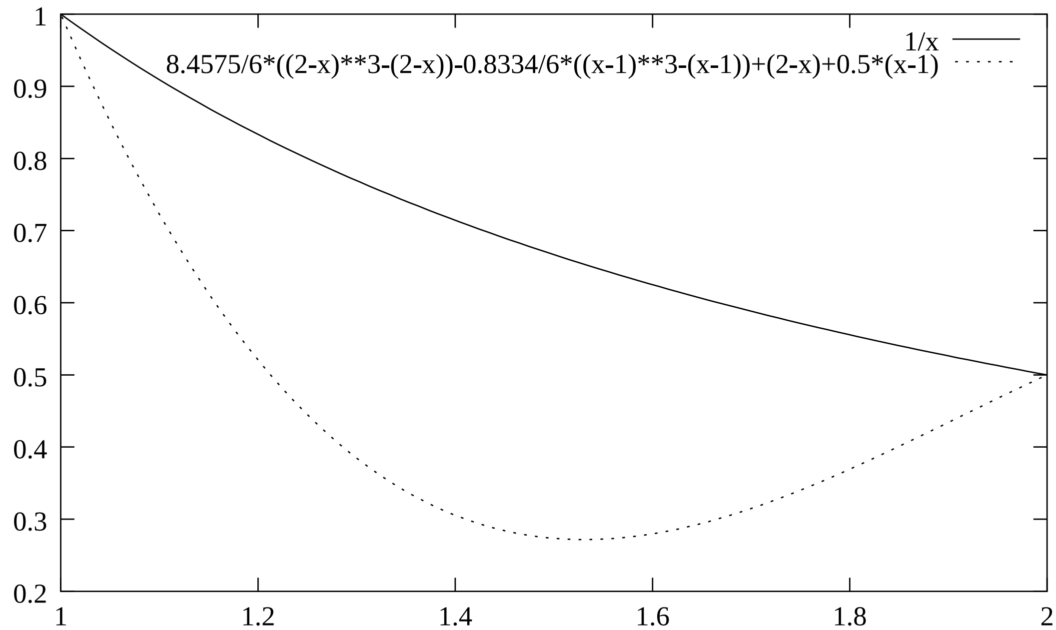

Figura 2.2: La función $1/x$ y $q_3(x)$ en el intervalo [1.0,2.0].

$$s(x) = \begin{cases} q_0(x)\,, & \text{si } x \in [0.1, 0.2], \quad (\text{también } (-\infty, 0.2]) \\ q_1(x)\,, & \text{si } x \in [0.2, 0.5], \\ q_2(x)\,, & \text{si } x \in [0.5, 1.0], \\ q_3(x)\,, & \text{si } x \in [1.0, 2.0], \\ q_4(x)\,, & \text{si } x \in [2.0, 5.0], \\ q_5(x)\,, & \text{si } x \in [5.0, 10.0], \quad (\text{también } [5.0, +\infty)) \end{cases}$$

En el intervalo $[1.0, 2.0]$, la representación de ambas funciones es la que aparece en la figura 2.2.

Ejemplo 20. *Interpolar por splines cúbicos la función $f(x) = 1/(x^2 + 1)$ en el intervalo $0 \leqslant x \leqslant 1$ tomando los seis puntos de abcisas $x_k = k/5,\ k = 0, 1, 2, 3, 4, 5.$*

Por cálculo directo tenemos :

$$\begin{array}{ll} y_0 = 1.00000000 & y_3 = 0.73529412 \\ y_1 = 0.96153846 & y_4 = 0.60975610 \\ y_2 = 0.86206896 & y_5 = 0.50000000 \end{array}$$

y es $h_0 = h_1 = h_2 = h_3 = h_4 = 1/5 = h$ y poniendo $\sigma_0 = \sigma_5 = 0$ y multiplicando ambas partes por $1/h$:

$$\begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \end{bmatrix} = \frac{6}{h^2} \begin{bmatrix} y_2 - 2\,y_1 + y_0 \\ y_3 - 2\,y_2 + y_1 \\ y_4 - 2\,y_3 + y_2 \\ y_5 - 2\,y_4 + y_3 \end{bmatrix} = \begin{bmatrix} -9.151194 \\ -4.095801 \\ 0.185523 \\ 2.367288 \end{bmatrix}$$

y resolviendo $\sigma_1 = -2.165814$, $\sigma_2 = -0.487920$, $\sigma_3 = 0.022536$, $\sigma_4 = 0.581866$ y tabulando la función entre 0 y 1.0 con paso 0.002 la gráfica de f(x) y el spline cúbico son indistinguibles (error máximo $\simeq 0.0040$ que se produce entre 0 y 0.2).

2.10 AJUSTE DE FUNCIONES. APROXIMACIÓN MÍNIMO CUADRÁTICA

Sean (x_i, y_i), $\quad i = 1, 2, \ldots, n$ una nube de puntos con pesos $w_i \geqslant 0$ y sea $\{g_j\}_{j=1}^m$ una familia de funciones básicas.
Si llamamos

$$\phi(x) = c_1 g_1(x) + c_2 g_2(x) + \cdots + c_m g_m(x) = \sum_{j=1}^{m} c_j g_j(x)$$

donde las c_j están por determinar, se trata de hacer mínimo el error cuadrático :

$$e = \sum_{i=1}^{n} w_i (\phi(x_i) - y_i)^2 \tag{2.27}$$

siendo $y_i = f(x_i)$, f desconocida.
Por tanto, una condición necesaria de existencia de mínimo para e es :

$$\boxed{\frac{\partial e}{\partial c_k} = 0, \quad \text{para} \quad k = 1, 2, \ldots, m} \tag{2.28}$$

que se denomina sistema de ecuaciones normales y es en definitiva un sistema de ecuaciones lineales com m incógnitas c_j.
Desarrollando (2.27) tenemos :

$$\begin{aligned} e =& w_1(\phi(x_1) - y_1)^2 + w_2(\phi(x_2) - y_2)^2 + \cdots + w_n(\phi(x_n) - y_n)^2 = \\ & w_1(c_1 g_1(x_1) + \cdots + c_m g_m(x_1) - y_1)^2 + \cdots + w_n(c_1 g_1(x_n) + \cdots + c_m g_m(x_n) - y_n)^2 \end{aligned}$$

Por tanto :

$$\begin{aligned}\frac{\partial e}{\partial c_k} &= 2w_1(\phi(x_1)-y_1)g_1(x_1)+2w_2(\phi(x_2)-y_2)g_k(x_2)+\cdots+2w_n(\phi(x_n)-y_n)g_k(x_n) = \\ &= 2\sum_{i=1}^{n} w_i(\phi(x_i)-y_i)g_k(x_i)=0, \quad k=1,2,\ldots,n\end{aligned} \tag{2.29}$$

de la cual se puede eliminar el 2 ya que está igualado a cero. O sea :

$$w_1(c_1g_1(x_1)+\cdots+c_mg_m(x_1)-y_1)g_k(x_1)+\cdots+w_n(c_1g_1(x_n)+\cdots+c_mg_m(x_n)-y_n)g_k(x_n)=0$$

Reordenando respecto a las c_k y pasando al segundo miembro los términos con signo negativo tenemos :

$$\begin{aligned}&= c_1(w_1g_1(x_1)g_k(x_1)+\cdots+w_ng_1(x_n)g_k(x_n))+\cdots+c_m(w_1g_m(x_1)g_k(x1)+\cdots+w_ng_m(x_n)g_k(x_n))\\ &= w_1y_1g_k(x_1)+\cdots+w_ny_ng_k(x_n)\end{aligned}$$

donde ya se han pasado al segundo miembro los sumandos con un signo negativo en la expresión anterior. Por tanto, queda :

$$\sum_{j=1}^{m} c_j\Big(\sum_{i=1}^{n} w_ig_j(x_i)g_k(x_i)\Big)=\sum_{i=1}^{n} w_iy_ig_k(x_i)$$

para $k=1,2,\ldots,m$. O también en forma matricial :

$$\begin{bmatrix} \sum\limits_{i=1}^{n} w_ig_1(x_i)g_1(x_i) & \sum\limits_{i=1}^{n} w_ig_2(x_i)g_1(x_i) & \cdots & \sum\limits_{i=1}^{n} w_ig_m(x_i)g_1(x_i) \\ \sum\limits_{i=1}^{n} w_ig_1(x_i)g_2(x_i) & \sum\limits_{i=1}^{n} w_ig_2(x_i)g_2(x_i) & \cdots & \sum\limits_{i=1}^{n} w_ig_m(x_i)g_2(x_i) \\ \vdots & \vdots & \ddots & \vdots \\ \sum\limits_{i=1}^{n} w_ig_1(x_i)g_m(x_i) & \sum\limits_{i=1}^{n} w_ig_2(x_i)g_m(x_i) & \cdots & \sum\limits_{i=1}^{n} w_ig_m(x_i)g_m(x_i) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} \sum\limits_{i=1}^{n} w_iy_ig_1(x_i) \\ \sum\limits_{i=1}^{n} w_iy_ig_2(x_i) \\ \vdots \\ \sum\limits_{i=1}^{n} w_iy_ig_m(x_i) \end{bmatrix}$$

que se denomina **sistema de ecuaciones normales** para la aproximación por mínimos cuadrados. De aquí se obtienen $c_1, c_2, \ldots, c_m$. Nótese que la matriz es simétrica, y definida positiva.

Un caso particular importante es cuando se quiere ajustar los datos a un **polinomio** de un determinado grado, o sea, cuando :

$$\phi(x) = p_n(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_m x^{m-1} = \sum_{j=0}^{m-1} \alpha_j x^j$$

En éste caso es $g_j(x) = x^{j-1}$ y por tanto $g_j(x_i)g_k(x_i) = x_i^{j+k-2}$ y el sistema normal queda de la siguiente forma :

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{m-1} \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^m \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \cdots & \sum x_i^{m+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum x_i^{m-1} & \sum x_i^m & \sum x_i^{m+1} & \cdots & \sum x_i^{2m-2} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{m-1} \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i x_i^2 \\ \vdots \\ \sum y_i x_i^{m-1} \end{bmatrix}$$

Ejemplo 21. *Dados los puntos* $P_1(1,6), P_2(2,1), P_3(4,2), P_4(5,3), P_5(10,4), P_6(16,5)$ *encontrar los polinomios de grados 1,2,3,4,5 y 6 que mejor se ajusten por mínimos cuadrados.*

Para el caso de grado uno tenemos $p_1(x) = \alpha_0 + \alpha_1 x$ y entonces :

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \end{bmatrix}$$

Y sustituyendo :

$$\begin{bmatrix} 6 & 38 \\ 38 & 402 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} 21 \\ 151 \end{bmatrix}$$

cuya solución nos da la **recta de regresión** $p_1(x) = 2.793 + 0.1116x$.
Análogamente, para el caso de grado dos tenemos:

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i x_i^2 \end{bmatrix}$$

Y sustituyendo :

$$\begin{bmatrix} 6 & 38 & 402 \\ 38 & 402 & 5294 \\ 402 & 5294 & 76434 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 21 \\ 151 \\ 1797 \end{bmatrix}$$

cuya solución nos da $p_2(x) = 4.0593 - 0.4159x + 0.03097x^2$.

$$\begin{bmatrix} 6 & 38 & 402 & 5294 \\ 38 & 402 & 5294 & 76434 \\ 402 & 5294 & 76434 & 1152758 \\ 5294 & 76434 & 1152758 & 17797002 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 21 \\ 151 \\ 1797 \\ 24997 \end{bmatrix}$$

cuya solución nos da $p_3(x) = 6.4822 - 2.1813x + 0.3075x^2 - 0.01106x^3$.
El resto de los polinomios resultantes del ajuste son:

$$\begin{aligned} p_3(x) &= 6.482 - 2.181x + 0.3075x^2 - 0.011076x^3 \\ p_4(x) &= 15.43 - 12.75x + 3.494x^2 - 0.33714x^3 + 0.010375x^4 \\ p_5(x) &= 20.72 - 21.74x + 8.28x^2 - 1.3582x^3 + 0.098876x^4 - 0.002579x^5 \\ &\quad \text{con error cuadrático de } 0.182823 \qquad \times 10^{-17}. \\ p_6(x) &= 13.2005 - 5.8561x - 3.3569x^2 + 2.4861x^3 - 0.5130x^4 + 0.04205x^5 - 0.001174x^5 \\ &\quad \text{con error cuadrático de } 0.229023 \times 10^{-19}. \end{aligned}$$

En algunas ocasiones la función de ajuste no aparece expresada como una combinación lineal de funciones básicas, debiéndose realizar previamente algunas transformaciones con el fin de obtener una ecuación transformada que sea lineal, como es el caso del siguiente ejemplo.

Ejemplo 22.

La intensidad de la radiación emitida de una fuente radiactiva viene dada por la fórmula $I = I_0 e^{-\alpha t}$. *Determinar las constantes* α *e* I_0 *usando los datos dados :*

t	0.2	0.3	0.4	0.5	0.6	0.7	0.8
I	3.16	2.38	1.75	1.34	1.00	0.74	0.56

Tomando logaritmos neperianos, tenemos :

$$I = I_0 e^{-\alpha t} \Longrightarrow \ln I = \ln I_0 - \alpha t \Longrightarrow Y = A + Bt$$

con $Y = \ln I$, $A = \ln I_0$ y $B = -\alpha$, con lo que hay que añadir una línea más en la tabla, la correspondiente a $\ln I$ que nos ha aparecido en la linealización :

$\ln \mathbf{I}$	1.1506	0.8671	0.5596	0.2927	0	−0.3011	−0.5798

El sistema que resulta es pues :

$$\begin{bmatrix} \sum 1 & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i t_i \end{bmatrix}$$

O sea :

$$\begin{bmatrix} 7 & 3.5 \\ 3.5 & 2.03 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 1.9891 \\ 0.18583 \end{bmatrix}$$

Y de aquí :

$$A = \frac{3.387468}{1.96} = 1.7283 \Longrightarrow I_0 = e^A = 5.631073 \simeq 5.63$$
$$B = \frac{-5.66104}{1.96} = -2.8883 \Longrightarrow \alpha = -B = 2.8883 \simeq 2.89$$

siendo el error cuadrático cometido de 0.366075×10^{-3}.
Otro ejemplo que supone linealización es:

Ejemplo 23. *Dada la siguiente nube de puntos :*

$\mathbf{x_i}$	0.1	0.2	0.3	0.4	0.5	0.6
$\mathbf{y_i}$	0.4204	0.4460	0.4726	0.5027	0.5378	0.5792

encontrar la función del tipo $y = \alpha\ 2^{\beta x^2 + \gamma\sqrt{x} - 1}$ *que mejor se ajuste por mínimos cuadrados. Estimar asimismo el error cuadrático cometido.*

Tenemos $\log_2 y = \log_2 \alpha + \beta x^2 + \gamma\sqrt{x} - 1 \Longrightarrow \underbrace{\log_2 y}_{Y} = \underbrace{(\log_2 \alpha - 1)}_{A} + \underbrace{\beta}_{B}\, x^2 + \underbrace{\gamma}_{C}\, \sqrt{x}$
o sea, $Y = A + Bx^2 + C\sqrt{x}$, con funciones básicas $g_1(x) = 1, \quad g_2(x) = x^2, \quad g_3(x) = \sqrt{x}$.
Construyendo entonces la lineas necesarias:

$\mathbf{x_i}$	0.1	0.2	0.3	0.4	0.5	0.6
$\mathbf{y_i}$	0.4204	0.4460	0.4726	0.5027	0.5378	0.5792
$\mathbf{Y} = \log_2 \mathbf{y_i}$	-1.25016506	-1.164884	-1.081308	-0.992230	-0.894858	-0.787866
$\sqrt{\mathbf{x_i}}$	0.316228	0.447214	0.547723	0.632456	0.707107	0.774597
$\mathbf{x_i^2}$	0.01	0.04	0.09	0.16	0.25	0.36

El sistema que resulta es:

$$\begin{bmatrix} \sum_{i=1}^{6} 1 & \sum_{i=1}^{6} x_i^2 & \sum_{i=1}^{6} \sqrt{x_i} \\ \sum_{i=1}^{6} x_i^2 & \sum_{i=1}^{6} x_i^4 & \sum_{i=1}^{6} x_i^2\sqrt{x_i} \\ \sum_{i=1}^{6} \sqrt{x_i} & \sum_{i=1}^{6} x_i^2\sqrt{x_i} & \sum_{i=1}^{6} x_i \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{6} y_i \\ \sum_{i=1}^{6} y_i x_i^2 \\ \sum_{i=1}^{6} y_i\sqrt{x_i} \end{bmatrix}$$

Y sustituyendo :

$$\begin{bmatrix} 6 & 0.91 & 3.425325 \\ 0.91 & 0.2275 & 0.6271071 \\ 3.425325 & 0.627171 & 2.1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} -6.171311 \\ -0.822518 \\ -3.379128 \end{bmatrix} \text{(o también} \begin{bmatrix} -4.277627 \\ -0.570126 \\ -2.342233 \end{bmatrix} \text{si es ln)}$$

con lo que queda:
$A = -1.414679 = \log_2 \alpha - 1 \Longrightarrow \alpha = s^{-0.414679} \simeq 0.750186$ y por otra parte $B = \beta = 0.667617, C = \gamma = 0.498999$ siendo entonces la función :

$$\boxed{y = 0.7501862^{0.667617x^2+0.498999\sqrt{x}-1}}$$

siendo el error cuadrático 0.161628×10^{-7} en éste caso.

Bibliografía

[Act90] F.S. Acton. *Numerical Methods That (Usually) Work*. The Mathematical Association of America, Washington, 1990.

[Atk89] K. E. Atkinson. *An Introduction to Numerical Analysis*. John Wiley, New York, 2nd. edition, 1989.

[BF80] R.L. Burden and D. Faires. *Análisis Numérico*. Grupo Editorial Iberoamericana, México, 1980.

[CC89] S.C. Chapra and R.P. Canale. *Numerical Methods for Engineers*. McGraw-Hill International, New York, second edition, 1989.

[CdB80] S. D. Conte and C. de Boor. *Elementary Numerical Analysis: An Algorithmic Approach*. McGraw–Hill, New York, third edition, 1980.

[DB74] Germund Dahlquist and Åke Björck. *Numerical Methods*. Prentice-Hall, Englewood Cliffs, New Jersey, 1974.

[Frö79] C.-E. Fröberg. *Introduction to Numerical Analysis*. Adison–Wesley, Reading, Massachusetts, 2nd. edition, 1979.

[GW89] C.F. Gerald and P.O. Wheatley. *Applied Numerical Analysis*. Addison–Wesley Publishing Co., Reading, Massachusets, fourth edition, 1989.

[Hen72] P. Henrici. *Elementos de Análisis Numérico*. Ed. Trillas, México, 1972.

[Hil74] F. B. Hildebrand. *Introduction to Numerical Analysis*. McGraw–Hill, New York, second edition, 1974.

[KC94] D. Kincaid and W. Cheney. *Análisis Numérico : las matemáticas del cálculo científico*. Addison-Wesley Iberoamericana, 1994.

[ML91] M. J. Maron and R. J. Lopez. *Numerical Analysis: A Practical Approach*. Wadsworth, Belmont, California, third edition, 1991.

[RR78] Anthony Ralston and Philip Rabinowitz. *A First Course in Numerical Analysis*. McGraw-Hill, New York, 2nd. edition, 1978.

[Sch89] H.R. Schwarz. *Numerical Analysis*. John Wiley & Sons, Chichester, 1989.

[YG73a] David M. Young and R.T. Gregory. *A Survey of Numerical Mathematics*, volume I. Dover Publications, New York, 1973.

[YG73b] David M. Young and R.T. Gregory. *A Survey of Numerical Mathematics*, volume II. Dover Publications, New York, 1973.

Capítulo 3

DERIVACIÓN E INTEGRACIÓN NUMÉRICAS

3.1 DERIVACIÓN NUMÉRICA

3.1.1 Introducción

Las fórmulas que surgen lógicamente para aproximar la derivada de una función en un punto a cuando se conoce el valor de la función en varios puntos son las fórmulas de tipo interpolatorio.
Vamos a verlas en los casos más usuales .
El caso más sencillo será aquel en que se usa solamente el valor de la función en un punto x_0. La fórmula es por tanto del tipo :

$$f'(a) \simeq a_0 \, f(x_0)$$

Para escribir el segundo miembro debemos hallar el polinomio de interpolación de f en x_0 y derivarlo. Como dicho polinomio es $f(x_0)$ constante, al derivarlo resulta cero. Por tanto, la fórmula anterior de tipo interpolatorio, exacta para los polinomios de grado 0 es :

$$f'(a) \simeq 0$$

que tiene poco interés y con error $R(f) = f'(a)$ naturalmente.
Supondremos que f es de clase C^n (para un cierto n que dependerá del caso) en un intervalo que contenga el valor a.

3.2 FÓRMULAS CON DOS PUNTOS

El caso en que se utilizan dos valores de f : $f(x_0)$ y $f(x_1)$, se construye el polinomio de interpolación de f en x_0 y x_1 y se deriva en el punto a

$$\begin{array}{lll} x_0 \longrightarrow & f(x_0) \searrow & \\ & & f[x_0,x_1] \Longrightarrow p(x) = f(x_0) + f[x_0,x_1](x-x_0) \\ x_1 \longrightarrow & f(x_1) \nearrow & \end{array}$$

con lo que

$$p'(x) = f[x_0,x_1] = \frac{f(x_1)-f(x_0)}{x_1-x_0}$$

Si se quiere aproximar $f'(a)$ tenemos :

- Eligiendo $x_0 = a, x_1 = a+h$:

$$f'(a) \simeq \frac{f(a+h)-f(a)}{h}$$

y el error cometido sería, aplicando Taylor :

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \frac{h^3}{3!}f'''(a) + \cdots$$

con lo que :

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

y de aquí :

$$f'(a) - \frac{f(a+h)-f(a)}{h} = \overbrace{-\frac{h}{2!}f''(a) - \frac{h^2}{3!}f'''(a) + \cdots}^{\text{Error de truncamiento, potencias } h^k}$$

o bien:

$$f'(a) - \frac{f(a+h)-f(a)}{h} = -h/2f''(\xi), \quad \text{donde } \xi \in [a, a+h]$$

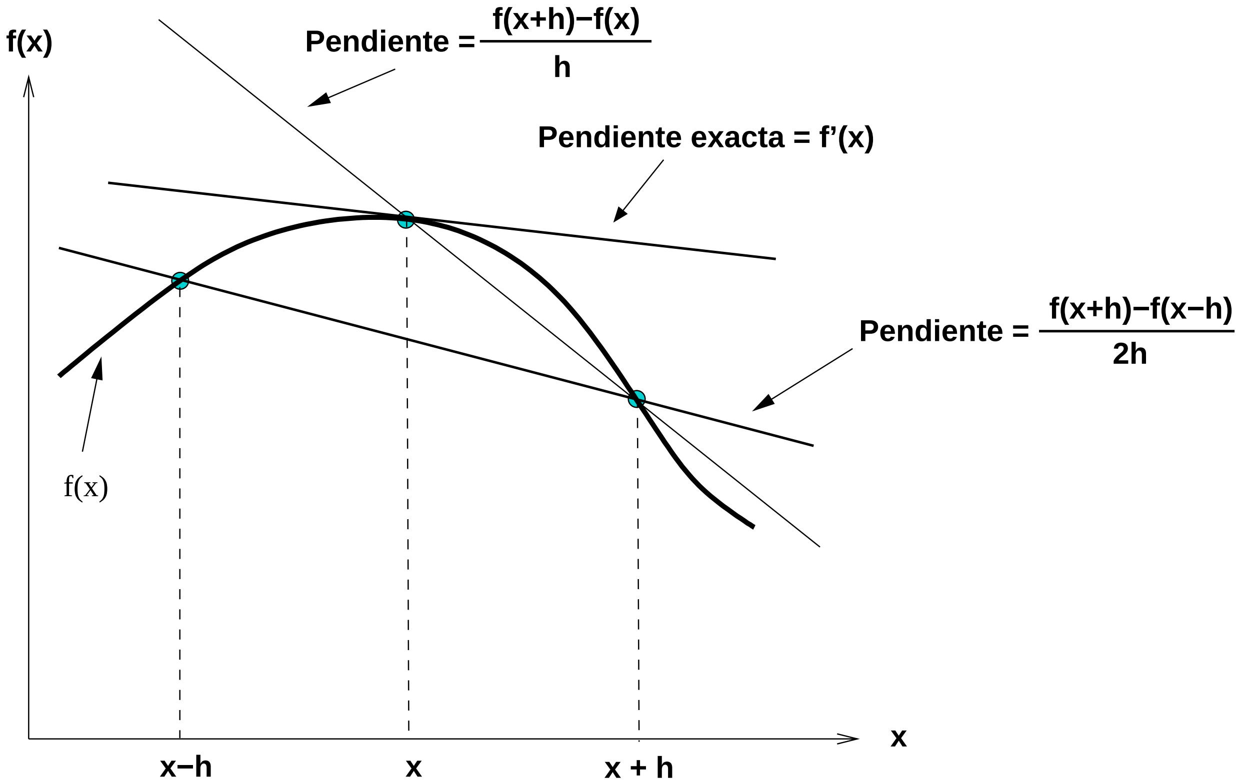

Figura 3.1: Derivada y aproximaciones.

- Eligiendo $x_1 = a + h, x_0 = a - h$:

$$f'(a) \simeq \frac{f(a+h) - f(a-h)}{2h}$$

y el error cometido en éste caso sería, aplicando de nuevo Taylor :

$$f(a-h) = f(a) - hf'(a) + \frac{h^2}{2!}f''(a) - \frac{h^3}{3!}f'''(a) + \cdots$$

con lo que :

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

y de aquí :

$$f'(a) = \frac{f(a+h) - f(a-h)}{2h} \overbrace{- \frac{h^2}{3!} f'''(a) - \frac{h^4}{5!} f^{(5)}(a) - \frac{h^6}{7!} f^{(7)}(a) - \cdots}^{\text{Error de truncamiento, potencias } h^{2k}}$$

o bien:

$$f'(a) - \frac{f(a+h) - f(a-h)}{2h} = -h^2/6 f'''(\xi), \quad \text{donde } \xi \in [a-h, a+h]$$

Como el error en la fórmula en diferencias centrales es de orden $O(h^2)$, ésta es más precisa que la fórmula que utiliza diferencias progresivas. En la figura [3.1] se aprecian las diferentes pendientes obtenidas por las diferentes estimaciones de la derivada.

3.3 FÓRMULAS CON TRES PUNTOS

En éste caso es $p(x) = f(x_0) + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1)$ y entonces:

$$\begin{aligned} p'(x) &= f[x_0, x_1] + f[x_0, x_1, x_2](2x - x_0 - x_1) \\ &= \frac{f(x_1) - f(x_0)}{x_1 - x_0} + \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} (2x - x_0 - x_1) \\ &= \frac{f(x_1) - f(x_0)}{x_1 - x_0} + \frac{\frac{f(x_2) - f(x_1)}{x_2 - x_1} - \frac{f(x_1) - f(x_0)}{x_1 - x_0}}{x_2 - x_0} (2x - x_0 - x_1) \end{aligned}$$

Si se quiere aproximar $f'(a)$ tenemos :

- Eligiendo entonces $x_0 = a, x_1 = a + h, x_2 = a + 2h$ (derivación en un extremo del intervalo) :

$$f'(a) = \frac{-f(a+2h) + 4f(a+h) - 3f(a)}{2h} + \overbrace{\frac{h^2}{3} f'''(a) + \frac{h^3}{4} f^{iv}(a) + \cdots}$$

- Eligiendo $x_0 = a - h, x_1 = a, x_2 = a + h$ (derivación en el centro del intervalo) :

$$f'(a) = \frac{f(a+h) - f(a-h)}{2h} \overbrace{- \frac{h^2}{3!} f^{(3)}(a) - \frac{h^4}{5!} f^{(5)}(a) - \frac{h^6}{7!} f^{(7)}(a) - \cdots}$$

obteniéndose de nuevo la fórmula en diferencias centrales y siendo el error el mismo que antes, por supuesto.

3.4 FÓRMULAS CON CUATRO PUNTOS

Tomando $x_0 = a - 2h, x_1 = a - h, x_2 = a + h, x_3 = a + 2h$ (derivación en el centro del intervalo) :

$$f'(a) = \frac{-f(a+2h) + 8f(a+h) - 8f(a-h) + f(a-2h)}{12h} + \overbrace{\frac{h^4}{30} f^{(4)}(a) + \cdots}$$

3.5 DERIVADAS DE ORDEN SUPERIOR

No hay dificultad en extender a la aproximación de $f^k(a)$ el procedimiento usado para la obtención de aproximaciones de $f'(a)$. Sin embargo, en primer lugar, hay que tener en cuenta que si para la aproximación de $f^k(a)$ vamos a usar n puntos y es $n \leqslant k$, el polinomio de interpolación será de grado $n-1$ a lo más y al derivarlo k veces para obtener $p^{(k)}(a)$ resultará el valor cero.
Por tanto, $n > k$ forzosamente. En particular, el caso más interesante $k = 2$ usando tres puntos x_0, x_1, x_2 y así obtener $f''(a)$.
Partiendo de nuevo de la expresión conocida del polinomio de interpolación de Newton en los tres puntos : $p(x) = f(x_0) + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1)$ y derivando dos veces obtenemos :

$$p''(x) = 2f[x_0, x_1, x_2]$$

- Tomando $x_0 = a - h, x_1 = a, x_2 = a + h$ (derivación en el centro del intervalo) :

$$\begin{aligned} f''(a) &\simeq 2f[a-h, a, a+h] = 2\frac{\frac{f(a+h)-f(a)}{h} - \frac{f(a)-f(a-h)}{h}}{2h} \\ &= \frac{f(a+h) - 2f(a) + f(a-h)}{h^2} \end{aligned}$$

  siendo el error :

$$R(f) = -\frac{h^2}{12} f^{(4)}(a) + \cdots$$

- Eligiendo $x_0 = a, x_1 = a + h, x_2 = a + 2h$ (derivación en el extremo del intervalo) :

$$f''(a) = \frac{f(a+2h) - 2f(a+h) + f(a)}{h^2} + \overbrace{-hf'''(\xi) + \frac{h^2}{6} f^{iv}(\eta)}$$

siendo $\xi, \eta \in]a, a+2h[$.

Ejemplo 24. *Suponiendo los siguientes datos empíricos, queremos encontrar diferentes estimaciones del valor de $f'(0.5)$:*

$\mathbf{x_i}$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\mathbf{f(x_i)}$	5.1234	5.3057	5.5687	5.9378	6.4370	7.0978	7.9493	9.0253	10.3627

(Nota : en realidad los valores dados en la tabla son perturbaciones de los valores del polinomio $f(x) = x^4 + 3x^3 + 2x^2 + x + 5$ que se pueden considerar que surgen de errores de observación, de redondeo o ambos. El valor exacto es $f'(0.5) = 5.75$)
Tenemos un paso $h = 0.1$ y utilizando la fórmula central para dos puntos :

$$f'(0.5) \simeq \frac{f(a+h) - f(a-h)}{2h} \overset{(a=0.5)}{=} \frac{f(0.6) - f(0.4)}{0.2} = 5.8$$

Eligiendo la fórmula central con cuatro puntos :

$$\begin{aligned} f'(0.5) &\simeq \frac{1}{12h}[f(a-2h) - 8f(a-h) + 8f(a+h) - f(a+2h)] \\ &= \frac{f(0.3) - 8f(0.4) + 8f(0.6) - f(0.7)}{1.2} = 5.7495 \end{aligned}$$

Si se ajustan los datos a un polinomio de grado cuatro obtendríamos :

$$y(x) = 0.9988x^4 + 2.9899x^3 + 2.0172x^2 + 0.9920x + 5.0010$$

que derivado y evaluado en $x = 0.5$ nos da $y'(0.5) = 5.7510$, que también es muy buena estimación.

3.6 FÓRMULA DE EXTRAPOLACIÓN DE RICHARDSON

3.6.1 Introducción

Sabemos que:

$$\begin{aligned}
f'(x) &= \frac{f(x+h)-f(x-h)}{2h} \overbrace{-\frac{h^2}{3!}f'''(x) - \frac{h^4}{5!}f^{v}(x) - \frac{h^6}{7!}f^{vii}(x) - \cdots}^{\text{Error de truncamiento } h^{2k}} \\
f'(x) &= \frac{f(x+h)-f(x)}{h} \overbrace{-\frac{h}{2!}f''(x) - \frac{h^2}{3!}f'''(x) - \frac{h^3}{4!}f^{iv}(x) + \cdots}^{\text{Error de truncamiento } h^{k}} \\
f''(x) &= \frac{f(x+h)-2f(x)+f(x-h)}{h^2} + \underbrace{\sum_{i=1}^{\infty} 2\,\frac{f^{(2i+2)}(x)}{(2i+2)!}\,h^{2i}}
\end{aligned} \tag{3.1}$$

Nuestra estrategia sería entonces evaluar una de las fórmulas para una secuencia de valores de h tendentes a cero, parando cuando obtuviésemos el resultado con la precisión deseada. Este procedimiento, sin embargo, no es posible ya que eventualmente el error de redondeo dominará el cálculo. Cuando $h \longrightarrow 0$, tanto $f(x+h)$ como $f(x-h)$ tienden a $f(x)$, así que su diferencia tiende a la diferencia de dos cantidades casi iguales y así contiene cada vez menos dígitos significativos.

Por tanto, no tiene sentido llevar a cabo éste cálculo más allá de un cierto valor umbral para h que depende de la precisión con que $f(x)$ se puede calcular.

Si una cota del error relativo en el cálculo de $f(x)$ viene dada por R, o sea, el valor calculado $\bar{f}(x) = f(x)(1+r)$ con $|r| \leqslant R$ entonces una cota del error relativo es R/h.

Si quisiéramos un error relativo de ϵ (o menor) en la primera derivada, no podemos permitir que h sea menor que R/ϵ. Pra la segunda derivada el umbral es $h^2 > 4R/\epsilon$.

Ejemplo 25. *Calcular $f'(3)$ siendo $f(x) = \ln x$.*

Usando la fórmula central con dos puntos :

$$\ln'(3) \simeq \frac{\ln(3+h)-\ln(3-h)}{2h}$$

obtenemos sucesivamente (suponiendo una mantisa de 10 dígitos) :

h	**f′(3)**
0.1	0.3334568725
0.01	0.333334568
0.001	0.333333345
0.0001	0.33333335
0.00001	0.33333335
0.000001	0.3333335
10^{-7}	0.33335
10^{-8}	0.3335
10^{-9}	0.335
10^{-10}	0.35
10^{-11}	0.5

Obsérvese que inicialmente se van obteniendo mejores estimaciones, pero pasado un determinado valor, las estimaciones empeoran hasta el punto de carecer totalmente de significación.

3.6.2 Extrapolación de Richardson

Supongamos en general que disponemos de una fórmula $f(h)$ que nos aproxima una cantidad A :

$$A = f(h) + \overbrace{a_1 h^{n_1} + a_2 h^{n_2} + a_3 h^{n_3} + \cdots}^{\text{error de truncamiento}} \quad n_1 < n_2 < n_3 < \ldots \tag{3.2}$$

Utilizando ésta fórmula dos veces para obtener dos aproximaciones para dos valores diferentes de h, digamos h y h_1, donde suponemos $h_1 > h$, o sea, disponemos de $f(h)$ y $f(h_1)$. Como $f(h) \longrightarrow A$ cuando $h \longrightarrow 0$, $f(h)$ será más precisa que $f(h_1)$. En ésta situación, introducimos el paso

$$r = \frac{h_1}{h} > 1 \Longrightarrow h_1 = rh$$

Reemplazando en (3.2) cada ocurrencia de h por $h_1 = rh$:

$$\begin{aligned} A &= f(rh) + a_1(rh)^{n_1} + a_2(rh)^{n_2} + a_3(rh)^{n_3} + \cdots \\ &= f(rh) + a_1 r^{n_1} h^{n_1} + a_2 r^{n_2} h^{n_2} + a_3 r^{n_3} h^{n_3} + \cdots \end{aligned} \tag{3.3}$$

Multiplicando la ecuación (3.2) por r^{n_1} y restándole la ecuación (3.3) tenemos :

$$\begin{aligned}
r^{n_1}A - A &= r^{n_1}f(h) - f(rh) + r^{n_1}a_2h^{n_2} - a_2r^{n_2}h^{n_2} + r^{n_1}a_3h^{n_3} - a_3r^{n_3}h^{n_3} + \cdots \\
&= r^{n_1}f(h) - f(rh) + a_2h^{n_2}(r^{n_1} - r^{n_2}) + a_3h^{n_3}(r^{n_1} - r^{n_2}) + \cdots \Longrightarrow \\
\Longrightarrow A &= \frac{r^{n_1}f(h) - f(rh)}{r^{n_1} - 1} + \underbrace{\frac{r^{n_1} - r^{n_2}}{r^{n_1} - 1}\, a_2}_{b_2} h^{n_2} + \underbrace{\frac{r^{n_1} - r^{n_3}}{r^{n_1} - 1}\, a_3}_{b_3} h^{n_3} + \cdots \\
&= \underbrace{\frac{r^{n_1}f(h) - f(rh)}{r^{n_1} - 1}}_{f_1(h)} + b_2h^{n_2} + b_3h^{n_3} + \cdots
\end{aligned} \tag{3.4}$$

O sea, que la cantidad A la podemos aproximar también por la nueva fórmula $f_1(h)$

$$A = f_1(h) + b_2h^{n_2} + b_3h^{n_3} + \cdots \tag{3.5}$$

que tiene por término principal del error $b_2h^{n_2}$, es decir, es una fórmula de $O(h^{n_2})$, más precisa que la anterior, $f(h)$.
La fórmula anterior (3.5) también se puede escribir así :

$$\boxed{A = \frac{(h_1/h)^{n_1}\, f(h) - f(h_1)}{(h_1/h)^{n_1} - 1} + b_2h^{n_2} + b_3h^{n_3} + \cdots} \tag{3.6}$$

La fórmula de Richardson se podría aplicar a un conjunto de valores de A obtenidos para diferentes valores h_i y que denotaremos T_0^i, todos ellos con error $O(h^{n_1})$. De ésta forma se podría obtener una segunda columna que sería de $O(h^{n_2})$ y por tanto con valores más precisos para A. Combinando ahora dos valores sucesivos de ésta nueva columna, aplicando de nuevo Richardson, podríamos obtener una tercera columna que sería de $O(h^{n_3})$, etc.

h	$O(h^{n_1})$	$O(h^{n_2})$	$O(h^{n_3})$	$O(h^{n_4})$	$\cdots$
$h_1 \longrightarrow$	T_0^1				
$h_2 \longrightarrow$	T_0^2				
$h_3 \longrightarrow$	T_0^3				
$\vdots$	$\vdots$				
$h_m \longrightarrow$	T_0^m				

Ahora bien, si los $n_j = \beta\, j$, siendo $\underline{\beta \text{ fijo}}$, o sea, el error de truncamiento para A fuera de la forma $\sum_{k=1}^{\infty} a_k\, h^{\beta\, k}$, se podría construir un arreglo triangular mediante la aplicación de:

$$\boxed{T_k^i = \frac{(\frac{h_i}{h_{i+k}})^\beta\, T_{k-1}^{i+1} - T_{k-1}^i}{(h_i/h_{i+k})^\beta - 1} \quad \text{para} \begin{cases} k = 1, 2, \cdots (columna) \\ i = 1, 2, \cdots (fila) \end{cases}} \tag{3.7}$$

Entonces, si se utiliza para la aproximación de la derivada la fórmula

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

sabemos que el desarrollo del error de truncamiento tiene la forma

$$-\sum_{k=1}^{\infty} \frac{h^{2k}}{(2k+1)!}\, f^{(2k+1)}(x)$$

y por lo tanto es $\beta = 2$ en la fórmula general quedando entonces :

$$T_k^i = \frac{(\frac{h_i}{h_{i+k}})^2\, T_{k-1}^{i+1} - T_{k-1}^i}{(h_i/h_{i+k})^2 - 1}$$

Para éste caso, **si además** se consideran las h_i de forma que cada h sea la mitad del anterior, tendríamos :

$$\frac{h_i}{h_{i+1}} = 2 \Longrightarrow \frac{h_i}{h_{i+2}} = 4 \Longrightarrow \frac{h_i}{h_{i+3}} = 8 \Longrightarrow \cdots \Longrightarrow \frac{h_i}{h_{i+k}} = 2^k$$

con lo que

$$T_k^i = \frac{(2^k)^2\, T_{k-1}^{i+1} - T_{k-1}^i}{(2^k)^2 - 1} \Longrightarrow \boxed{T_k^i = \frac{4^k\, T_{k-1}^{i+1} - T_{k-1}^i}{4^k - 1}}$$

que es como generalmente aparece en los libros de texto la fórmula de extrapolación de Richardson, siendo éste un caso particular de aplicación de la fórmula.
La siguiente tabla indica el orden de cálculo de los elementos a construir para la obtención de las diferentes estimaciones sucesivas de la cantidad A :

$\mathbf{h}$	$\mathbf{T_0^i}$	$\mathbf{T_1^i}$	$\mathbf{T_2^i}$	$\mathbf{T_3^i}$	$\mathbf{T_4^i}$	$\mathbf{T_5^i}$	$\mathbf{T_6^i}$
$h_1 \longrightarrow$	$T_0^1 \searrow$						
		$T_1^1 \searrow$					
$h_2 \longrightarrow$	$T_0^2 \nearrow$		$T_2^1 \searrow$				
		$T_1^2 \nearrow$		$T_3^1 \searrow$			
$h_3 \longrightarrow$	T_0^3		$T_2^2 \nearrow$		$T_4^1 \searrow$		
		T_1^3		$T_3^2 \nearrow$		$T_5^1 \searrow$	
$h_4 \longrightarrow$	T_0^4		T_2^3		$T_4^2 \nearrow$		T_6^1
		T_1^4		T_3^3		$T_5^2 \nearrow$	
$h_5 \longrightarrow$	T_0^5		T_2^4		T_4^3		
		T_1^5		T_3^4			
$h_6 \longrightarrow$	T_0^6		T_2^5				
		T_1^6					
$h_7 \longrightarrow$	T_0^7						

Ejemplo 26. *Calcular la primera derivada de* $f(x) = \cot x$ *para* $x = 0.04$. *(valor exacto a 7 decimales : 625.33344002) Utilizar la fórmula central para dos puntos con los siguientes valores de* h *:*

$\mathbf{h_i/h_{i+1}}$	$\mathbf{h}$	$\mathbf{T_0^i}$	$\mathbf{T_1^i}$	$\mathbf{T_2^i}$	$\mathbf{T_3^i}$	$\mathbf{T_4^i}$	$\mathbf{T_5^i}$	$\mathbf{T_6^i}$
	$h_1 = 0.0256 \longrightarrow$	1058.9377906						
4/3			495.5225783					
	$h_2 = 0.0192 \longrightarrow$	812.4436352		640.1425224				
3/2			603.9875364		624.4282997			
	$h_3 = 0.0128 \longrightarrow$	696.6346914		626.6381123		625.3572216		
4/3			620.9754683		625.2991640		625.3330932	
	$h_4 = 0.0096 \longrightarrow$	663.5337813		625.4479360		625.3339415		625.3334398
3/2			624.3298191		625.3317679		625.3334344	
	$h_5 = 0.0064 \longrightarrow$	641.7538023		625.3481040		625.3334485		
4/3			625.0935328		625.3333435			
	$h_6 = 0.0048 \longrightarrow$	634.4649344		625.3349836				
3/2			625.2746209					
	$h_7 = 0.0032 \longrightarrow$	629.3592047						

donde en la primera columna aparecen los cocientes entre dos h_i consecutivas.

Ejemplo 27. *Queremos calcular $f'(3)$ siendo $f(x) = \ln x$ usando 6 dígitos y utilizando la fórmula central :*

$$f'(3) \simeq \frac{f(3+h) - f(3-h)}{2h}$$

Como se va a usar la fórmula central, es $\beta = 2$ y la fórmula de extrapolación de Richardson sabemos que queda :

$$T_k^i = \frac{4^k \, T_{k-1}^{i+1} - T_{k-1}^i}{4^k - 1}$$

si se utilizan los h_i de forma que h_{i+1} sea la mitad que el h_i.
La tabla con los cálculos queda de la siguiente forma :

$\mathbf{h}$	$\mathbf{T_0^i}$	$\mathbf{T_1^i}$	$\mathbf{T_2^i}$	$\mathbf{T_3^i}$
$h_1 = 0.8 \longrightarrow$	0.341590 $\searrow$			
		0.333243 $\searrow$		
$h_2 = 0.4 \longrightarrow$	0.335330 $\nearrow$		0.333336 $\searrow$	
		0.333330 $\nearrow$		0.333333
$h_3 = 0.2 \longrightarrow$	0.333830		0.333330 $\nearrow$	
		0.333330		
$h_4 = 0.1 \longrightarrow$	0.333455			

3.7 INTEGRACIÓN NUMÉRICA

3.7.1 Caso lineal (dos puntos)

Supongamos los puntos $(x_0, y_0), (x_1, y_1)$ siendo $x_1 = x_0 + h$, siendo h la longitud del intervalo de integración.
El polinomio de interpolación de Newton en diferencias progresivas es :

$$p_1(x) = y_0 + \Delta y_0 \, s, \text{ siendo } s = \frac{x - x_0}{h}$$

Por tanto $x = sh + x_0 \Longrightarrow dx = hds$ y tenemos :

$$\int_{x_0}^{x_1} p_1(x)dx = h \int_0^1 [y_0 + (y_1 - y_0)s]ds =$$

$$= h \left[y_0 s + (y_1 - y_0)\frac{s^2}{2} \right] \Bigg|_0^1 = h \left[y_0 + (y_1 - y_0)\frac{1}{2} \right] = \boxed{\frac{h}{2}\,(y_0 + y_1)}$$

conocida como **fórmula del trapecio** ya que geométricamente representa el área del trapecio bajo los puntos $(x_0, y_0), (x_1, y_1)$.

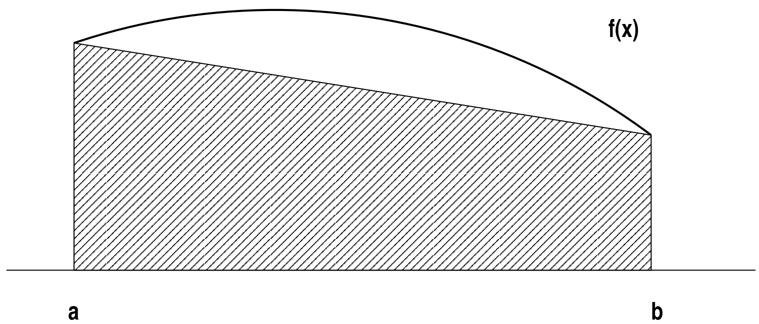

Figura 3.2: Fórmula del trapecio.

Vamos ahora a calcular el error cometido E al aproximar la integral mediante la regla del trapecio. Para ello, consideraremos que F es una primitiva de la función $y(x)$ que estamos integrando y entonces $F'(x) = y(x)$ y se verifica :

$$I = \int_{x_0}^{x_1} y(x)dx \overset{\text{(Barrow)}}{=} F(x_1) - F(x_0) = F(x_0) + hF'(x_0) + \frac{h^2}{2!}F''(x_0) + \cdots \; - F(x_0)$$
$$= hF'(x_0) + \frac{h^2}{2!}F''(x_0) + \frac{h^3}{3!}F'''(x_0) + \cdots = h\,y_0 + \frac{h^2}{2!}\,y_0' + \frac{h^3}{3!}\,y_0'' + \cdots$$

Por otra parte tenemos :

$$\frac{h}{2}\,(y_0 + y_1) = \frac{h}{2}\,(y_0 + y_0 + hy_0' + \frac{h^2}{2!}\,y_0'' + \frac{h^3}{3!}\,y_0''' + \cdots) \Longrightarrow$$
$$\Longrightarrow E = I - \frac{h}{2}\,(y_0 + y_1) = -\frac{h^3}{12}\,y_0'' - \frac{h^4}{24}\,y_0''' - \cdots - \frac{2-n}{2\,n!}\,h^n\,y_0^{(n-1)} - \cdots)$$

o también :

$$\boxed{E = -\frac{h^3}{12}\,y''(\xi), \quad \text{donde } \xi \in \,]x_0, x_1[}$$

3.7.2 Caso cuadrático (tres puntos)

Sean los puntos $(x_0, y_0), (x_1, y_1), (x_2, y_2)$ siendo $x_1 = x_0 + h, x_2 = x_0 + 2h$, siendo h la mitad de la longitud total del intervalo de integración.
Ahora el polinomio de Newton es:

$$p_2(x) = y_0 + \Delta y_0\, s + \frac{\Delta^2 y_0}{2!} s(s-1), \text{ siendo } s = \frac{x - x_0}{h}$$

como antes. Tenemos entonces :

$$\begin{aligned}
\int_{x_0}^{x_2} p_2(x)dx &= h \int_0^2 \left[y_0 + (y_1 - y_0)s + \frac{1}{2}(y_2 - 2y_1 + y_0)s(s-1) \right] ds = \\
&= h \left[y_0 s + (y_1 - y_0)\frac{s^2}{2} + \frac{1}{2}(y_2 - 2y_1 + y_0)\left(\frac{s^3}{3} - \frac{s^2}{2}\right) \right] \Bigg|_0^2 \\
&= h \left[2y_0 + (y_1 - y_0)2 + \frac{1}{2}(y_2 - 2y_1 + y_0)\left(\frac{8}{3} - 2\right) \right] \\
&= \boxed{\frac{h}{3}\,(y_0 + 4y_1 + y_2)}
\end{aligned}$$

que se denomina **regla de Simpson**.

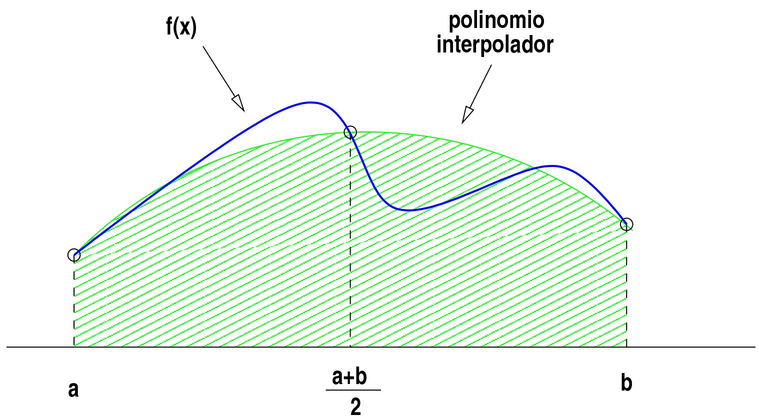

Figura 3.3: Fórmula de Simpson.

El error cometido E al aproximar la integral mediante la regla de Simpson es :

$$\begin{aligned}
\frac{h}{3}\,(y_0 + 4\,y_1 + y_2) &= \frac{h}{3}\,(y_0 + 4\,(y_0 + hy_0' + \frac{h^2}{2!}\,y_0'' + \frac{h^3}{3!}\,y_0''' + \cdots) + \\
&+ (y_0 + 2h\,y_0' + \frac{(2h)^2}{2!}\,y_0'' + \frac{(2h)^3}{3!}\,y_0''' + \cdots) = \\
&= \frac{h}{3}\,(6\,y_0 + 6h\,y_0' + 4\,h^2\,y_0'' + 2\,h^3\,y_0''' + \frac{5}{6}\,h^4\,y_0^{(4)} + \cdots)
\end{aligned}$$

y por otra parte :

$$I = \int_{x_0}^{x_2} y(x)\, dx \overset{\text{(Barrow)}}{=} F(x_2) - F(x_0) = 2h\, F'(x_0) + \frac{(2h)^2}{2!} F''(x_0) + \frac{(2h)^3}{3!} F'''(x_0) + \cdots$$
$$= 2h\, y_0 + 2h^2\, y_0' + \frac{4}{3}\, h^3\, y_0'' + \frac{2}{3}\, h^4\, y_0''' + \frac{4}{15}\, h^5\, y_0^{(4)} + \cdots$$

y restando :

$$\boxed{E = \int_{x_0}^{x_2} y(x)\, dx - \frac{h}{3}\,(y_0 + 4\, y_1 + y_2) = -\frac{h^5}{90}\, y^{(4)}(\xi), \quad \text{donde } \xi \in\,]x_0, x_2[}$$

3.8 TABLA DE FÓRMULAS ELEMENTALES DE INTEGRACIÓN

Veremos sólo las fórmulas cerradas de Newton-Cotes, donde n se refiere al grado del polinomio.
Son todas del tipo :

$$\int_{x_0}^{x_n} y(x)\, dx = k\, h\, (c_0\, y_0 + c_1\, y_1 + \cdots + c_n\, y_n)$$

donde k es una constante que multiplica al paso $h = \frac{x_n - x_0}{n}$.

Nombre	**n**	**Fórmula**	**Error**
Trapecio	1	$\frac{1}{2}\, h\,(y_0 + y_1)$	$-\frac{h^3}{12}\, y''(\xi)$
Simpson	2	$\frac{1}{3}\, h\,(y_0 + 4y_1 + y_2)$	$-\frac{h^5}{90}\, y^{iv}(\xi)$
3/8 Simpson	3	$\frac{3}{8}\, h\,(y_0 + 3\, y_1 + 3\, y_2 + y_3)$	$-\frac{3h^5}{80}\, y^{iv}(\xi)$
Milne / Boole	4	$\frac{2}{45}\, h\,(7y_0 + 32\, y_1 + 12\, y_2 + 32y_3 + 7y_4)$	$-\frac{8h^7}{945}\, y^{(6)}(\xi)$
	5	$\frac{5}{288}\, h\,(19y_0 + 75\, y_1 + 50\, y_2 + 50y_3 + 75y_4 + 19y_5)$	$-\frac{275h^7}{12096}\, y^{(6)}(\xi)$
Weddle	6	$\frac{1}{140}\, h\,(41y_0 + 216\, y_1 + 27(4)\, y_2 + 272y_3 + 27y_4 + 216y_5 + 41y_6)$	$-\frac{9h^9}{1400}\, y^{(8)}(\xi)$

o también en forma más detallada

$$\int_{x_0}^{x_1} f(x)\,dx = \frac{1}{2}\,h\,(y_0+y_1) - \frac{h^3}{12}\,y''(\xi) \ \ \text{(Trapecio)}$$
$$\int_{x_0}^{x_2} f(x)\,dx = \frac{1}{3}\,h\,(y_0+4\,y_1+y_2) - \frac{h^5}{90}\,y^{iv}(\xi) \ \ \text{(Simpson)}$$
$$\int_{x_0}^{x_3} f(x)\,dx = \frac{3}{8}\,h\,(y_0+3\,y_1+3\,y_2+y_3) - \frac{3h^5}{80}\,y^{iv}(\xi) \ \ \text{(3/8 Simpson)}$$
$$\int_{x_0}^{x_4} f(x)\,dx = \frac{2}{45}\,h\,(7\,y_0+32\,y_1+12\,y_2+32\,y_3+7\,y_4) - \frac{8h^7}{945}\,y^{(6)}(\xi) \ \ \text{(Milne)}$$
$$\int_{x_0}^{x_5} f(x)\,dx = \frac{5}{288}\,h\,(19\,y_0+75\,y_1+50\,y_2+50\,y_3+75\,y_4+19\,y_5) - \frac{275h^7}{12096}\,y^{(6)}(\xi)$$
$$\int_{x_0}^{x_6} f(x)\,dx = \frac{1}{140}\,h\,(41\,y_0+216\,y_1+27\,y_2+272\,y_3+27\,y_4+216\,y_5+41\,y_6)$$
$$- \frac{9h^9}{1400}\,y^{(8)}(\xi) \ \ \text{(Weddle)}$$
$$\int_{x_0}^{x_7} f(x)\,dx = \frac{7}{17280}\,h\,[751\,(y_0+y_7)+3577\,(y_1+y_6)+1323\,(y_2+y_5)+2989\,(y_3+y_4)$$
$$- \frac{8183\,h^9}{518400}\,y^{(8)}(\xi)$$
$$\int_{x_0}^{x_8} f(x)\,dx = \frac{4}{14175}\,h\,[989\,(y_0+y_8)+5888\,(y_1+y_7)-928\,(y_2+y_6)+10496\,(y_3+y_5)$$
$$-4540\,y_4] - \frac{2368\,h^{11}}{467775}\,y^{(10)}(\xi)$$
$$\int_{x_0}^{x_9} f(x)\,dx = \frac{9}{89600}\,h\,[2857\,(y_0+y_9)+15741\,(y_1+y_8)+1080\,(y_2+y_7)+19344\,(y_3+y_6)$$
$$+5778\,(y_4+y_5)] - \frac{173\,h^{11}}{14620}\,y^{(10)}(\xi)$$
$$\int_{x_0}^{x_{10}} f(x)\,dx = \frac{5}{299376}\,h\,\Big[16067\,(y_0+y_{10})+106300\,(y_1+y_9)-48525\,(y_2+y_8)+272400\,(y_3+y_7)$$
$$-260550\,(y_4+y_6)+427368\,y_5] + -\frac{1346350\,h^{13}}{326918592}\,y^{(12)}(\xi)$$

Ejemplo 28. *Calcular*

$$\int_0^{\pi/2} \frac{dx}{\sin^2 x + \frac{1}{4}\cos^2 x} = \pi$$

Los valores que se obtienen por la diferentes fórmulas son :

$$\text{Trapecio}:\ (0,\pi/2) \quad I = \frac{h}{2}(y_0+y_1) = \frac{\pi/2}{2}(4+1) = \frac{5}{4}\,\pi$$

Simpson : $(0, \pi/4, \pi/2)$ $\quad I = \frac{h}{3}\,(y_0 + 4y_1 + y_2) = \frac{\pi/4}{3}\,(4 + 4 \times 1.6 + 1) = 0.95\,\pi$

3/8 de Simpson : $(0, \pi/6, \pi/3, \pi/2)$ $\quad I = \frac{3}{8}\frac{\pi}{6}(4{+}2.285714286{\times}3{+}1.230769231{\times}3{+}1) = 0.9718406593\pi$

Milne : $(0, \pi/8, \pi/4, 3\pi/8, \pi/2)$:

$$I = \frac{2}{45}\frac{\pi}{8}(7 \times 4 + 32 \times 2.77905184 + 12 \times 1.6 + 32 \times 1.12338718 + 7 \times 1) = 0.9948780488\,\pi$$

3.9 FÓRMULAS COMPUESTAS

Queremos obtener una estimación de $\int_a^b f(x)dx$ más exacta que la que nos proporciona la aplicación de una de las fórmulas elementales de integración. Una forma sería subdividir el intervalo [a,b] en n subintervalos de forma que $x_0 = a, x_n = b, h = \frac{b-a}{n}$

$$\int_a^b f(x)dx = \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \cdots + \int_{x_{n-1}}^{x_n} f(x)dx = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k-1}} f(x)dx$$

- Para el caso de la regla del trapecio :

$$\int_a^b f(x)dx = \frac{(x_1 - x_0)}{2}\,(y_0 + y_1) + \frac{(x_2 - x_1)}{2}\,(y_1 + y_2) + \cdots + \frac{(x_n - x_{n-1})}{2}\,(y_{n-1} + y_n)$$

que se denomina fórmula general trapezoidal (véase figura 3.4).

Si las abcisas son equidistantes, o sea, $x = x_0 + ih$ tenemos :

$$\begin{aligned} &= \frac{h}{2}\,(y_0 + y_1) + \frac{h}{2}\,(y_1 + y_2) + \cdots + \frac{h}{2}\,(y_{n-1} + y_n) \\ &= \boxed{h\left(\frac{1}{2}\,y_0 + y_1 + \cdots + y_{n-1} + \frac{1}{2}\,y_n\right)} \end{aligned}$$

que se conoce con el nombre de **regla trapezoidal**.

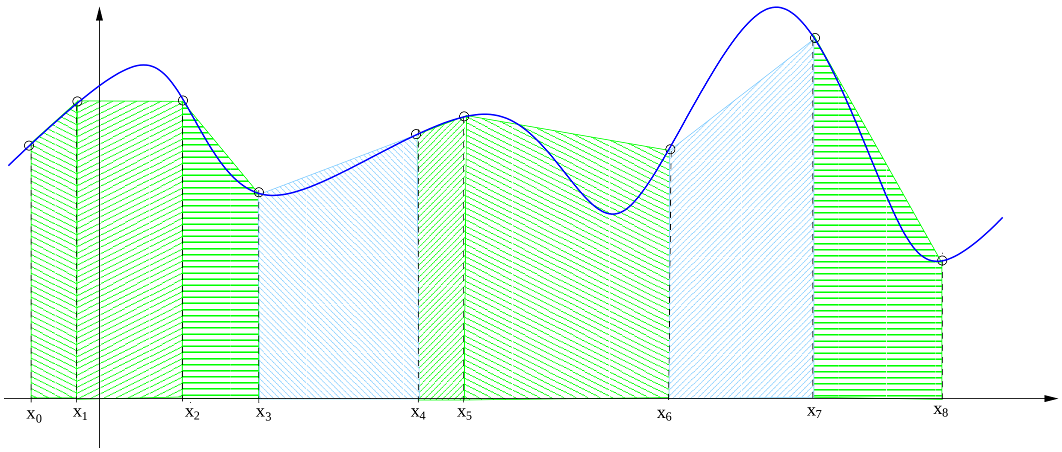

Figura 3.4: Fórmula general trapezoidal.

- Para la regla de Simpson :

$$\begin{aligned} I &= \int_a^b f(x)dx = \int_{x_0}^{x_2} f(x)dx + \int_{x_2}^{x_4} f(x)dx + \cdots + \int_{x_{n-2}}^{x_n} f(x)dx = \\ &\text{(las abcisas han de ser equidistantes)} \\ &= \frac{h}{3}(y_0 + 4y_1 + y_2) + \frac{h}{3}(y_2 + 4y_3 + y_4) + \cdots + \frac{h}{3}(y_{n-2} + 4y_{n-1} + y_n) \\ &= \boxed{\frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n)} \end{aligned}$$

cuyo nombre es **regla parabólica**.

3.9.1 Error en las fórmulas compuestas

Teorema 3.9.1. *Sea f una función continua en $[a, b]$ tal que $\xi_i \in [a, b]$, $\forall i = 1, 2, \ldots, m$ y siendo $\alpha_i \in \mathbb{R}$ no nulos y todos del mismo signo. En éstas condiciones se verifica :*

$$\sum_{i=1}^{m} \alpha_i f(\xi_i) = \alpha f(\xi), \quad \textit{siendo } \xi \in [a, b], \;\; \alpha = \sum_{i=1}^{m} \alpha_i$$

Demostración. Suponemos que todos los α_i son positivos.

Sean

$$\begin{cases} \underline{f(x)} = \min\limits_{1\leqslant i\leqslant n} f(\xi_i) \\ \\ \overline{f(x)} = \max\limits_{1\leqslant i\leqslant n} f(\xi_i) \end{cases} \tag{3.8}$$

entonces

$$\begin{aligned}
&\underline{f(x)} \leqslant f(\xi_i) \leqslant \overline{f(x)} \quad \forall i = 1, 2, \cdots, m \quad \text{(multiplicando por } \alpha_i) \\
&\alpha_i \underline{f(x)} \leqslant \alpha_i f(\xi_i) \leqslant \alpha_i \overline{f(x)} \quad \text{(sumando de 1 a n)} \\
&\sum_1^n \alpha_i \underline{f(x)} \leqslant \sum_1^n \alpha_i f(\xi_i) \leqslant \sum_1^n \alpha_i \overline{f(x)} \quad \left(\text{sustituyendo } \alpha = \sum_{i=1}^m \alpha_i \right) \\
&\alpha \underline{f(x)} \leqslant \sum_1^n \alpha_i f(\xi_i) \leqslant \alpha \overline{f(x)} \text{ (dividiendo por } \alpha) \\
&\underline{f(x)} \leqslant \sum_1^n \frac{\alpha_i f(\xi_i)}{\alpha} \leqslant \overline{f(x)}
\end{aligned}$$

y por el teorema del Valor Medio

$$\exists\, \xi \in [a, b] \quad \text{tal que} \quad f(\xi) = \sum_1^n \frac{\alpha_i f(\xi_i)}{\alpha}$$

□

- Error en la fórmula trapezoidal. El error en cada subintervalo es

$$E_i = -\frac{h^3}{12} f''(\xi_i)$$

Por lo tanto el error total será la suma de los errores en cada intervalo :

$$E_T = \sum_1^m E_i = \sum_1^m -\frac{h^3}{12} f''(\xi_i) \overset{(\alpha_i=1)}{=} -\frac{h^3}{12} \sum_1^m 1 \,.\, f''(\xi_i) \overset{\text{(teorema)}}{=} -\frac{h^3}{12}\, m\, f''(\xi)$$

$$E_T = -\frac{h^3}{12}\, m\, f''(\xi) = -\left(\frac{b-a}{m} \right)^3 \frac{m}{12}\, f''(\xi) = -\frac{(b-a)h^2}{12}\, f''(\xi)$$

- Error en la fórmula parabólica. En éste caso es

$$E_i = -\frac{h^5}{90} f^{(4)}(\xi_i)$$

Y sumando los errores de cada subintervalo :

$$E_T = \sum_1^{m/2} E_i = \sum_1^{\frac{m}{2}} -\frac{h^5}{90} f''(\xi_i) \overset{(\alpha_i=1)}{=} -\frac{h^5}{90} \sum_1^{m/2} 1 \,.\, f''(\xi_i) \overset{(\text{teorema})}{=} -\frac{h^5}{90}\,\frac{m}{2}\, f''(\xi)$$
$$E_T = -\frac{(b-a)^5}{18m^4}\, f^{(4)}(\xi) = -\frac{mh^5}{180}\, f^{(4)}(\xi) = -\frac{(b-a)}{180} h^4\, f^{(4)}(\xi)$$

Ejemplo 29. *Obtener diferentes estimaciones por las diferentes reglas de*

$$\int_{-4}^{4} \frac{dx}{1+x^2} = 2\arctan 4 \simeq 2.6516353$$

Nº abcisas	**Newton-Cotes**	**Parabólica**	**Trapezoidal**	**Romberg**
3	5.490	5.490	4.235	5.490
5	2.278	2.478	2.918	2.278
7	3.329	2.908	2.701	
9	1.941	2.573	2.659	2.584
11	3.596	2.695	2.6511	
17		2.6477	2.6505	2.6542
33		2.651627	2.65135	2.65186
85		2.656353	2.65156	2.651631
129		2.6516353	2.651617	2.6516353

3.10 INTEGRACIÓN DE ROMBERG

En el caso de la regla del trapecio, la subdivisión del intervalo de integración en subintervalos más pequeños produce una mejor estimación de la integral (en éste caso es un proceso convergente) Es la aplicación de la extrapolación de Richardson a la regla trapezoidal.
Sabemos que :

$$\int_a^b f(x)dx = h\left(\frac{1}{2}f_0 + f_1 + \cdots + f_{n-1} + \frac{1}{2}f_n\right) + \sum_{i=1}^{\infty} a_i h^{2i}$$

con $h = (b-a)/n$, donde los a_i dependen sólo de a, b y $f(x)$.
Sea ahora

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

y sean

$$T_0^k = h\left(\frac{1}{2}f_0 + f_1 + \cdots + f_{2^k-1} + \frac{1}{2}f_{2^k}\right), \quad h = \frac{b-a}{2^k}, \quad f_i = f(x_i) \ (2^k \text{ puntos })$$

diferentes estimaciones obtenidas subdividiendo el intervalo de integración en $n = 2^k$ subintervalos, aproximaciones a I por la regla del trapecio.
De forma que tendremos :

$$\begin{aligned}
T_0^0 &= \frac{b-a}{2^0}\left(\frac{1}{2}f_0 + \frac{1}{2}f_1\right)\\
T_0^1 &= \frac{b-a}{2}\left(\frac{1}{2}f_0 + f_1 + \frac{1}{2}f_2\right)\\
T_0^2 &= \frac{b-a}{2^2}\left(\frac{1}{2}f_0 + f_1 + f_2 + f_3 + \frac{1}{2}f_4\right)\\
&\text{(etcétera)}
\end{aligned}$$

Entonces :

$$T_0^k = I - \sum_{i=1}^{\infty} a_i \left(\frac{b-a}{2^k}\right)^{2i} \tag{3.9}$$

y definimos :

$$T_1^k = \frac{4\,T_0^{k+1} - T_0^k}{3}, \quad k = 0, 1, 2, \ldots$$

Por tanto, usando (3.9) tenemos :

$$
\begin{aligned}
T_1^k &= \frac{1}{3}\left(4\left(I-\sum_{i=1}^{\infty} a_i\left(\frac{b-a}{2^{k+1}}\right)^{2i}\right)-\left(I-\sum_{i=1}^{\infty} a_i\left(\frac{b-a}{2^{k}}\right)^{2i}\right)\right)= \\
&= \frac{1}{3}\left[4\,I-4\sum_{i=1}^{\infty} a_i\left(\frac{b-a}{2^{k+1}}\right)^{2i}-I+\sum_{i=1}^{\infty} a_i\left(\frac{b-a}{2^{k}}\right)^{2i}\right]= \\
&= I-\frac{1}{3}\left(\sum_{i=1}^{\infty}\frac{4}{2^{2i}}\,a_i\left(\frac{b-a}{2^{k}}\right)^{2i}-\sum_{i=1}^{\infty} a_i\left(\frac{b-a}{2^{k}}\right)^{2i}\right)= \\
&= I-\frac{1}{3}\sum_{i=1}^{\infty}\left(\frac{4}{2^{2i}}-1\right)\,a_i\left(\frac{b-a}{2^{k}}\right)^{2i}= \\
&= I-\sum_{i=1}^{\infty}\frac{1}{3}\left(\frac{4}{2^{2i}}-1\right)\,a_i\left(\frac{b-a}{2^{k}}\right)^{2i}= \\
&= I-\sum_{\underbrace{\mathbf{i=2}}_{!}}^{\infty}\underbrace{\frac{1}{3}\left(\frac{4}{2^{2i}}-1\right)\,a_i}_{b_i}\left(\frac{b-a}{2^{k}}\right)^{2i}
\end{aligned}
$$

Esta última ecuación afirma que si aproximamos por la regla del trapecio con $h_1=(b-a)/2^{k+1}$ y $h_2=(b-a)/2^k$ la aproximación resultante por T_1^k tiene como término de error principal $O(h_2^4)$.

La aproximación T_1^k es en efecto la regla parabólica para 2^k subintervalos.

Definimos entonces en general la fórmula recursiva :

$$
\boxed{T_m^k=\frac{4^m\,T_{m-1}^{k+1}-T_{m-1}^k}{4^m-1}\qquad \text{para}\;\begin{cases} k=0,1,\cdots \\ m=1,2,\cdots\end{cases}} \tag{3.10}
$$

y la disposición de los cálculos es como sigue :

h	$\mathbf{T_0^i}$	$\mathbf{T_1^i}$	$\mathbf{T_2^i}$	$\mathbf{T_3^i}$
$h=b-a$	$T_0^0\searrow$			
		$T_1^0\searrow$		
$h=\frac{b-a}{2}$	$T_0^1\nearrow$		$T_2^0\searrow$	
		$T_1^1\nearrow$		T_3^0
$h=\frac{b-a}{2^2}$	T_0^3		$T_2^1\nearrow$	
		T_1^2		
$h=\frac{b-a}{2^3}$	T_0^3			

La aproximación T_2^k es la regla compuesta correspondiente a la fórmula 3/8 de Simpson !

, pero para $m > 2$ no hay relación directa entre T_m^k y alguna regla compuesta de Newton-Cotes.

Con respecto al error en la integración de Romberg se puede probar (Bauer, Rutishauser y Stiefel (1963)) que éste error se puede expresar en forma estándar como :

$$\boxed{E = k\, h^{2m+2} f^{(2m+2)}(\xi), \quad \text{con } h = \frac{b-a}{2^{m+k}}}$$

El esquema de Romberg es de gran interés teórico y ha sido usado en algunas rutinas de integración e incluso en algunos esquemas adaptativos; sin embargo, el número de evaluaciones funcionales crece muy rápidamente con m.

Se puede elegir otra sucesión $\{h_n\}$ distinta a la $\{h/2^n\}$ y aún tiene beneficios de la extrapolación de Richardson. La elección particular de $\{h_n\} = \{(b-a)/p_n\}$, donde $\{p_n\} = \{1, 2, 3, 4, 6, 8, 12, \dots\}$ ha sido recomendada por Oliver (1971) como óptima en el sentido de proporcionar la mejor precisión con el menor error de redondeo para una "cantidad fija de cálculo". Aquí, los p_n son los enteros de la forma 2^k y $3.(2^k)$, de forma que todas las evaluaciones del integrando se usan en el cálculo de las siguientes sumas trapezoidales.

Ejemplo 30. *Calcular por Romberg*

$$\int_1^{2.2} \ln x \, dx$$

Solución : Es $\quad h = 2.2 - 1 = 1.2$

$$\begin{aligned}
T_0^0 &= \frac{1.2}{2^0}\left(\frac{1}{2}\ln 1 + \frac{1}{2}\ln 2.2\right) = 0.4730744162 \\
T_0^1 &= \frac{1.2}{4}\,(\ln 1 + \underline{2\ln 1.6} + \ln 2.2) = 0.5185393857 \\
T_0^2 &= \frac{1.2}{8}\,(\ln 1 + \underline{2\ln 1.3} + 2\ln 1.6 + \underline{2\ln 1.9} + \ln 2.2) = 0.5305351381 \\
T_0^3 &= \frac{1.2}{16}(\ln 1 + \underline{2\ln 1.15} + 2\ln 1.3 + \underline{2\ln 1.45} + 2\ln 1.6 + \underline{2\ln 1.75} + 2\ln 1.9 + \\
&\quad + \underline{2\ln 2.05} + \ln 2.2) = 0.533584731 \\
T_0^4 &= \frac{1.2}{32}(\ln 1 + \underline{2\ln 1.075} + 2\ln 1.15 + \underline{2\ln 1.225} + 2\ln 1.3 + \underline{2\ln 1.375} + 2\ln 1.45 + \\
&\quad + \underline{2\ln 1.525} + 2\ln 1.6 + \underline{2\ln 1.675} + 2\ln 1.75 + \underline{2\ln 1.825} + 2\ln 1.9 + \underline{2\ln 1.975} + \\
&\quad + 2\ln 2.05 + \underline{2\ln 2.125} + \ln 2.2) = 0.5343505905
\end{aligned}$$

(etcétera)

Los sumandos que aparecen subrayados son los que aparecen nuevos al pasar del elemento T_0^{k-1} al T_0^k. Por tanto se podrían aprovechar las evaluaciones ya realizadas de la función f y los cálculos ya hechos para obtener T_0^{k-1} y así calcular T_0^k

$$T_0^k = \frac{T_0^{k-1}}{2} + \frac{(b-a)}{2^k} \sum_{i=0}^{2^{k-1}} f\left[a + \frac{b-a}{2^k}(2\,i-1)\right]$$

Los cálculos se podrían ordenar de la siguiente forma :

h	$\mathbf{T_0^k}$	$\mathbf{T_1^k}$	$\mathbf{T_2^k}$	$\mathbf{T_3^k}$	$\mathbf{T_4^k}$
$h = b-a$	0.4730744162	0.5336943755	0.5345896786	0.53460602	0.534606192
$h = (b-a)/2$	0.5185393857	0.5345337221	0.5346057646	0.5346061913	□
$h = (b-a)/4$	0.5305351381	0.534601262	0.5346061846		
$h = (b-a)/8$	0.5335847310	0.534605877			
$h = (b-a)/16$	0.5343505905				

3.11 INTEGRACIÓN DE GAUSS

Definición 4. Polinomios ortogonales
Sea $\{g_i(x)\}_{i=0}^{\infty}$ una familia de polinomios definida en [a,b]. Se dice que $\{g_i(x)\}_{i=0}^{\infty}$ es una **familia ortogonal de polinomios** *en $[a,b]$ respecto de la función peso $w(x)$ si :*

$$\int_a^b w(x)g_n(x)g_m(x)dx = 0, \ \text{para } n \neq m \ \text{ y además}$$
$$\int_a^b w(x)[g_n(x)]^2 dx \neq 0$$

Definición 5. *Los* **polinomios de Legendre** *son una familia de polinomios $\{P_n(x)\}_{n=0}^{\infty}$ ortogonales en el intervalo $[-1,1]$ respecto de la función peso $w(x) = 1$ y se definen recursivamente como:*

LEGENDRE

$$\begin{aligned} P_n(x) &= \frac{2n-1}{n}\,x\,P_{n-1}(x) - \frac{n-1}{n}\,P_{n-2}(x), \quad n \geqslant 2 \\ P_0(x) &= 1\,; \quad P_1(x) = x \end{aligned}$$

Los primeros polinomios de Legendre son :

$$P_2(x) = \frac{1}{2}(3x^2 - 1) \qquad P_3(x) = \frac{1}{2}(5x^3 - 3x)$$
$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3) \qquad P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$$
$$P_6(x) = \frac{1}{16}(231x^6 - 315x^4 + 105x^2 - 5) \quad P_7(x) = \frac{1}{16}(429x^7 - 693x^5 + 315x^3 - 35x)$$

Tienen las siguientes propiedades :

Propiedad 3. *$P_n(x)$ es un polinomio de grado n.*

Propiedad 4. *Sus raíces están distribuidas simétricamente en el intervalo $[-1, 1]$*

Propiedad 5. *$\{P_0(x), P_1(x), \cdots, P_n(x)\}$ es una base del espacio vectorial de los polinomios de grado menor o igual que n. Esto es,*
$\forall p(x) \in \mathbb{R}[x], \exists \alpha_i \in \mathbb{R}, \ \ i = 0, 1, \ldots, n$ tales que $p(x) = \sum\limits_{i=0}^{n} \alpha_i P_i(x)$ de forma única.

Propiedad 6. *Se verifica que*

$$\int_{-1}^{+1} P_n(x) P_m(x) dx = 0, \ \ \textit{para } n \neq m$$
$$\int_{-1}^{+1} [P_n(x)]^2 dx = \frac{2}{2n+1}$$

Propiedad 7. *En el caso de la integración de Gauss-Legendre :*

$$w_k = \frac{1}{[P_n'(x_k)]^2} \frac{2}{1 - x_k^2}, \ \ \textit{(pesos)}$$

$$R_n = \frac{2^{2n+1}(n!)^4}{(2n+1)[(2n)!]^3} f^{(2n)}(\xi) \ \ (-1 < \xi < 1) \ \ \textit{(error)}$$

Teorema 3.11.1. *Dados los $n+1$ puntos $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$ es posible construir una fórmula de integración que sea exacta hasta polinomios de grado $2n + 1$.*

Demostración. Sean los $n + 1$ puntos $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$ y se desea calcular

$$\int_a^b f(x)\, dx \tag{3.11}$$

El polinomio de interpolación de Lagrange en estos $n + 1$ puntos será :

$$p_n(x) = \sum_{i=0}^{n} \overbrace{f(x_i)}^{y_i} \, . \, L_i(x) = \sum_{i=0}^{n} f(x_i) \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j}$$

donde $\xi \in [a, b]$.
Como usaremos los polinomios de Legendre, debemos tenet como intervalo de integracion el [-1,1] en lugar del [a,b], para los cual haremos el cambio de variable :

$$\begin{array}{ll} x = a \longrightarrow \quad z = -1 \searrow & \\ & \frac{(b-a)}{(1+1)} \Longrightarrow x = a + (b-a)\,(z+1)/2 \\ x = b \longrightarrow \quad z = +1 \nearrow & \end{array}$$

Por tanto

$$x = \frac{z(b-a) + b + a}{2} \Longrightarrow f(x) = f\left(\frac{z(b-a) + b + a}{2}\right) = F(z) = p_n(z) + E_n(z)$$

Por tanto, si la fórmula de integración ha de ser exacta para polinomios hasta grado $2n+1$ tendríamos :

$$\underbrace{f(x)}_{\text{pol. grado } 2n+1} = \underbrace{\sum_{i=0}^{n} f(x_i) \underbrace{\prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j}}_{\text{pol. grado } n}}_{\text{pol. grado máximo } n} + \overbrace{\frac{f^{(n+1)}(\xi)}{(n+1)!} \underbrace{(x - x_0)(x - x_1)\ldots(x - x_n)}_{\text{pol. grado } n+1}}^{\text{Error de interpolación } E_n(x)}$$

integrando entonces :

$$\int_{-1}^{+1} f(x)\,dx = \int_{-1}^{+1} \sum_{i=0}^{n} f(x_i) \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j}\,dx + \int_{-1}^{+1} \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)(x - x_1)\ldots(x - x_n)\,dx \tag{3.12}$$

Por otra parte :

$$\begin{aligned} \prod_{i=0}^{n}(x - x_i) &= (x - x_0)(x - x_1)\ldots(x - x_n) = \\ &= \alpha_0\, P_0(x) + \alpha_1\, P_1(x) + \cdots + \alpha_n\, P_n(x) + \alpha_{n+1}\, P_{n+1}(x) \end{aligned} \tag{3.13}$$

y además :

$$\frac{f^{(n+1)}(\xi)}{(n+1)!} = \beta_0\, P_0(x) + \beta_1\, P_1(x) + \cdots + \beta_n\, P_n(x) \tag{3.14}$$

Por lo que (3.12) queda :

$$\int_{-1}^{+1} f(x)\,dx = \sum_{i=0}^{n} f(x_i) \overbrace{\int_{-1}^{+1} \prod_{\substack{j=0\\ j\neq i}}^{n} \frac{x-x_j}{x_i-x_j}\,dx}^{w_i} + \overbrace{\int_{-1}^{+1} \sum_{i=0}^{n+1}\sum_{j=0}^{n} \alpha_i\beta_j P_i(x)P_j(x)\,dx}^{\textbf{tiene que ser cero}} \quad (3.15)$$

Por tanto, ya sabemos que :

$$\boxed{w_i = \int_{-1}^{+1} \prod_{\substack{j=0\\ j\neq i}}^{n} \frac{x-x_j}{x_i-x_j}\,dx}$$

y el segundo sumando de la ecuación (3.15) se transforma en :

$$\sum_{j=0}^{n} \int_{-1}^{+1} \alpha_{n+1}\beta_j P_{n+1}(x)P_j(x)\,dx$$

y tiene que ser :

$$\sum_{i=0}^{n} \int_{-1}^{+1} \alpha_i\beta_i[P_i(x)]^2\,dx = 0 \Longrightarrow \alpha_i = 0, \forall i = 0,1,\ldots,n$$

Por tanto :

$$\prod_{i=0}^{n}(x-x_i) = \overbrace{\alpha_0\,P_0(x)}^{0} + \overbrace{\alpha_1\,P_1(x)}^{0} + \cdots + \overbrace{\alpha_n\,P_n(x)}^{0} + \alpha_{n+1}\,P_{n+1}(x)$$

y entonces los x_i han de ser las raíces del polinomio de Legendre de grado $n+1$.
Por tanto :

$$\int_{-1}^{+1} f(x)\,dx = \sum_{i=0}^{n} w_i f(x_i), \quad \text{donde} \begin{cases} w_i = \int_{-1}^{+1} \prod\limits_{\substack{j=0\\ j\neq i}}^{n} \frac{x-x_j}{x_i-x_j}\,dx \\ \\ x_i \text{ es la i-ésima raiz de } P_{\underline{n+1}}(x) \end{cases}$$

□

La fórmula para un intervalo arbitrario queda

$$\boxed{\int_a^b f(x)\,dx = \frac{b-a}{2} \sum_{k=1}^{n} w_k\, f\left(\frac{(b-a)}{2}\,x_k + \frac{b+a}{2}\right) + R_n}$$

siendo en éste caso el error

$$R_n = \frac{(b-a)^{2n+1}(n!)^4}{(2n+1)[(2n)!]^3} f^{(2n)}(\xi) \qquad (a < \xi < b)$$

Una subrutina que calcula los pesos y las raíces del polinomio de Legendre de grado n y que se puede encontrar en [PFTV86] es la siguiente :

```
      SUBROUTINE GAULEG(X1,X2,X,W,N)
      IMPLICIT REAL*8 (A-H,O-Z)
      REAL*4 X1,X2,X(N),W(N)
      PARAMETER (EPS=3.D-14)
      M=(N+1)/2
      XM=0.5D0*(X2+X1)
      XL=0.5D0*(X2-X1)
      DO 12 I=1,N
         Z=COS(3.141592654D0*(I-.25D0)/(N+.5D0))
 1    CONTINUE
         P1=1.D0
         P2=0.D0
         DO 11 J=1,N
            P3=P2
            P2=P1
            P1=((2.D0*J-1.D0)*Z*P2-(J-1.D0)*P3)/J
 11      CONTINUE
         PP=N*(Z*P1-P2)/(Z*Z-1.D0)
         Z1=Z
         Z=Z1-P1/PP
         IF (ABS(Z-Z1).GT.EPS) GO TO 1
         X(I)=XM-XL*Z
         X(N+1-I)=XM+XL*Z
         W(I)=2.D0*XL/((1.D0-Z*Z)*PP+PP)
         W(N+1-I)=W(I)
 12   CONTINUE
      RETURN
      END
```

Definición 6. *Los* **polinomios de Laguerre** *son una familia de polinomios* $\{\mathcal{L}_n(x)\}_{n=0}^{\infty}$ *ortogonales en el intervalo* $[0, +\infty]$ *respecto de la función peso* $w(x) = e^{-x}$ *y se definen recursivamente como :*

LAGUERRE

$$\begin{aligned}\mathcal{L}_n(x) &= (2n-x-1)\,\mathcal{L}_{n-1}(x) - (n-1)^2\,\mathcal{L}_{n-2}(x), \quad n \geqslant 2 \\ \mathcal{L}_0(x) &= 1\,; \quad \mathcal{L}_1(x) = 1-x\end{aligned}$$

Los primeros polinomios de Laguerre son :

$$\begin{aligned}
L_2(x) &= \frac{1}{2}(x^2-4x+2)\\
L_3(x) &= \frac{1}{6}(-x^3+9x^2-18x+6)\\
L_4(x) &= \frac{1}{24}(x^4-16x^3+72x^2-96x+24)\\
L_5(x) &= \frac{1}{120}(-x^5+25x^4-200x^3+600x^2-600x+120x)\\
L_6(x) &= \frac{1}{720}(x^6-36x^5+450x^4-2400x^3+5400x^2-4320x+720)\\
L_7(x) &= \frac{1}{5040}(-x^7+49x^6-882x^5+7350x^4-29400x^3+52920x^2-35280x+5040)
\end{aligned}$$

Tienen las siguientes propiedades :

Propiedad 8. $\mathcal{L}_n(x)$ *es un polinomio de grado* n.

Propiedad 9. $\{\mathcal{L}_0(x), \mathcal{L}_1(x), \cdots, \mathcal{L}_n(x)\}$ *es una base del espacio vectorial de los polinomios de grado menor o igual que* n. *Esto es,*
$\forall p(x) \in \mathbb{R}[x], \exists \alpha_i \in \mathbb{R}, \quad i = 0, 1, \ldots, n$ *tales que* $p(x) = \sum_{i=0}^{n} \alpha_i \mathcal{L}_i(x)$ *de forma única.*

Propiedad 10. *Se verifica que*

$$\int_0^{+\infty} \mathcal{L}_n(x)\mathcal{L}_m(x)dx = 0, \text{ para } n \neq m$$
$$\int_0^{+\infty} [\mathcal{L}_n(x)]^2 dx = 1$$

Propiedad 11. *En el caso de la integración de Gauss-Laguerre :*

$$w_k = \frac{(n!)^2}{x_k[\mathcal{L}'_n(x_k)]^2} = \frac{(n!)^2\,x_k}{(n+1)^2\,[\mathcal{L}_{n+1}(x_k)]^2}, \text{ (pesos)}$$

$$R_n = \frac{(n!)^2}{(2n)!} f^{(2n)}(\xi), (0 < \xi < +\infty) \text{ (error)}$$

La fórmula general para la integración de Gauss-Laguerre queda

$$\boxed{\int_0^{+\infty} e^{-x}\, f(x)\, dx = \sum_{i=1}^{n} w_i\, f(x_i) + R_n}$$

siendo x_i el i-ésimo cero de $L_n(x)$.

Definición 7. *Los* **polinomios de Hermite** *son una familia de polinomios* $\{H_n(x)\}_{n=0}^{\infty}$ *ortogonales en el intervalo* $[-\infty, +\infty]$ *respecto de la función peso* $w(x) = e^{-x^2}$ *y se definen recursivamente como :*

HERMITE

$$\begin{aligned} H_n(x) &= 2x\, H_{n-1}(x) - 2(n-1)\, H_{n-2}(x), \quad n \geqslant 2 \\ H_0(x) &= 1\,; \quad H_1(x) = 2x \end{aligned}$$

Los primeros polinomios de Hermite son :

$$\begin{aligned} &H_2(x) = 4x^2 - 2 && H_3(x) = 8x^3 - 12x \\ &H_4(x) = 16x^4 - 48x^2 + 12 && H_5(x) = 32x^5 - 160x^3 + 120x \\ &H_6(x) = 64x^6 - 480x^4 + 720x^2 - 120 && H_7(x) = 128x^7 - 1344x^5 + 3360x^3 - 1680x \end{aligned}$$

Tienen las siguientes propiedades :

Propiedad 12. $H_n(x)$ *es un polinomio de grado* n.

Propiedad 13. $\{H_0(x), H_1(x), \cdots, H_n(x)\}$ *es una base del espacio vectorial de los polinomios de grado menor o igual que* n. *Esto es,*
$\forall p(x) \in \mathbb{R}[x], \exists \alpha_i \in \mathbb{R},\ \ i = 0, 1, \ldots, n$ *tales que* $p(x) = \sum_{i=0}^{n} \alpha_i H_i(x)$ *de forma única.*

Propiedad 14. *Se verifica que*

$$\int_{-\infty}^{+\infty} H_n(x) H_m(x) dx = 0, \ \textit{para } n \neq m$$

$$\int_{-\infty}^{+\infty} [H_n(x)]^2 dx = \sqrt{\pi}\, 2^n\, n!$$

Propiedad 15. *En el caso de la integración de Gauss-Hermite :*

$$w_k = \frac{2^{n+1}.n!\sqrt{\pi}}{[H_n'(x_k)]^2} = \frac{2^{n-1}.n!\sqrt{\pi}}{n^2\,[H_{n-1}(x_k)]^2}, \ \textit{(pesos)}$$

$$R_n = \frac{(n!)\sqrt{\pi}}{2^n(2n)!} f^{(2n)}(\xi), (-\infty < \xi < +\infty) \ \textit{(Error)}$$

Por tanto, la integración de Gauss-Hermite queda :

$$\int_{-\infty}^{+\infty} e^{-x^2} f(x)\,dx = \sum_{i=1}^{n} w_i f(x_i) + R_n$$

siendo x_i la i-ésima raíz del polinomio $H_n(x)$.

Definición 8. *Los* **polinomios de Tchebychev** *son una familia de polinomios* $\{T_n(x)\}_{n=0}^{\infty}$ *ortogonales en el intervalo* $[-1,+1]$ *respecto de la función peso* $w(x) = \frac{1}{\sqrt{1-x^2}}$ *y se definen recursivamente como :*

TCHEBYCHEV

$$\begin{aligned} T_n(x) &= 2x\,T_{n-1}(x) - T_{n-2}(x), \quad n \geqslant 2 \\ T_0(x) &= 1\,; \quad T_1(x) = x \end{aligned}$$

Los primeros polinomios de Tchebychev son :

$$\begin{array}{ll} T_2(x) = 2x^2 - 1 & T_3(x) = 4x^3 - 3x \\ T_4(x) = 8x^4 - 8x^2 + 1 & T_5(x) = 16x^5 - 20x^3 + 5x \\ T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1 & T_7(x) = 64x^7 - 112x^5 + 56x^3 - 7x \end{array}$$

Tienen las siguientes propiedades :

Propiedad 16. $T_n(x)$ *es un polinomio de grado* n.

Propiedad 17. *Sus raíces están distribuidas simétricamente en el intervalo* $[-1,+1]$

Propiedad 18. $\{T_0(x), T_1(x), \cdots, T_n(x)\}$ *es una base del espacio vectorial de los polinomios de grado menor o igual que* n. *Esto es,*
$\forall p(x) \in \mathbb{R}[x], \exists \alpha_i \in \mathbb{R}, \ \ i = 0, 1, \ldots, n$ *tales que* $p(x) = \sum_{i=0}^{n} \alpha_i T_i(x)$ *de forma única.*

Propiedad 19. *Se verifica que*

$$\int_{-1}^{+1} T_n(x) T_m(x) dx = 0, \ \textit{para } n \neq m$$

$$\int_{-1}^{+1} [T_n(x)]^2 dx = \begin{cases} \frac{\pi}{2}, \pi/2 & \textit{si } n \neq 0 \\ \pi & \textit{si } n = 0 \end{cases}$$

Propiedad 20. *En el caso de la integración de Gauss-Tchebychev :*

$$x_k = \cos\left(\frac{(2k-1)\pi}{2n}\right), \quad k = 1, 2, \ldots, n$$
$$w_k = \frac{\pi}{n}, \ \textit{(pesos todos iguales)}$$

$$R_n = \frac{2\pi}{2^{2n}(2n)!} f^{(2n)}(\xi), (-1 < \xi < 1) \ \textit{(Error)}$$

Por tanto, la integración de Gauss-Tchebychev queda

$$\boxed{\int_{-1}^{+1} \frac{f(x)}{\sqrt{1-x^2}}\, dx = \frac{\pi}{n} \sum_{k=1}^{n} f\left(\cos\left(\frac{2k-1}{2n}\pi\right)\right) + \frac{2\pi}{2^{2n}(2n)!} f^{(2n)}(\xi)}$$

Ejemplo 31. GAUSS-LEGENDRE *Calcular usando tres puntos*

$$\int_1^3 \frac{dx}{x}$$

Solución : haciendo el cambio $y = x - 2$ tenemos

$$\int_{-1}^{+1} \frac{dy}{y+2} \simeq \frac{5}{9}\frac{1}{1.225403} + \frac{8}{9}\frac{1}{2} + \frac{5}{9}\frac{1}{2.774597} \simeq 1.098039$$

mientras el verdadero valor es $\ln 3 = 1.099612$.

El error es :

$$E = \frac{f^{(6)}(\xi)}{6!} \int_{-1}^{+1} [p_3(x)]^2 dx = \frac{4}{25} \int_{-1}^{+1} [p_3(x)]^2 dx = \frac{8}{175}$$
$$\Longrightarrow E = \frac{1}{15750}\frac{6!}{(\xi+2)^7} = \frac{8}{175}\frac{1}{(\xi+2)^7} \Longrightarrow 0.000021 \simeq \frac{8}{175}\frac{1}{(3)^7} < E < \frac{8}{175} \simeq 0.045714$$

Ejemplo 32. GAUSS-LAGUERRE *Calcular usando tres puntos*

$$\int_0^{+\infty} x^7 e^{-x} dx$$

Solución :

$$\int_0^{+\infty} x^7 e^{-x} dx \simeq (0.711093).(0.415775)^7 + (0.278518).(2.294280) + (0.010389).(6.289945)^7 = 4319.9$$

siendo el valor exacto $7! = 5040$.

El error es :

$$E =$$

Ejemplo 33. GAUSS-TCHEBYCHEV *Calcular con seis cifras decimales correctas*

$$\int_{-1}^{+1} \frac{e^x dx}{\sqrt{1-x^2}}$$

Solución :

$$E_n = \frac{2\pi}{2^{2n}(2n)!}e^{\eta}, \quad -1 < \eta < 1 \Longrightarrow |E_n| \leqslant \frac{2\pi e}{2^{2n}(2n)!} = B_n$$

Para $n = 4, B_n = 1.66 \times 10^{-6}$ y para $n = 5, B_n = 4.6 \times 10^{-9}$. Aplicando la fórmula para $n = 5$ obtenemos :

$$I \simeq \frac{\pi}{5} \sum_{j=1}^{5} \; exp\left(\cos\left(\frac{2j-1}{10}\pi\right)\right) = 3.977463$$

Bibliografía

[Act90] F.S. Acton. *Numerical Methods That (Usually) Work.* The Mathematical Association of America, Washington, 1990.

[Atk89] K. E. Atkinson. *An Introduction to Numerical Analysis.* John Wiley, New York, 2nd. edition, 1989.

[BF80] R.L. Burden and D. Faires. *Análisis Numérico.* Grupo Editorial Iberoamericana, México, 1980.

[CC89] S.C. Chapra and R.P. Canale. *Numerical Methods for Engineers.* McGraw-Hill International, New York, second edition, 1989.

[CdB80] S. D. Conte and C. de Boor. *Elementary Numerical Analysis: An Algorithmic Approach.* McGraw–Hill, New York, third edition, 1980.

[DB74] Germund Dahlquist and Åke Björck. *Numerical Methods.* Prentice-Hall, Englewood Cliffs, New Jersey, 1974.

[Frö79] C.-E. Fröberg. *Introduction to Numerical Analysis.* Adison–Wesley, Reading, Massachusetts, 2nd. edition, 1979.

[GW89] C.F. Gerald and P.O. Wheatley. *Applied Numerical Analysis.* Addison–Wesley Publishing Co., Reading, Massachusets, fourth edition, 1989.

[Hen72] P. Henrici. *Elementos de Análisis Numérico.* Ed. Trillas, México, 1972.

[Hil74] F. B. Hildebrand. *Introduction to Numerical Analysis.* McGraw–Hill, New York, second edition, 1974.

[KC94] D. Kincaid and W. Cheney. *Análisis Numérico : las matemáticas del cálculo científico.* Addison-Wesley Iberoamericana, 1994.

[ML91] M. J. Maron and R. J. Lopez. *Numerical Analysis: A Practical Approach*. Wadsworth, Belmont, California, third edition, 1991.

[PFTV86] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, New York, 1986. 3.11

[RR78] Anthony Ralston and Philip Rabinowitz. *A First Course in Numerical Analysis*. McGraw-Hill, New York, 2nd. edition, 1978.

[Sch89] H.R. Schwarz. *Numerical Analysis*. John Wiley & Sons, Chichester, 1989.

[YG73a] David M. Young and R.T. Gregory. *A Survey of Numerical Mathematics*, volume I. Dover Publications, New York, 1973.

[YG73b] David M. Young and R.T. Gregory. *A Survey of Numerical Mathematics*, volume II. Dover Publications, New York, 1973.