

[75.12] Análisis Numérico I

María Inés Parnisari

26 de abril de 2017

Índice

1. Errores	2
2. Sistemas de ecuaciones lineales	4
3. Ecuaciones de una variable	10
4. Sistemas de ecuaciones no lineales	13
5. Interpolación	14
6. Ajuste de funciones y mejor aproximación	18
7. Diferenciación numérica	18
8. Integración numérica	19
9. Ecuaciones diferenciales ordinarias	22

Algunas definiciones preliminares:

Análisis numérico: diseño de algoritmos para resolver problemas de matemática continua.

Error: incertidumbre tanto de los datos de ingreso como de los resultados obtenidos en un procedimiento. Está asociado a la aproximación.

Exactitud: error absoluto o relativo de un valor. No está limitada por la precisión.

Precisión: exactitud con que se llevan a cabo operaciones aritméticas. No siempre mejora una aproximación.

Orden de convergencia de un algoritmo: la velocidad con la cual una sucesión converge a su límite. Este concepto es, desde el punto de vista práctico, muy importante si necesitamos trabajar con secuencias de sucesivas aproximaciones de un método iterativo.

Supongamos que la secuencia $\{x_k\}$ converge al número ξ . Decimos que la sucesión converge con orden q a ξ si

$$\lim_{k \rightarrow \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^q} = \lambda \text{ con } \lambda > 0$$

El número q es llamado orden de convergencia. En particular, convergencia de orden 1 es llamada convergencia lineal, la de orden 2 convergencia cuadrática y la convergencia de orden 3 convergencia cúbica.



Alta exactitud, pero baja precisión.



Alta precisión pero baja exactitud.

Figura 1: Precisión y exactitud.

1 Errores

1.1 Fuentes de error

1.1.1 Error inherente

Error inherente: error de los datos de entrada que puede provenir de la precisión en la medición de los datos, la representación numérica, etc. Dado un valor exacto m y una aproximación \tilde{m} , definimos:

- **Error absoluto:** $|e_a| = |m - \tilde{m}|$
- **Error relativo:** $|e_r| = \frac{|e_a|}{|\tilde{m}|} = \frac{|m - \tilde{m}|}{|\tilde{m}|}$ (mejor aproximación que el error absoluto)

Número de condición / factor de amplificación: sea la función $y = f(x)$. El error relativo de y es $e_{r_y} \approx \underbrace{\frac{\partial f(x)}{\partial x} \cdot \frac{x}{f(x)}}_{C_P} \cdot e_{r_x}$. Al coeficiente que multiplica el error relativo de x lo llamamos *número de condición*.

Si la función es de dos o más variables, tendremos $C_{P_i} = \frac{\partial f(x)}{\partial x_i} \cdot \frac{x_i}{f(x)}$. Es decir:

$$C_P = \sum_i \left| \frac{\partial f(x)}{\partial x_i} \right| \cdot \frac{|x_i|}{|f(x)|}$$

Ejemplo: sea el algoritmo $z = a^2 + b^2$. Su número de condición está dado por

$$C_P = \left[\frac{\partial z}{\partial a} \frac{a}{z} \right] + \left[\frac{\partial z}{\partial b} \frac{b}{z} \right] = \left[2a \cdot \frac{a}{a^2 + b^2} \right] + \left[2b \cdot \frac{b}{a^2 + b^2} \right] = \frac{2a^2}{a^2 + b^2} + \frac{2b^2}{a^2 + b^2} = 2$$

1.1.2 Error de redondeo

Error de redondeo / corte: error debido a la representación numérica utilizada.

El error de redondeo surge por la notación de punto flotante:

$$N \cong d, d_1 d_2 \dots d_t \times \beta^p \text{ con } \begin{cases} t \text{ cantidad de dígitos} \\ \beta \text{ base} \\ p \text{ exponente} \end{cases}$$

Corte: forma de redondear que consiste en “borrar” dígitos.

Redondeo: forma de redondear que consiste en modificar el último dígito según el siguiente criterio:

$$d_k = \begin{cases} d_k & \text{si } 0 \leq d_{k+1} \leq 4 \\ d_k + 1 & \text{si } 5 \leq d_{k+1} \leq 9 \end{cases}$$

La cota para el error de redondeo es:

	Para números de la forma $d, d_1 d_2 \dots d_k$	Para números de la forma $0, d_1 d_2 \dots d_k$
Corte	10^{-t}	10^{1-t}
Redondeo	$\frac{1}{2} \cdot 10^{-t}$	$\frac{1}{2} \cdot 10^{1-t}$

Término de estabilidad: nos da una idea de la influencia del error de redondeo.

El error relativo por el error de redondeo es $e_r = \sum_{i=1}^k |T(x)| \mu_i = T_e \mu$

1.1.3 Otras fuentes de error

Error de truncamiento / discretización: error que aparece al transformar un procedimiento infinito en uno finito.

Error del modelo matemático: se debe a las simplificaciones e hipótesis introducidas para definir el modelo matemático que representa el modelo físico.

Error humano: error producido por la intervención humana y/o por programas de computación mal hechos.

1.1.4 Error total

$$e_{r_T} = C_P \cdot r + T_e \cdot \mu$$

1.2 Algoritmos

Factores de amplificación de errores de entrada: $y = x_1(op)x_2 \implies r_y = f_1 x_1 + f_2 x_2$

op	+	-	\times	\div
f_1	$\frac{x_1}{y}$	$\frac{x_1}{y}$	1	1
f_2	$\frac{x_2}{y}$	$-\frac{x_2}{y}$	1	-1

Algoritmo bien condicionado: un problema matemático está bien condicionado cuando al introducir pequeñas variaciones en los datos de entrada se producen pequeñas variaciones en los resultados.

$C_P \gg 1 \implies$ problema mal condicionado	$C_P \not\gg 1$ problema bien condicionado
--	--

Si f representa al algoritmo “real” y f^* al algoritmo “computacional”, y x a la variable “real” y x^* a la variable “computacional” entonces el error en los resultados se define como:

$$|f(x) - f^*(x^*)| \leq |f(x) - f(x^*)| + |f(x^*) - f^*(x)| + |f^*(x) - f^*(x^*)|$$

Algoritmo estable: tiene la propiedad de que la propagación de los errores de redondeo es lineal o cuasi-lineal.

$$E_n \approx c \cdot n \cdot E_0$$

$T_e \gg 1 \implies$ algoritmo inestable	$T_e \not\gg 1 \implies$ algoritmo estable
--	--

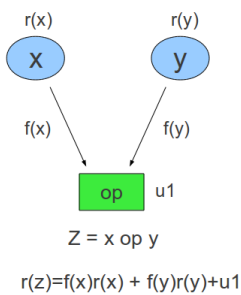
Para obtener algoritmos estables se recomienda:

1. Evitar la resta de números con errores
2. Minimizar el tamaño de los números intermedios relativos al resultado final
3. $valor_{nuevo} = valor_{viejo} + corrección$
4. Usar transformaciones bien condicionadas

Resumiendo:

$1 + \frac{T_e}{C_P} \gg 1 \implies$ mal algoritmo	$1 + \frac{T_e}{C_P} \not\gg 1 \implies$ buen algoritmo
--	---

1.3 Gráfica de proceso y perturbaciones experimentales



Una forma de obtener los coeficientes T_e y C_P es la **gráfica de proceso**: un diagrama de flujo que representa la propagación de los errores relativos y de redondeo.

La gráfica de proceso no siempre es útil y práctica. También podemos estimar el C_P y el T_e mediante **perturbaciones experimentales**:

- Si suponemos $T_e \approx 0 \implies C_P = \frac{e_r}{r}$
- Si suponemos $r \approx 0 \implies T_e = \frac{y_\mu - y_\xi}{y_\xi(\mu - \xi)}$

1.3.1 Estimación del C_p

Consiste en suponer que no existen errores de redondeo y perturbar los valores de entrada. La idea es: se toman los datos de entrada y se aplica el algoritmo a analizar, obteniendo el resultado “exacto” x . Luego se “perturban” los datos de entrada introduciéndoles un error $\pm k$ y se vuelven a calcular los resultados, obteniéndose \hat{x}_1 y \hat{x}_2 . Entonces se calculan los C_p :

- $C_{p1} = \left| \frac{x - \hat{x}_1}{x} \cdot \frac{1}{+k} \right|$
- $C_{p2} = \left| \frac{x - \hat{x}_2}{x} \cdot \frac{1}{-k} \right|$

Y finalmente tenemos que $C_P = \max\{C_{p1}, C_{p2}\}$. Como hemos supuesto que los errores de redondeo son despreciables, los C_p deberían ser similares, con lo cual tendremos una estimación de la condición del problema.

1.3.2 Estimación del T_e

Consiste en suponer que no existen errores inherentes.

Ejemplo: aproximamos la función $f(x) = \sin(x)$ con su serie de MacLaurin $g(x) = x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} + \frac{x^9}{362880}$. Calculemos el valor de $f(\frac{\pi}{4})$ con tres precisiones utilizando $g(x)$:

- $u = 15 \rightarrow \mu_u = 10^{-14} \rightarrow \bar{y}_u = 0,70710678293687$
- $t = 8 \rightarrow \mu_t = 10^{-7} \rightarrow \bar{y}_t = 0,7071068$
- $s = 4 \rightarrow \mu_s = 10^{-3} \rightarrow \bar{y}_s = 0,706$

Tomando como valor “exacto” a \bar{y}_u tenemos que:

- $T_{e_s} = \frac{\bar{y}_u - \bar{y}_s}{\bar{y}_u(\mu_s - \mu_u)} = 1,565$
- $T_{e_t} = \frac{\bar{y}_u - \bar{y}_t}{\bar{y}_u(\mu_t - \mu_u)} = 0,241$

En este caso concluimos que calcular $f(y)$ con más precisión mejora el resultado final, pero esto no siempre es cierto.

2 Sistemas de ecuaciones lineales

Sea el sistema de ecuaciones dado por

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_n \end{cases} \implies Ax = B \implies \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Si A es una matriz cuadrada tal que $\det(A) \neq 0$, podemos resolver el sistema mediante $x = A^{-1}B$. Sin embargo, los coeficientes de la matriz A puede provenir de mediciones imprecisas, lo que haría que calcular A^{-1} resulte en más errores. Por este motivo surgen alternativas a esta forma de resolución del sistema.

Norma infinito de una matriz A : es la máxima suma absoluta de las filas de una matriz.

$$\|A\|_{\infty} = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|$$

Número de condición de una matriz A : se define como $\kappa(A) = \|A\|_{\infty} \|A^{-1}\|_{\infty}$. Puede estimarse mediante $\frac{\|x - \hat{x}\|_{\infty}}{\|\hat{x}\|_{\infty}} \cdot 10^t \leq \kappa(A)$. El error relativo de \hat{x} al resolver $Ax = B$ depende de $\kappa(A)$.

Matriz bien condicionada: si $\kappa(A)$ es chico entonces la matriz A es bien condicionada, y no hay alta propagación de errores de redondeo al resolver $Ax = B$. Caso contrario, A es **mal condicionada**. Generalmente, las matrices mal condicionadas mezclan coeficientes chicos con coeficientes grandes, o satisfacen $\det(A) \approx 0$. Si A es tal que $\exists A^{-1}$ entonces $\kappa(A) = \infty$.

2.1 Sustitución directa e inversa

Dada una matriz A que es triangular superior o triangular inferior, existen dos métodos para resolver el sistema $Ax = B$.

Si A es triangular inferior, la **sustitución directa** consiste en:

$$x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij} x_j}{l_{ii}}$$

Si A es triangular superior, la **sustitución inversa** consiste en:

$$x_i = \frac{b_i - \sum_{j=i+1}^n u_{ij} x_j}{u_{ii}}$$

2.2 Métodos directos

Cuando queremos resolver el sistema $Ax = B$ y la matriz A tiene en su mayoría coeficientes no nulos (**matriz densa**) podemos aplicar un método directo para resolverlo. Estos métodos consisten en transformar la matriz original A en una matriz triangular y luego aplicar sustitución directa o inversa.

Ventajas	Desventajas
- sólo hay errores de redondeo	- no resuelven bien matrices mal condicionadas
- el número de operaciones es finito	- se vuelven inestables con matrices grandes

2.2.1 Método de eliminación de Gauss

Consiste en transformar la matriz A de $Ax = B$ en una matriz triangular superior mediante operaciones entre filas y luego aplicar sustitución inversa.

- Eliminación con pivoteo parcial: es necesario intercambiar de lugar filas. Es del orden $O(cn^3)$.
- Eliminación con pivoteo total: es necesario intercambiar de lugar filas y columnas.

Convergencia: cuando A es diagonal dominante, es decir $|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}|$

Ventajas	Desventajas
- fácil de programar	- si cambia el vector B es necesario retriangular A .
	- $O(n^3)$

2.2.2 Factorización LU; métodos de Doolittle y Crout

La factorización LU consiste en transformar la matriz A en $A = \underbrace{L}_{\Delta} \underbrace{U}_{\nabla}$. Entonces el sistema $Ax = B$ se convierte en:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ m_{2,1} & 1 & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ m_{n,1} & \cdots & m_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} a_{11}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & & \vdots \\ \vdots & \ddots & \\ 0 & \cdots & 0 & a_{nn}^{(n)} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \mapsto \begin{cases} Ux = y & (I) \\ Ly = B & (II) \end{cases}$$

donde $m_{i,j} = \frac{a_{ji}^{(i)}}{a_{ii}^{(i)}}$ son los coeficientes de la triangulación de Gauss. El sistema (I) se resuelve aplicando sustitución inversa y para (II) sustitución directa.

Teorema: si se puede efectuar la eliminación de Gauss en el sistema lineal $Ax = B$ sin intercambios de filas, entonces A se puede factorizar en el producto LU . Dos ejemplos de estos tipos de matrices son: estrictamente diagonal dominantes, definidas positivas.

Ventajas	Desventajas
- si cambia el vector B no hay que refactorizar A	- determinar L y U requiere $O(n^3)$ pasos
	- resolver $Ax = B$ requiere $O(n^2)$ pasos
	- no sirve para todas las A

Método de Doolittle: consiste en factorizar A asumiendo que $l_{ii} = 1$.

$$A = LU = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & \cdots & u_{1n} \\ 0 & \ddots & \vdots \\ \vdots & \ddots & \\ 0 & \cdots & 0 & u_{nn} \end{pmatrix}$$

La fórmula general es:

- $l_{ji} = \frac{1}{u_{ii}} \left[a_{ji} - \sum_{k=1}^{i-1} l_{jk} u_{ki} \right]$
- $u_{ii} = a_{ii} - \sum_{k=1}^{i-1} l_{ik} u_{ki}$
- $u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}$

Método de Crout: consiste en factorizar A asumiendo que $u_{ii} = 1$.

2.2.3 Método de Cholesky

Si A es una matriz simétrica y definida positiva, entonces planteamos que $A = \underbrace{S}_{\Delta} \underbrace{S^T}_{\nabla}$. La fórmula es:

- $s_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} s_{ik}^2}$
- $s_{ji} = \frac{1}{s_{ii}} \left[a_{ji} - \sum_{k=1}^{i-1} s_{jk} s_{ik} \right]$ para $i \neq j$

Ventajas	Desventajas
- menor cantidad de cálculos que Crout y Doolittle	- $O(n^3)$

2.3 Métodos iterativos estacionarios

Supongamos el sistema $Ax = B$. Tenemos que:

$$\begin{aligned} B - Ax &= 0 \\ B - Ax + Px &= Px \\ B + (-A + P)x &= Px \\ x &= (I - P^{-1}A)x + P^{-1}B \\ x^{(i+1)} &= \underbrace{(I - P^{-1}A)x^{(i)}}_T + \underbrace{P^{-1}B}_C \end{aligned}$$

Los métodos iterativos:

- Son muy útiles para resolver sistemas con matrices **rafas** (con muchos coeficientes nulos). También son útiles para resolver sistemas con matrices mal condicionadas (pero si $\kappa(A) \gg 10^t$, siendo t la precisión utilizada para realizar los cálculos, entonces debe aumentarse t para obtener una solución aceptable). Y por último, son útiles cuando A es grande.

- Para que el método converja a la única solución con cualquier $x^{(0)}$, T debe ser tal que $\rho(T) < 1$, o lo que es lo mismo $\|T\| < 1$ para cualquier norma natural. Idealmente, debería ser $\|T\| \ll 1$.
- Nunca dan una solución exacta, incluso si usamos una precisión infinita.

Refinamiento iterativo: es una técnica que, dada una solución obtenida por el método de Gauss, factorización LU o Cholesky, mejora el resultado.

$$Ax = B \implies A\tilde{x}_0 = \tilde{B} \implies B - A\tilde{x}_0 = r_0 \implies Ax - A\tilde{x}_0 = r_0 \implies \underbrace{A(x - \tilde{x}_0)}_{d_0} = r_0 \implies Ad_0 = r_0 \implies \tilde{d}_0 + \tilde{x}_0 = \tilde{x}_1 \implies \dots$$

Cuando la matriz A es muy mal condicionada, la primera solución obtenida mediante un método directo puede no estar muy cerca de la solución correcta. Entonces podemos usar el refinamiento iterativo:

$$x = x^{(1)} + \sum_{i=1}^{n \rightarrow \infty} d^{(i)}$$

Un criterio de corte aceptado para parar las iteraciones es:

$$\frac{|x^{(n+1)} - x^{(n)}|}{|x^{(n+1)}|} \leq TOL$$

Y además, puesto que $\frac{|x^{(n+1)} - x^{(n)}|}{|x^{(n+1)}|} \leq \kappa(A) \frac{\|R\|}{\|B\|}$, si A es mal condicionada la tolerancia debe ser más chica.

Los siguientes métodos se llaman **métodos iterativos estacionarios** porque las matrices T y C no se modifican.

2.3.1 Método de Jacobi

Dado el sistema inicial $Ax = B$, el **método de Jacobi** consiste en despejar de cada ecuación las incógnitas:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_n \end{cases} \implies \begin{cases} x_1 = \frac{b_1 - a_{12}x_2 - \dots - a_{1n}x_n}{a_{11}} \\ \vdots \\ x_n = \frac{b_n - a_{n1}x_1 - \dots - a_{nn-1}x_{n-1}}{a_{nn}} \end{cases} \implies x_i^{(i+1)} = \frac{b_i - \sum_{k=1, k \neq i}^n a_{ik}x_k^{(i)}}{a_{ii}}$$

En forma matricial esto es

$$x^{(i+1)} = Tx^{(i)} + C = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1,n}}{a_{11}} \\ & \ddots & & \vdots \\ \vdots & & \ddots & \\ -\frac{a_{n,1}}{a_{n,n}} & \dots & -\frac{a_{n,n-1}}{a_{n,n}} & 0 \end{pmatrix} x^{(i)} + \begin{pmatrix} \frac{b_1}{a_{11}} \\ \vdots \\ \frac{b_n}{a_{n,n}} \end{pmatrix}$$

siendo:

- $P = D$
- $T = -D^{-1}(L + U)$
- $C = D^{-1}B$

Convergencia: cuando A es estrictamente diagonal dominante, es decir $|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}|$

2.3.2 Método de Gauss-Seidel

El **método de Gauss-Seidel** es similar al método de Jacobi pero más rápido, con la diferencia de que en el cálculo de x_i ($i > 0$) se aprovecha lo calculado en el paso anterior:

$$x_i^{(i+1)} = \frac{b_i - \sum_{k=1}^{i-1} a_{ik}x_k^{(i+1)} - \sum_{k=1}^n a_{ik}x_k^{(i)}}{a_{ii}}$$

En forma matricial esto es $x^{(i+1)} = Tx^{(i)} + C$ siendo:

- $P = D + L$
- $T = I - (D + L)^{-1}A$
- $C = (D + L)^{-1}B$

Convergencia: cuando A es estrictamente diagonal dominante, es decir $|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}|$, o bien cuando A es simétrica y definida positiva.

2.3.3 Métodos de las sobrerelajaciones (SOR)

El **método de las sobrerelajaciones** es una variante del método de Gauss-Seidel, pero que converge más rápido.

$$x^{(i+1)} = (1 - \omega)x^{(i)} + \omega \underbrace{\left(\frac{b_i - \sum_{k=1}^{i-1} a_{ik}x_k^{(i+1)} - \sum_{k=1}^n a_{ik}x_k^{(i)}}{a_{ii}} \right)}_{x_{GAUSS-SEIDEL}}$$

En forma matricial esto es $x^{(i+1)} = (1 - \omega)x^{(i)} + \omega D^{-1}(B - Lx^{(i+1)} - Ux^{(i)})$ siendo:

- $P = \frac{1}{\omega}D + L$
- $0 < \omega < 2$ (notar que si $\omega = 1$ estamos en el método de Gauss-Seidel)

Convergencia: cuando A es simétrica ($A = A^T$) y definida positiva ($x^T A x > 0 \forall x \neq 0$).

Ventajas	Desventajas
	- obtener ω no es fácil

Teorema: si A es una matriz tribanda, es decir:

$$A = \begin{bmatrix} a_{11} & a_{12} & & & 0 \\ & \ddots & \ddots & & \\ a_{21} & & \ddots & \ddots & \\ & \ddots & & \ddots & a_{n-1,n} \\ 0 & & & a_{n,n-1} & a_{nn} \end{bmatrix}$$

entonces:

$$\omega_{optimo} = \frac{2}{1 + \sqrt{1 - \rho(T_j)^2}}$$

donde $\rho(T_j)$ es el mayor autovalor de la matriz T obtenida mediante el método de Jacobi.

2.4 Métodos iterativos no estacionarios

En estos métodos se aplica $x^{(i+1)} = Tx^{(i)} + C$, y las matrices T y C varían en cada iteración.

2.4.1 Método de los residuos mínimos

El **método de los residuos mínimos** consiste en buscar que $R^{(i+1)}$ sea mínimo en cada iteración. Su fórmula es:

$$x^{(i+1)} = x^{(i)} + \alpha_i R^{(i)}$$

Siendo:

- $R^{(i+1)} = R^{(i)} - \alpha_i A R^{(i)}$
- $\alpha_i = \frac{[A R^{(i)}]^T R^{(i)}}{[A R^{(i)}]^T A R^{(i)}}$

- $R^{(0)} = B - Ax^{(0)}$

Las iteraciones finalizan cuando $R^{(i+1)} < TOL$.

Convergencia: cuando A es simétrica y definida positiva.

Ventajas	Desventajas
- fácil de programar	- convergencia similar a Jacobi

2.4.2 Método del descenso más rápido

El **método del descenso más rápido** determina un mínimo local para una función de varias variables de la forma $g : \mathbb{R}^n \rightarrow \mathbb{R}$. Intuitivamente puede describirse así:

1. Evalúe g en una aproximación inicial $x^{(0)}$.
2. Determine una dirección desde $x^{(0)}$ que origine una disminución en el valor de g .
3. Desplace una cantidad apropiada hacia esta dirección y llame al nuevo vector $x^{(1)}$.
4. Repita los pasos 1 a 3 reemplazando $x^{(0)}$ con $x^{(1)}$.

La fórmula general es:

$$x^{(i+1)} = x^{(i)} + \alpha_i R^{(i)}$$

siendo:

- $R^{(0)} = B - Ax^{(0)}$
- $\alpha_i = \frac{R^{(i)T} R^{(i)}}{R^{(i)T} A R^{(i)}}$
- $R^{(i+1)} = R^{(i)} - \alpha_i A R^{(i)}$

Convergencia: cuando A es simétrica y definida positiva.

Ventajas	Desventajas
- converge independientemente del $x^{(0)}$	- convergencia lineal
- fácil de programar	- puede converger en algo que no es el mínimo de g
- mejor velocidad que MRM	- no usa bien las direcciones más empinadas
- menos cantidad de operaciones	

2.4.3 Método de los gradientes conjugados

El **método del gradiente conjugado** es muy útil como método iterativo de aproximación para resolver sistemas de gran tamaño con entradas no nulas. Supondremos que A es simétrica y definida positiva.

El cálculo multivariable nos dice que la dirección de máximo descenso en el valor de $g(x)$ es la dirección dada por $-\nabla g(x)$; es decir, en la dirección del residual $r = b - Ax$.

$$x^{(i+1)} = x^{(i)} + \alpha_i d^{(i)}$$

con:

- $d^{(0)} = R^{(0)} = b - Ax^{(0)}$
- $\alpha_i = \frac{R^{(i)T} R^{(i)}}{d^{(i)T} A d^{(i)}}$
- $R^{(i+1)} = R^{(i)} - \alpha_i A d^{(i)}$
- $\beta^{(i+1)} = \frac{R^{(i+1)T} R^{(i+1)}}{R^{(i)T} R^{(i)}}$
- $d^{(i+1)} = R^{(i+1)} + \beta^{(i+1)} \cdot d^{(i)}$

Convergencia: cuando A es simétrica y definida positiva.

Ventajas	Desventajas
- convergencia muy rápida para matrices grandes	- suele ser inestable
- Si $A \in R^{n \times n}$, n iteraciones alcanzan	
- Si $A \in R^{n \times n}$ tiene k autovalores repetidos, $n - k$ iteraciones alcanzan	

Si la matriz A es mal condicionada, este método es altamente susceptible a los errores de redondeo. Es por esto que existe el **método de gradientes conjugados preconditionado**. Un ejemplo de su uso es, en vez de resolver $Ax = B$, podemos resolver $A^T Ax = A^T B$. Notar que en este caso $\tilde{A} = A^T A$ es simétrica y definida positiva.

3 Ecuaciones de una variable

Sea el problema $f(x) = 0$. Resolver esta ecuación equivale a encontrar las raíces de la función f . Sin embargo, no siempre es posible despejar x e igualar a 0. Una forma de obtener la solución sería graficando la función y viendo donde ésta corta al eje X . Pero esto es poco práctico en casos complejos.

Si tenemos como datos a $f(x)$, una tolerancia máxima y un intervalo inicial $[a, b]$, podemos utilizar los **métodos de arranque** (bisección y Regula Falsi) para achicar este intervalo, y luego utilizar otros métodos para aproximar la solución.

Teorema del valor intermedio: sea f una función continua en un intervalo $[a, b]$ y supongamos que $f(a) < f(b)$. Entonces para cada u tal que $f(a) < u < f(b)$, existe al menos un c dentro de (a, b) tal que $f(c) = u$.

3.1 Método de la bisección

Método de la bisección: se basa en el teorema del valor intermedio y en la búsqueda binaria. Si f es una función continua definida en $[a, b]$ con $f(a)$ y $f(b)$ con signos diferentes, entonces existe $p \in [a, b]$ tal que $f(p) = 0$. El método requiere dividir varias veces a la mitad el intervalo $[a, b]$ y, en cada paso, localizar a la mitad que contenga a p . Conviene elegir el intervalo inicial lo más pequeño posible.

Ejemplo: supongamos el intervalo inicial $[a_1, b_1]$. Luego $p_1 = \frac{a_1 + b_1}{2}$. Si $f(p_1) = 0$ hemos llegado a la solución. Si $f(a_1)$ y $f(p_1)$ tienen el mismo signo, el nuevo intervalo será $[p_1, b_1]$. Si $f(a_1)$ y $f(p_1)$ tienen distinto signo, el nuevo intervalo es $[a_1, p_1]$.

Teorema: sea $f \in C[a, b]$ y $f(a) \cdot f(b) < 0$. El método de la bisección genera una sucesión p_1, \dots, p_n que aproxima a un cero de f tal que:

$$|p_n - p| \leq \frac{b - a}{2^n} \implies n \geq \frac{\ln(b - a) - \ln(TOL)}{\ln(2)}$$

En este caso, tenemos que la condición de corte es:

$$TOL \leq \frac{|b_0 - a_0|}{2^{n+1}}$$

Es importante destacar que para elegir el intervalo donde se encuentra la raíz debe verificarse que $f(a) \cdot f(b) < 0$. Sin embargo, si esto no se cumple no implica que la raíz no exista (pues puede ser una raíz doble).

Ventajas	Desventajas
- fácil de entender	- convergencia lenta
- converge siempre	- podemos rechazar una buena aproximación intermedia

3.2 Método del de punto fijo o de las aproximaciones sucesivas

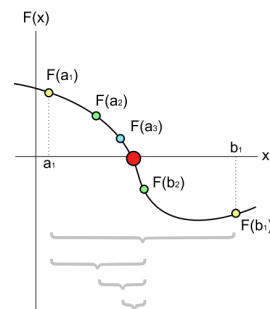
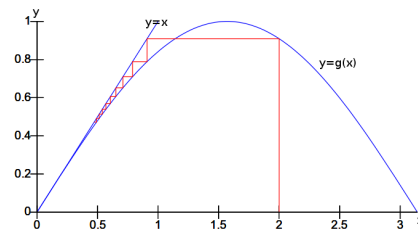


Figura 2: Bisección aplicada en el rango $[a_1, b_1]$.

Un **punto fijo** de una función g es un número p que satisface $g(p) = p$.

Teorema de suficiencia para la existencia y unicidad del punto fijo:

1. Si $g \in C[a, b]$ y $g(x) \in [a, b]$ para toda $x \in [a, b]$ entonces g tiene un punto fijo en $[a, b]$.
2. Si $g'(x)$ existe en (a, b) y existe una constante $0 < k < 1$ con $|g'(x)| \leq k$ para toda $x \in (a, b)$ entonces el punto fijo en $[a, b]$ es único.



Método de iteración de punto fijo: supongamos que deseamos resolver $f(x) = 0$. Para hallar el x tal que $g(x) = x$ escogemos una aproximación inicial x_0 y generamos la sucesión x_1, \dots, x_n haciendo

Figura 3: Método del punto fijo.

$$x_{i+1} = g(x_i) \text{ para } i \geq 0$$

Si la secuencia converge en x y g es continua, entonces obtenemos una solución con $x = g(x)$. El error está dado por:

$$|x_n - x| \leq k^n |b - a|$$

Ventajas	Desventajas
- si $k \approx 0$ converge rápidamente	- convergencia lineal
	- si $k \approx 1$ converge lentamente

3.3 Método de Newton-Raphson

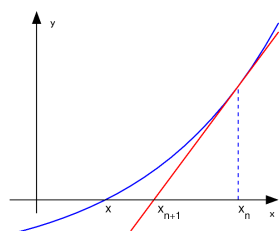


Figura 4: Una iteración del método de Newton-Raphson.

El **método de Newton-Raphson** es una de las técnicas más poderosas que existen para resolver un problema del tipo $f(x) = 0$. Este método comienza con una aproximación inicial x_0 y genera la sucesión x_1, \dots, x_n definida por:

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

El método obtiene las aproximaciones utilizando tangentes sucesivas. Comenzando con x_0 , la aproximación x_1 es la intersección entre la línea tangente a la gráfica de f en $(x_0, f(x_0))$ y el eje X . La aproximación x_2 es la intersección entre la tangente a la gráfica de f en $(x_1, f(x_1))$ y el eje X , y así sucesivamente.

Es importante notar que no es posible continuar con este método si $f'(x_{n-1}) = 0$ para alguna n . Además, $f''(x)$ debe estar acotada.

La deducción de este método proviene del desarrollo por Taylor:

$$\begin{aligned} f(x) &= f(\bar{x}) + (x - \bar{x})f'(\bar{x}) + \dots = 0 \\ (x - \bar{x})f'(\bar{x}) &= -f(\bar{x}) \\ x &= \bar{x} - \frac{f(\bar{x})}{f'(\bar{x})} \end{aligned}$$

Teorema de convergencia: sea $f \in C^2[a, b]$. Si $x \in [a, b]$ es tal que $f(x) = 0$ y $f'(x) \neq 0$, entonces existe $\delta > 0$ tal que el método de Newton genera una sucesión x_1, \dots, x_n que converge a p para cualquier aproximación inicial $x_0 \in [x - \delta, x + \delta]$.

Ventajas	Desventajas
- muy poderoso	- necesitamos conocer la derivada de f
- convergencia cuadrática para raíces simples	- necesita una buena aproximación inicial
	- convergencia lineal para raíces dobles

3.4 Método de la secante

Si queremos evitarnos el problema de evaluar la derivada en el método de Newton, podemos aproximarla mediante la definición: $f'(x_i) \approx \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$. Entonces tenemos:

$$x_{i+1} = x_i - \frac{f(x_i)(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}$$

La técnica que se utiliza en esta fórmula se llama **método de la secante**. Comenzando con las dos aproximaciones iniciales x_0 y x_1 , la aproximación x_2 es la intersección entre la línea que une $(x_0, f(x_0))$ y $(x_1, f(x_1))$ con el eje X .

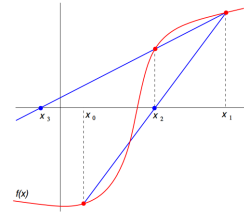


Figura 5: Método de la secante.

Ventajas	Desventajas
- es más rápido que el método del punto fijo	- es más lento que el método de Newton
- fácil de usar	
- no necesitamos conocer f'	

3.5 Método de la posición falsa / Regla falsa

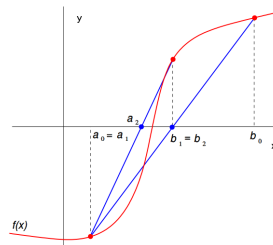


Figura 6: Método de la *regula Falsi*.

El **método de la posición falsa** combina el método de bisección y el método de la secante, pero ofrece una prueba para asegurarnos que la raíz queda entre dos iteraciones sucesivas.

Primero elegimos p_0 y p_1 tales que $f(p_0) \cdot f(p_1) < 0$. p_2 será la intersección en X de la recta secante. Luego calculamos $f(p_1) \cdot f(p_2)$, si es mayor a 0 la raíz está entre p_0 y p_1 , sino está entre p_1 y p_2 . Su fórmula general es:

$$p_n = a_n - \frac{f(a_n)(b_n - a_n)}{f(b_n) - f(a_n)} \quad \text{ó} \quad p_n = b_n - \frac{f(b_n)(b_n - a_n)}{f(b_n) - f(a_n)}$$

Ventajas	Desventajas
- converge siempre	- requiere más cálculos que el método de la secante
- más rápido que el método de la bisección	- converge lentamente
	- puede fallar si el denominador se anula

3.6 Método Δ^2 de Aitken

Diferencias progresivas: dada la sucesión p_0, \dots, p_n la diferencia progresiva Δp_n está dada por $\Delta p_n = p_{n+1} - p_n$ con $n \geq 0$.

Supongamos que p_0, \dots, p_n es una sucesión linealmente convergente con un límite p . El **método Δ^2 de Aitken** se basa en la suposición de que la sucesión $\hat{p}_0, \dots, \hat{p}_n$ definida por

$$\hat{p}_n = p_n - \frac{(p_{n+1} - p_n)^2}{p_{n+2} - 2p_{n+1} + p_n} = p_n - \frac{(\Delta p_n)^2}{\Delta^2 p_n}$$

converge más rápidamente a p que la sucesión original.

Ventajas	Desventajas
- convergencia acelerada	- poco práctico para programar
- no evaluamos la derivada	

3.7 Método de Steffensen

Existe una mejora para el método de Aitken, que se basa en aplicar este método para una sucesión dada por aproximaciones sucesivas, y se conoce como **método de Steffensen**.

La sucesión que genera está dada por:

$$\begin{matrix} p_0^{(0)} & p_0^{(1)} & p_0^{(2)} & \cdots \\ p_1^{(0)} & p_1^{(1)} & p_1^{(2)} & \cdots \\ p_2^{(0)} & & & \end{matrix}$$

Ventajas	Desventajas
- Convergencia cuadrática	

Algoritmo 1 Método de Steffensen

```
función steffensen(p0,TOL,itermax,g(p)):
    iter = 1
    while (iter < itermax):
        p1 = g(p0)
        p2 = g(p1)
        p = p0 - (p1 - p0)^2/(p2 - 2p1 + p0)
        if (|p - p0| < TOL):
            return p
        else:
            iter = iter + 1
            p0 = p
    return 'El método fracasó después de'+itermax+'iteraciones.'
```

4 Sistemas de ecuaciones no lineales

Un sistema de ecuaciones no lineales tiene la forma

$$\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ f_2(x_1, \dots, x_n) = 0 \\ \vdots \\ f_n(x_1, \dots, x_n) = 0 \end{cases} \implies \mathbf{F}(x_1, \dots, x_n) = (f_1, \dots, f_n)^T = \mathbf{0}$$

4.1 Adaptación del método de Jacobi

Ejemplo: supongamos el sistema dado por
$$\begin{cases} f(x_1, x_2) = x_1^2 + x_2^2 - 1 = 0 \\ g(x_1, x_2) = x_1^2 - 2x_1 - x_2 + 1 = 0 \end{cases}$$

Podemos resolverlo mediante una adaptación del método de Jacobi a sistemas de ecuaciones mediante:

$$\begin{aligned} x_1^{(i+1)} &= \pm \sqrt{1 - (x_2^{(i)})^2} \\ x_2^{(i+1)} &= (x_1^{(i)})^2 - 2x_1^{(i)} + 1 \end{aligned}$$

4.2 Adaptación del método de Gauss-Seidel

Ejemplo: supongamos el sistema dado por
$$\begin{cases} f(x_1, x_2) = x_1^2 + x_2^2 - 1 = 0 \\ g(x_1, x_2) = x_1^2 - 2x_1 - x_2 + 1 = 0 \end{cases}$$

Podemos resolverlo mediante una adaptación del método de Gauss-Seidel a sistemas de ecuaciones mediante:

$$x_1^{(i+1)} = \pm \sqrt{1 - (x_2^{(i)})^2}$$

$$x_2^{(i+1)} = (x_1^{(i+1)})^2 - 2x_1^{(i+1)} + 1$$

4.3 Método de Newton

El **método de Newton** para sistemas no lineales consiste en

$$x^{(i+1)} = x^{(i)} - J(x^{(i)})^{-1} \cdot F(x^{(i)})$$

siendo $J(x)$ la matriz jacobiana de $F(x)$.

Ventajas	Desventajas
- convergencia cuadrática	- debe conocerse un valor inicial bueno
	- necesita calcular e invertir $J(x)$ en cada paso
	- $O(n^3)$ pasos

4.4 Método de Broyden

El **método de Broyden** es una variante del método de Newton. En este método se reemplaza la matriz jacobiana con una matriz de aproximación que se actualiza en cada iteración.

Supongamos que se da una aproximación inicial $x^{(0)}$ a la solución p de $F(x) = 0$. Calculamos $x^{(1)}$ como en el método de Newton, y para calcular $x^{(2)}$ nos apartamos de Newton y examinamos el método de la secante para una ecuación no lineal: reemplazamos la matriz $J(x^{(1)})$ por una matriz A que tiene la propiedad de que $A_1(x^{(1)} - x^{(0)}) = F(x^{(1)}) - F(x^{(0)})$, es decir que $x^{(2)} = x^{(1)} - A_1^{-1}F(x^{(1)})$. En general, calculamos:

$$x^{(i+1)} = x^{(i)} - A_i^{-1}F(x^{(i)})$$

siendo

$$A_i = A_{i-1} + \frac{F(x^{(i)}) - F(x^{(i+1)}) - A_{i-1}(x^{(i)} - x^{(i+1)})}{\|x^{(i)} - x^{(i+1)}\|_2^2}$$

Ventajas	Desventajas
- no hay que invertir matrices	- hay que invertir una matriz al comienzo
- no hay que calcular el jacobiano	- convergencia superlineal
- $O(n^2)$ pasos	- debe conocerse un valor inicial bueno

5 Interpolación

En ingeniería y algunas ciencias es frecuente disponer de un cierto número de puntos obtenidos por muestreo o a partir de un experimento y pretender construir una función que los ajuste. Se denomina **interpolación** a la obtención de nuevos puntos partiendo del conocimiento de un conjunto discreto de puntos.

Polinomios: funciones de la forma $P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$. Dado que las derivadas e integrales de los polinomios son fáciles de calcular, éstos se usan para aproximar funciones continuas.

Para curvas que no pueden definirse mediante polinomios contamos con curvas paramétricas llamadas **curvas de Bezier**.

Supongamos que tenemos los puntos $y_0 = f(x_0)$, $y_1 = f(x_1)$, ..., $y_n = f(x_n)$ y queremos obtener un polinomio de grado n que pase lo más cerca posible de y_0, \dots, y_n . Lo que podemos hacer es plantear el sistema de ecuaciones:

$$\begin{cases} a_0 + a_1x_0 + \dots + a_nx_0^n = y_0 \\ \vdots \\ a_0 + a_1x_n + \dots + a_nx_n^n = y_n \end{cases} \Rightarrow \begin{pmatrix} 1 & x_0 & \dots & x_0^n \\ 1 & x_1 & \dots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_0 \\ \vdots \\ y_n \end{pmatrix}$$

La matriz de este sistema se llama **matriz de Vandermonde**. Esta matriz es mal condicionada. Por lo tanto, este generalmente no es un buen método.

Fenómeno de Runge: se da cuando aparecen oscilaciones no deseadas en los extremos del intervalo a interpolar.

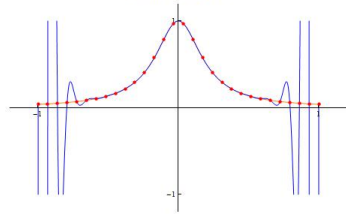


Figura 7: Fenómeno de Runge.

5.1 Interpolación y polinomio de Lagrange

Teorema: si x_0, \dots, x_n son $n + 1$ números distintos y si f es una función cuyos valores están dados en esos números, entonces existe un *único* polinomio $P(x)$ de grado a lo sumo n que satisface $f(x_k) = P(x_k)$ con $k = 0, 1, \dots, n$.

Este polinomio está dado por:

$$L_{n;i} = \prod_{j=0, j \neq i}^n \frac{(x - x_j)}{(x_i - x_j)}$$

$$P_n(x) = \sum_{i=0}^n [f(x_i) \cdot L_{n;i}]$$

Ventajas	Desventajas
- la distribución de puntos no incide	- $O(n^2)$ operaciones para evaluar $P_n(x)$
- es sencillo	- agregar (x_{n+1}, y_{n+1}) requiere recalcular todo
	- es inestable

5.2 Método de Newton

El **método de las diferencias divididas de Newton** consiste en definir

- $f(x_i) = y_i$
- $f(x_i, x_{i+1}) = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}$
- $f(x_i, x_{i+1}, x_{i+2}) = \frac{f(x_{i+1}, x_{i+2}) - f(x_i, x_{i+1})}{x_{i+2} - x_i}$
- \vdots
- $f(x_k, \dots, x_n) = \frac{f(x_{k+1}, \dots, x_n) - f(x_k, \dots, x_{n-1})}{x_n - x_k}$

El método de las diferencias divididas “**progresivas**” de Newton de grado 3 consiste en:

$$P(x) = f(x_0) + f(x_0, x_1)(x - x_0) + f(x_0, x_1, x_2)(x - x_0)(x - x_1) + f(x_0, x_1, x_2, x_3)(x - x_0)(x - x_1)(x - x_2)$$

donde $x_0 > x_1 > x_2 > x_3$.

El método de las diferencias divididas “**regresivas**” de Newton de grado 3 consiste en:

$$P(x) = f(x_3) + f(x_2, x_3)(x - x_3) + f(x_1, x_2, x_3)(x - x_3)(x - x_2) + f(x_0, x_1, x_2, x_3)(x - x_3)(x - x_2)(x - x_1)$$

donde $x_0 < x_1 < x_2 < x_3$.

Ventajas	Desventajas
- agregar un punto inicial o final es muy sencillo	- $O(n^2)$ operaciones para obtener $f(x_k, \dots, x_n)$
- $O(n)$ operaciones para evaluar $P_n(x)$	- agregar un punto intermedio requiere recalcular todo
	- se exigen datos x_i ordenados

5.3 Interpolación baricéntrica de Lagrange

Sea el polinomio genérico $L(x) = \prod_{i=0}^n (x - x_i)$

Sean los pesos baricéntricos

$$w_i = \frac{1}{\prod_{k=0, k \neq i}^n x_i - x_k} \text{ para todo } i = 0 \dots n$$

Entonces podemos definir al polinomio interpolante como:

$$P_n(x) = L(x) \cdot \sum_{i=0}^n \left[f(x_i) \frac{w_i}{x - x_i} \right]$$

Este es el llamado **método mejorado de Lagrange**.

Ventajas	Desventajas
- $O(n)$ operaciones para evaluar $P_n(x)$	
- $O(n)$ operaciones para agregar un par de datos	
- w_i no depende de $f(x_{n+1})$	
- no es necesario ordenar los datos	

El **método baricéntrico de Lagrange** consiste en:

$$P_n(x) = \frac{\sum_{i=0}^n f(x_i) \frac{w_i}{x - x_i}}{\sum_{i=0}^n \frac{w_i}{x - x_i}}$$

Ventajas	Desventajas
- $O(n)$ operaciones para agregar un par de datos	- es inestable
- es más estable que Lagrange original	

5.4 Interpolación de Hermite

La **interpolación de Hermite** se utiliza cuando conocemos el valor de la función en el punto y la derivada.

Teorema: sea $f \in C^1[a, b]$ y sean $x_0, \dots, x_n \in [a, b]$ distintos, el polinomio *único* que concuerda con $f(x_i)$ y $f'(x_i)$ para $i = 0 \dots n$ es el polinomio de Hermite de grado, a lo sumo, $2n + 1$, que está dado por

$$H_{2n+1}(x) = \sum_{i=0}^n f(x_i) H_{n,i}(x) + \sum_{i=0}^n f'(x_i) \hat{H}_{n,i}(x)$$

donde

- $H_{n,i}(x) = [1 - 2(x - x_i)L'_{n,i}(x_i)] [L_{n,i}(x)]^2$
- $\hat{H}_{n,i}(x) = (x - x_i) [L_{n,i}(x)]^2$

Este polinomio también se puede calcular mediante el método de diferencias divididas Newton. Definimos una sucesión $\{z_0, \dots, z_{2n+1}\}$ dada por $z_{2i} = z_{2i+1} = x_i$, con $i = 0 \dots n$. Entonces:

$$H_{2n+1}(x) = f(z_0) + \sum_{k=1}^{2n+1} \left[f(z_0, \dots, z_k) \cdot \prod_{j=0}^{k-1} (x - z_j) \right]$$

Y podemos formar la tabla correspondiente:

z_i	$f(z_i)$	$f(z_i, z_{i+1})$	$f(z_i, z_{i+1}, z_{i+2})$	$f(z_i, z_{i+1}, z_{i+2}, z_{i+3})$
$z_0 = x_0$	$f(z_0) = f(x_0)$	$f(z_0, z_1) = f'(x_0)$		
$z_1 = x_0$	$f(z_1) = f(x_0)$	$f(z_1, z_2)$	$f(z_0, z_1, z_2)$	$f(z_0, z_1, z_2, z_3)$
$z_2 = x_1$	$f(z_2) = f(x_1)$	$f(z_2, z_3) = f'(x_1)$	$f(z_1, z_2, z_3)$	
$z_3 = x_1$	$f(z_3) = f(x_1)$	$f(z_3, z_4)$		

Ventajas	Desventajas
	- procedimiento tedioso incluso para valores chicos de n

5.5 Interpolación por “splines”

La **interpolación por “splines”** consiste en interpolar datos mediante segmentos de curvas. En este caso los segmentos serán polinomios de tercer grado, denominados **trazadores cúbicos**. Sean los $n + 1$ puntos, definimos las n curvas interpolantes como:

$$S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3 \text{ con } i = 0..n - 1$$

Este polinomio debe cumplir las siguientes condiciones:

1. $S_i(x_i) = f(x_i)$ para $i = 0, \dots, n$ (la curva pasa por los puntos)
2. $S_{i+1}(x_{i+1}) = S_i(x_{i+1})$ para $i = 0, \dots, n - 2$ (S_i es continua)
3. $S'_{i+1}(x_{i+1}) = S'_i(x_{i+1})$ para $i = 0, \dots, n - 2$ (S'_i es continua)
4. $S''_{i+1}(x_{i+1}) = S''_i(x_{i+1})$ para $i = 0, \dots, n - 2$ (S''_i es continua)
5. Alguna de las siguientes condiciones de borde:

a) **Frontera libre / natural:** $S''_0(x_0) = S''_{n-1}(x_n) = S''_n(x_n) = 0$

b) **Frontera sujeta:** $\begin{cases} S'_0(x_0) = f'(x_0) = \alpha \\ S'_{n-1}(x_n) = S'_n(x_n) = f'(x_n) = \beta \end{cases}$

Para hallar los coeficientes a_i , b_i , c_i y d_i con spline de frontera libre debe plantearse:¹

$$\begin{aligned} & \blacksquare a_i = f(x_i) \\ & \blacksquare h_i = x_{i+1} - x_i \\ & \blacksquare \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ h_0 & 2(h_0 + h_1) & h_1 & h_2 & \vdots \\ 0 & h_1 & 2(h_1 + h_2) & h_2 & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & h_{n-2} & 2(h_{n-2} + h_{n-1}) & h_{n-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ 3[f_N(x_1, x_2) - f_N(x_0, x_1)] \\ \vdots \\ 3[f_N(x_{n-1}, x_n) - f_N(x_{n-2}, x_{n-1})] \\ 0 \end{bmatrix} \\ & \blacksquare b_i = f_N(x_i, x_{i+1}) - \frac{h_i}{3}(2c_i + c_{i+1}) \\ & \blacksquare d_i = \frac{c_{i+1} - c_i}{3h_i} \end{aligned}$$

Para hallar los coeficientes a_i , b_i , c_i y d_i con spline de frontera sujeta debe plantearse:

$$\begin{aligned} & \blacksquare a_i = f(x_i) \\ & \blacksquare h_i = x_{i+1} - x_i \\ & \blacksquare \begin{bmatrix} 2h_0 & h_0 & 0 & \dots & 0 \\ h_0 & 2(h_0 + h_1) & h_1 & h_2 & \vdots \\ 0 & h_1 & 2(h_1 + h_2) & h_2 & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & h_{n-2} & 2(h_{n-2} + h_{n-1}) & h_{n-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 3f_N(x_0, x_1) - 3f'(a) \\ 3f_N(x_1, x_2) - 3f_N(x_0, x_1) \\ \vdots \\ 3f_N(x_{n-1}, x_n) - 3f_N(x_{n-2}, x_{n-1}) \\ 3f'(b) - 3f_N(x_{n-1}, x_n) \end{bmatrix} \\ & \blacksquare b_i = f_N(x_i, x_{i+1}) - \frac{h_i}{3}(2c_i + c_{i+1}) \\ & \blacksquare d_i = \frac{c_{i+1} - c_i}{3h_i} \end{aligned}$$

La interpolación con spline de frontera sujeta es más precisa que la natural, pero requiere los valores de la derivada en los extremos, o al menos buenas aproximaciones.

Puede demostrarse que las matrices asociadas a los dos sistemas mencionados son estrictamente diagonal dominantes, definidas positivas, tribandas y simétricas, y que la solución de los dos es única.

¹ f_N denota la diferencia dividida de Newton

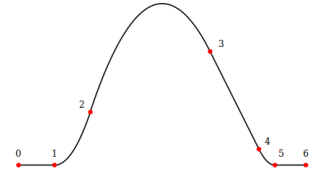


Figura 8: Spline.

6 Ajuste de funciones y mejor aproximación

6.1 Mejor aproximación

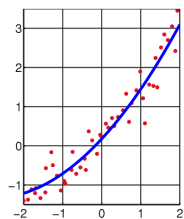


Figura 9: Ajuste.

Si la cantidad de puntos x_i dados es muy grande, interpolar mediante polinomios crea curvas de alto grado que se vuelven muy inestables en los extremos. Esto se puede resolver, en parte, mediante interpolaciones fragmentadas o segmentos de curvas (polinomios de Hermite o splines). Sin embargo, una de las condiciones fundamentales para estos métodos es que los puntos x_i sean distintos.

Cuando para un mismo x_i tenemos varios valores de $f(x_i)$ podemos construir una curva que ajuste lo mejor posible los datos disponibles, sin que la curva pase por los puntos dados.

La aproximación la hacemos con una función $g(x) = \sum_{i=0}^m c_i \phi_i(x)$, donde $m < n$.

6.2 Método de los cuadrados mínimos

El problema general de aproximar un conjunto de datos *distintos* $(x_1, y_1), \dots, (x_m, y_m)$ con un polinomio $P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ de grado $n < m - 1$ puede resolverse mediante el **método de los cuadrados mínimos**. Este método consiste en resolver las $n + 1$ ecuaciones normales para las $n + 1$ incógnitas a_j .

$$\sum_{k=0}^n a_k \sum_{i=1}^m x_i^{j+k} = \sum_{i=1}^m y_i x_i^j \text{ para cada } j = 0, \dots, n$$

En forma matricial esto es

$$\begin{bmatrix} \sum_{i=1}^m x_i^0 & \sum_{i=1}^m x_i^1 & \dots & \sum_{i=1}^m x_i^n \\ \sum_{i=1}^m x_i^1 & \sum_{i=1}^m x_i^2 & & \vdots \\ \vdots & & \ddots & \\ \sum_{i=1}^m x_i^n & \dots & & \sum_{i=1}^m x_i^{2n} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m y_i x_i^0 \\ \sum_{i=1}^m y_i x_i^1 \\ \vdots \\ \sum_{i=1}^m y_i x_i^n \end{bmatrix}$$

Nótese que el coeficiente A_{11} nos dice la cantidad de puntos que intervienen en el ajuste, y que la matriz es simétrica y definida positiva, y es similar a una matriz de Vandermonde. De ahí que cualquier ajuste de curvas hecho con polinomios es un problema mal condicionado, y por ello no se recomienda trabajar con polinomios de grado mayor a 4 o 5.

Puede demostrarse que las ecuaciones normales tienen una solución única, a condición de que las x_i sean distintas.

Si sospechamos que los datos (x_i, y_i) tienen una relación exponencial, podemos proponer la función de ajuste $y = ae^{bx}$.

7 Diferenciación numérica

Dado que no siempre las soluciones analíticas al problema de diferenciar son aplicables a los problemas, y no siempre existe tal solución numérica, existen métodos numéricos que aproximan las soluciones con distinto grado de exactitud.

Sin embargo, debe notarse que la diferenciación numérica es inestable y depende mucho de la precisión utilizada.

7.1 Aproximación por diferencias

Se pueden clasificar según 3 tipos:

1. **Diferencias progresivas:** $f'(x) \approx \frac{f(x+h)-f(x)}{h}$. Su orden de convergencia es $O(h)$.
2. **Diferencias regresivas:** $f'(x) \approx \frac{f(x)-f(x-h)}{h}$. Su orden de convergencia es $O(h)$.

3. **Diferencias centradas:** $f'(x) \approx \frac{f(x+h)-f(x-h)}{2h}$. Su orden de convergencia es $O(h^2)$. Es el que mejor aproxima de los tres.

Tener en cuenta que achicar el paso h en los tres algoritmos anteriores no siempre es una manera de obtener una aproximación mejor, ya que empieza a incidir el error de redondeo, dado que el numerador se vuelve cada vez más chico. Además, no es posible calcular un valor óptimo para h .

7.2 Método de los tres y cinco puntos

El algoritmo del **método de los tres puntos** es:

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

Su orden de convergencia es $O(h^2)$.

El algoritmo llamado **método de los cinco puntos** es:

$$f'(x) = \frac{f(x-2h) - 8f(x-h) + 8f(x+h) - f(x+2h)}{12h} + \frac{h^4}{30} f^{(5)}(\xi)$$

Su orden de convergencia es $O(h^4)$, es decir, proporcional a la derivada quinta, lo que lo hace muy preciso.

7.3 Extrapolación de Richardson

El **método de extrapolación de Richardson** es aplicable solo si $E(h) = k_1 h_1 + k_2 h_1^2 + k_3 h_1^3 + \dots$, y se puede generalizar de la siguiente manera:

$$N_j(h) = N_{j-1}\left(\frac{h}{q}\right) + \frac{N_{j-1}\left(\frac{h}{q}\right) - N_{j-1}(h)}{q^{j-1} - 1}$$

siendo N_1 el resultado de aproximar la derivada con alguno de los métodos mencionados anteriormente (por ejemplo, diferencias progresivas).

Ejemplo: aproximar $f'(1)$ siendo $f(x) = \sin\left(\frac{\pi}{3}x\right)$ con extrapolación de Richardson, con $q = 2$.

h_i	$N_1 = \frac{f(1+h)-f(1)}{h}$	$N_2 = N_1\left(\frac{h}{2}\right) + \frac{N_1\left(\frac{h}{2}\right) - N_1(h)}{2^1 - 1}$	$N_3 = N_2\left(\frac{h}{2}\right) - \frac{N_2\left(\frac{h}{2}\right) - N_2(h)}{2^2 - 1}$	N_4
0,2	$N_1(0,2) = 0,4252$			
0,1	$N_1(0,1) = 0,4752$	$N_2(0,2) = 0,5252$		
0,05	$N_1(0,05) = 0,4996$	$N_2(0,1) = 0,5240$	$N_3(0,2) = 0,5236$	
0,025	$N_1(0,025) = 0,5117$	$N_2(0,05) = 0,5238$	$N_3(0,1) = 0,5237$	$N_4(0,2) = 0,5237$

La extrapolación puede aplicarse si el error de truncamiento es $\sum_{j=1}^{n-1} k_j h_j^\alpha + O(h^\alpha m)$.

8 Integración numérica

Cuando necesitamos obtener el valor de una integral definida, pueden ocurrir dos cosas: que la función no tenga una antiderivada explícita, o que ésta no sea fácil de obtener. Entonces podemos obtener una aproximación utilizando una cuadratura.

Cuadratura numérica: $\int_a^b f(x) dx \approx \sum_{i=1}^n c_i f(x_i)$, donde c_i son los “coeficientes de peso” y los $f(x_i)$ son los “puntos de cuadratura”.

Los siguientes métodos aproximan la función $f(x)$ con un polinomio interpolante, cuyas antiderivadas son fáciles de calcular.

8.1 Fórmulas cerradas de Newton-Cotes

Fórmula cerrada de Newton-Cotes: la fórmula cerrada de $n + 1$ puntos de Newton-Cotes utiliza los nodos equidistantes $x_i = x_0 + ih$, para $i = 0..n$, donde $x_0 = a$, $x_n = b$ y $h = \frac{b-a}{n}$. Se llama “cerrada” porque se incluyen los extremos del intervalo como nodos.

- **Regla del rectángulo:**

$$\underbrace{\int_a^b f(x) dx \approx (b-a) \cdot f(a)}_{\text{por defecto}} \text{ o } \underbrace{\int_a^b f(x) dx \approx (b-a) \cdot f(b)}_{\text{por exceso}}$$

Su orden de convergencia es $O(h)$. Esta regla da un resultado exacto para funciones del tipo $f(x) = c$.

- **Regla del trapecio (n=1):** se obtiene al integrar en $[a, b]$ el primer polinomio de Lagrange. Se aproxima el área bajo una función f de valores positivos con un trapecio.

$$\int_{x_0}^{x_1} f(x) dx = \frac{h}{2} \cdot [f(x_0) + f(x_1)] - \frac{h^3}{12} f''(\xi)$$

siendo $h = b - a$. Su convergencia es $O(h^2)$. Esta regla da un resultado exacto para polinomios de grado 1 o menos.

- **Regla de Simpson (n=2):** se obtiene al integrar el segundo polinomio de Lagrange.

$$\int_{x_0}^{x_2} f(x) dx = \frac{h}{3} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right] - \frac{h^5}{90} f^{(4)}(\xi)$$

siendo $h = \frac{b-a}{2}$. Su convergencia es $O(h^4)$. Esta regla proporciona resultados exactos al aplicarla a polinomios de grado 3 o menos.

- **Regla de tres octavos de Simpson (n=3):**

$$\int_{x_0}^{x_3} f(x) dx = \frac{3h}{8} [f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)] - \frac{3h^5}{80} f^{(4)}(\xi)$$

siendo $h = \frac{b-a}{3}$.

8.2 Fórmulas abiertas de Newton-Cotes

Fórmula abierta de Newton-Cotes: la fórmula abierta de $n + 1$ puntos de Newton-Cotes utiliza los nodos equidistantes $x_i = x_0 + ih$, para $i = 0..n$, donde $x_0 = a + h$, $x_n = b - h$ y $h = \frac{b-a}{n+2}$. Se llama “abierta” porque no se incluyen los extremos del intervalo como nodos.

- **Regla del punto medio (n=0):**

$$\begin{aligned} \int_{x_0}^{x_1} f(x) dx &= 2hf(x_0) + \frac{h^3}{3} f''(\xi) \\ &= (x_1 - x_0) \cdot f\left(\frac{x_0 + x_1}{2}\right) \end{aligned}$$

8.3 Métodos compuestos

En general, las fórmulas de Newton-Cotes no son adecuadas para utilizarse en intervalos de integración grandes. Para estos casos se utilizan los siguientes métodos, que tienen la ventaja de que son estables respecto del error de redondeo. Todos estos algoritmos son $O(h^4)$, pero tienen una desventaja: el tiempo de procesamiento. Esto quiere decir que reducir el paso h mejora la precisión, pero llegará un punto en que la representación numérica nos impedirá afinar el paso. Además, cada vez que querremos afinar el paso, deberemos calcular todo otra vez.

- **Regla compuesta del rectángulo:**

$$\int_a^b f(x) dx \approx h \left[\sum_{i=1}^{n-1} f(x_i) + f(b) \right]$$

siendo $h = \frac{b-a}{n}$ y $x_j = a + jh$.

- **Regla compuesta del trapecio:** sea $f \in C^2[a, b]$. Entonces la fórmula compuesta del trapecio para n subintervalos puede escribirse como

$$\int_a^b f(x) dx = \frac{h}{2} \left[f(a) + 2 \sum_{j=1}^{n-1} f(x_j) + f(b) \right] - \frac{b-a}{12} h^2 f''(\mu)$$

siendo $h = \frac{b-a}{n}$ y $x_j = a + jh$ para cada $j = 0..n$.

- **Regla compuesta de Simpson:** sea $f \in C^4[a, b]$ y n par. Entonces la fórmula compuesta de Simpson para n subintervalos puede escribirse como

$$\int_a^b f(x) dx = \frac{h}{3} \left[f(a) + 2 \sum_{j=1}^{\frac{n}{2}-1} f(x_{2j}) + 4 \sum_{j=1}^{\frac{n}{2}} f(x_{2j-1}) + f(b) \right] - \frac{b-a}{180} h^4 f^{(4)}(\mu)$$

siendo $h = \frac{b-a}{n}$ y $x_j = a + jh$ para cada $j = 0..n$.

- **Regla compuesta del punto medio:** sea $f \in C^2[a, b]$ y n par. Entonces la fórmula compuesta del punto medio para $n + 2$ subintervalos puede escribirse como

$$\int_a^b f(x) dx = 2h \sum_{j=0}^{\frac{n}{2}} f(x_{2j}) + \frac{b-a}{6} h^2 f''(\mu)$$

siendo $h = \frac{b-a}{n+2}$ y $x_j = a + (j+1)h$ para cada $j = -1..(n+1)$.

8.4 Método de Romberg

El **método de integración de Romberg** utiliza la regla compuesta del trapecio para obtener aproximaciones preliminares y luego se aplica el proceso de extrapolación de Richardson para mejorar las aproximaciones. Si se verifica que $f \in C^{2(k+1)}[a, b]$, la fórmula está dada por

$$R_{k,j} = R_{k,j-1} + \frac{R_{k,j-1} - R_{k-1,j-1}}{4^{j-1} - 1}$$

donde $k = 2, 3, \dots, n$ y $j = 2, \dots, k$. Su convergencia es $O(h_k^{2j})$. Los resultados generados con éstas fórmulas se incluyen en la siguiente tabla.

$R_{1,1}$			
$R_{2,1}$	$R_{2,2}$		
\vdots	\vdots	\ddots	
$R_{n,1}$	$R_{n,2}$	\cdots	$R_{n,n}$

El “mejor” valor es $R_{n,n}$. Otra ventaja de este método es que permite agregar una nueva fila a la tabla con sólo hacer una aplicación más de la regla compuesta del trapecio (y luego promediar los demás valores).

8.5 Cuadratura de Gauss

En todas las fórmulas de Newton-Cotes se emplean valores de la función en puntos equidistantes. Esta práctica es útil cuando utilizamos dichos métodos en reglas compuestas, pero esta restricción puede afectar la exactitud de la aproximación.

El **método de cuadratura de Gauss** selecciona los puntos de la evaluación de manera óptima y *no* en una forma igualmente espaciada. Recordando la fórmula de una cuadratura $\sum_{i=1}^n c_i f(x_i)$, los valores x_i son las raíces del polinomio de Legendre, y las constantes c_i se obtienen operando con los x_i . Estos valores están tabulados:

n	x_i	c_i
1	$x_0 = 0$	$c_0 = 2$
2	$x_0 = 0,57735$	$c_0 = 1$
	$x_1 = -0,57735$	$c_1 = 1$
3	$x_0 = 0,77459$	$c_0 = 0,55555$
	$x_1 = 0$	$c_1 = 0,88888$
	$x_2 = -0,77459$	$c_2 = 0,55555$
\vdots	\vdots	\vdots

La fórmula para este método es

$$\int_a^b f(x) dx = \int_{-1}^1 f\left(\frac{(b-a)t + (b+a)}{2}\right) \frac{b-a}{2} dt$$

Este método arroja resultados exactos cuando $g \leq 2n - 1$, donde g es el grado del polinomio a integrar, y n es la cantidad de puntos de Gauss. Además, el error cometido es proporcional a la derivada de orden $2n$.

Ventajas	Desventajas
- pocas evaluaciones del polinomio	
- fácil de programar	
- sólo desconocemos el error de redondeo	

8.6 Integrales múltiples

Dada la integral $\iint_A f(x, y) dx dy$, donde A es una región rectangular del plano tal que $A = \{(x, y) : a \leq x \leq b \wedge c \leq y \leq d\}$, podemos obtener una aproximación utilizando una adaptación de los métodos anteriormente estudiados.

■ Método del trapecio

$$\int_c^d \int_a^b f(x, y) dx dy \approx \frac{(b-a)(d-c)}{4} [f(a, c) + f(a, d) + f(b, c) + f(b, d)]$$

■ Método de Simpson

$$\begin{aligned} \int_c^d \int_a^b f(x, y) dx dy \approx & \frac{(b-a)(d-c)}{36} \{f(x_0, y_0) + f(x_0, y_2) + f(x_2, y_0) + f(x_2, y_2) \\ & + 4[f(x_0, y_1) + f(x_1, y_0) + f(x_1, y_2) + f(x_2, y_1) + 4f(x_1, y_1)]\} \end{aligned}$$

donde $x_0 = a$, $x_1 = \frac{a+b}{2}$, $x_2 = b$, $y_0 = c$, $y_1 = \frac{c+d}{2}$, $y_2 = d$.

■ Método de cuadratura de Gauss

$$\int_c^d \int_a^b f(x, y) dx dy \approx \frac{(b-a)(d-c)}{4} \sum_{i=1}^n \sum_{j=1}^m c_i c_j f(x_i, y_j)$$

con $x_i = \frac{b-a}{2} t_i + \frac{b+a}{2}$, $y_j = \frac{d-c}{2} t_j + \frac{d+c}{2}$, siendo t_i , t_j las raíces de los polinomios de Legendre y c_i , c_j los coeficientes de peso.

9 Ecuaciones diferenciales ordinarias

Estudiaremos cómo aproximar la solución $y(t)$ a un problema de la forma

$$\begin{cases} \frac{dy}{dt} = f(t, y) & \text{para } a \leq t \leq b \\ y(a) = \alpha \end{cases}$$

Condición de Lipschitz: una función $f(t, y)$ satisface la condición de Lipschitz en la variable y en un conjunto $D \subset \mathbb{R}^2$ si existe una constante $L > 0$ tal que

$$\frac{|f(t, y_1) - f(t, y_2)|}{|y_1 - y_2|} \leq L \text{ para todo } (t, y_1), (t, y_2) \in D$$

Teorema: si $f(t, y)$ está definida en un conjunto convexo $D \subset \mathbb{R}^2$ y existe una constante $L > 0$ tal que $\left| \frac{\partial f(t, y)}{\partial y} \right| \leq L$ para toda $(t, y) \in D$, entonces f satisface la condición de Lipschitz.

Teorema (de suficiencia): supongamos que $D = \{(t, y) | a \leq t \leq b; -\infty < y < \infty\}$ y que $f(t, y)$ es continua en D . Si f satisface la condición de Lipschitz, entonces el problema de valor inicial $y'(t) = f(t, y)$, $a \leq t \leq b$, $y(a) = \alpha$ tiene solución única.

Problema bien planteado: el problema de valor inicial $\frac{dy}{dt} = f(t, y)$, $a \leq t \leq b$, $y(a) = \alpha$ es bien planteado si:

1. El problema tiene solución única $y(t)$,
2. Si perturbamos el problema, la solución también se perturbará, pero poco.

Los siguientes métodos no nos proporcionan la función $y(t)$, sino que nos da aproximaciones a $y(t_i)$ para puntos $a \leq t_i \leq b$ igualmente espaciados. Si queremos valores intermedios deberemos interpolarlos, generalmente utilizando el método de Hermite.

9.1 Métodos de un paso

En los **métodos de un paso** la aproximación del punto de red t_{i+1} contiene información proveniente de uno solo de los puntos anteriores de red t_i .

9.1.1 Método de Euler

El método de Euler es un caso particular del método de Taylor para $n = 1$. Este método no es lo suficientemente exacto para utilizarlo en la vida real, pero resulta simple para analizarlo. Sin embargo, son métodos estables.

- **Método de Euler explícito:** siendo el paso $h = \frac{b-a}{N}$, donde N es la cantidad de intervalos, la expresión de este método es:

$$\begin{aligned} w_i &= y(t_i) \\ w_{i+1} &= w_i + h \cdot f(t_i, w_i) \text{ para } i = 0, 1, \dots, N-1 \end{aligned}$$

El error local es $O(h)$. Converge siempre.

- **Método de Euler implícito:** siendo el paso $h = \frac{b-a}{N}$, donde N es la cantidad de intervalos, la expresión de este método es:

$$\begin{aligned} w_i &= y(t_i) \\ w_{i+1} &= w_i + h \cdot f(t_{i+1}, w_{i+1}) \text{ para } i = 0, 1, \dots, N-1 \end{aligned}$$

El error local es $O(h)$. El método de Euler implícito suele dar mejores resultados que el explícito (aunque esto no siempre es así). La desventaja del método implícito es que no siempre es posible implementarlo.

- **Método predictor-corrector de Euler:** este método consiste en obtener una primera aproximación con el método explícito, y con ésta obtener una nueva aproximación con el método implícito, que “corrige” la anterior.

$$\begin{aligned} w_{i+1}^* &= w_i + h \cdot f(t_i, w_i) \\ w_{i+1} &= w_i + h \cdot f(t_i, w_{i+1}^*) \end{aligned}$$

El error local de este método es $O(h)$.

9.1.2 Métodos de Taylor de orden superior

El **método de Taylor de orden n** tiene la siguiente expresión:

$$\begin{aligned} w_0 &= \alpha \\ w_i &= y(t_i) \\ w_{i+1} &= w_i + h \cdot f(t_i, w_i) + \frac{h^2}{2} f'(t_i, w_i) + \dots + \frac{h^n}{n!} f^{(n-1)}(t_i, w_i) \text{ para } i = 0, \dots, n-1 \end{aligned}$$

Puede demostrarse que si $f \in C^{n+1}[a, b]$, entonces el error de truncamiento es $O(h^n)$. Este método rara vez se emplea en la práctica, pues tiene la desventaja de requerir el cálculo y evaluación de las derivadas de $f(t, y)$.

9.1.3 Métodos de Runge-Kutta

Teorema: sea $f(t, y) \in C^{n+1}$ en $D = \{(t, y) | a \leq t \leq b, c \leq y \leq d\}$ y sea $(t_0, y_0) \in D$. Entonces, para toda $(t, y) \in D$, existe $\xi \in [t_0, t]$ y $\mu \in [y_0, y]$ con $f(t, y) = P_n(t, y) + R_n(t, y)$ tal que $P_n(t, y)$ es el n -ésimo polinomio de Taylor en dos variables para la función f alrededor de (t_0, y_0) , y $R_n(t, y)$ es el término residual asociado a $P_n(t, y)$.

Los **métodos de Runge-Kutta** se basan en aproximar el polinomio de Taylor para una variable mediante polinomios de Taylor de dos variables. Existen varios métodos de Runge Kutta, que se clasifican según su orden de convergencia. Los más sencillos son los de orden 2:

- **Método del punto medio:**

$$\begin{aligned}w_0 &= \alpha \\w_{i+1} &= w_i + h \cdot f\left(t_i + \frac{h}{2}, w_i + \frac{h}{2} \cdot f(t_i, w_i)\right) \text{ para cada } i = 0, 1, \dots, N-1\end{aligned}$$

- **Método modificado de Euler:**

$$\begin{aligned}w_0 &= \alpha \\w_{i+1} &= w_i + \frac{h}{2} [f(t_i, w_i) + f(t_{i+1}, w_i + h \cdot f(t_i, w_i))] \text{ para cada } i = 0, 1, \dots, N-1\end{aligned}$$

- **Método de Crank-Nicolson (implícito):**

$$\begin{aligned}w_0 &= \alpha \\w_{i+1} &= w_i + \frac{h}{2} [f(t_i, w_i) + f(t_{i+1}, w_{i+1})] \text{ para } i = 0, 1, \dots, N-1\end{aligned}$$

- **Método de Heun:**

$$\begin{aligned}w_0 &= \alpha \\w_{i+1} &= w_i + \frac{h}{4} \left[f(t_i, w_i) + 3f\left(t_i + \frac{2}{3}h, w_i + \frac{2}{3}h \cdot f(t_i, w_i)\right) \right] \text{ para } i = 0, 1, \dots, N-1\end{aligned}$$

Para obtener métodos con mayor orden de convergencia se aplica el teorema antes mencionado, y con él se obtiene el **método de Runge-Kutta de orden 4**, que es muy preciso siempre y cuando la función $y(t)$ tenga al menos cinco derivadas continuas:

$$\begin{aligned}w_0 &= \alpha \\k_1 &= h \cdot f(t_i, w_i) \\k_2 &= h \cdot f\left(t_i + \frac{h}{2}, w_i + \frac{1}{2} \cdot k_1\right) \\k_3 &= h \cdot f\left(t_i + \frac{h}{2}, w_i + \frac{1}{2} \cdot k_2\right) \\k_4 &= h \cdot f(t_{i+1}, w_i + k_3) \\w_{i+1} &= w_i + \frac{1}{6} (k_1 + 2k_2 + 2k_3 + k_4) \text{ para cada } i = 0, 1, \dots, N-1\end{aligned}$$

9.2 Métodos multipasos

Los **métodos multipasos** son aquellos que emplean la aproximación en más de uno de los puntos de red precedentes para determinar la aproximación en el siguiente punto. Se pueden deducir de las series de Taylor o del polinomio interpolante de Lagrange.

Para obtener $w_i, w_{i-1}, w_{i-2}, \dots$ debemos utilizar un método de paso simple del mismo orden que el método de paso múltiple.

9.2.1 Métodos explícitos de Adams-Bashforth

Dado un orden de convergencia p , los **métodos explícitos de Adams-Bashforth** usan los puntos $w_i, w_{i-1}, \dots, w_{i+1-p}$ para obtener el w_{i+1} . El paso no puede ser arbitrario.

- **Método de Adams-Bashforth de orden 2** (2 pasos):

$$\begin{aligned}w_0 &= \alpha \\w_1 &= \alpha_1 \\w_{i+1} &= w_i + \frac{h}{2} [3f(t_i, w_i) - f(t_{i-1}, w_{i-1})] \text{ para } i = 1, 2, \dots, N-1\end{aligned}$$

- **Método de Adams-Bashforth de orden 4** (4 pasos):

$$\begin{aligned}w_0 &= \alpha \\w_1 &= \alpha_1 \\w_2 &= \alpha_2 \\w_3 &= \alpha_3 \\w_{i+1} &= w_i + \frac{h}{24} [55f(t_i, w_i) - 59f(t_{i-1}, w_{i-1}) + 37f(t_{i-2}, w_{i-2}) - 9f(t_{i-3}, w_{i-3})] \text{ para } i = 3, 4, \dots, N-1\end{aligned}$$

9.2.2 Métodos implícitos de Adams-Moulton

Dado un orden de convergencia p , los **métodos implícitos de Adams-Moulton** usan los puntos $w_{i+1}, w_i, w_{i-1}, \dots, w_{i+2-p}$ para obtener el w_{i+1} . El tamaño del paso puede ser arbitrario.

- **Método de Adams-Moulton de orden 2** (1 paso):

$$\begin{aligned}w_0 &= \alpha \\w_1 &= \alpha_1 \\w_{i+1} &= w_i + \frac{h}{2} [f(t_{i+1}, w_{i+1}) - f(t_i, w_i)] \text{ para } i = 1, 2, \dots, N-1\end{aligned}$$

Este método es equivalente al método de Crank-Nicolson.

- **Método de Adams-Moulton de orden 4** (3 pasos):

$$\begin{aligned}w_0 &= \alpha \\w_1 &= \alpha_1 \\w_2 &= \alpha_2 \\w_{i+1} &= w_i + \frac{h}{24} [9f(t_{i+1}, w_{i+1}) + 19f(t_i, w_i) - 5f(t_{i-1}, w_{i-1}) + f(t_{i-2}, w_{i-2})] \text{ para } i = 2, 3, \dots, N-1\end{aligned}$$

Como no siempre es posible transformar la expresión de Adams-Moulton a una de forma explícita, existen los **métodos predictores-correctores de Adams**, que combinan un método de Adams-Bashforth y un método de Adams-Moulton, ambos del mismo orden de convergencia.

<i>Adams-Bashforth</i>	<i>Adams-Moulton</i>
- más sencillos de programar	- más estables
	- menos errores de redondeo
	- más precisos

9.3 EDOs con condiciones de contorno

Trataremos de resolver ecuaciones del tipo $\frac{d^n y}{dt^n} = f(t, y, y', \dots, y^{(n-1)})$ en el intervalo $[a, b]$ y con las condiciones iniciales $y(a) = \alpha, y'(a) = \beta, \dots, y^{(n-1)}(a) = \gamma$.

9.3.1 Método del disparo lineal

Teorema: sea el problema $y'' = f(x, y, y')$ con $a \leq x \leq b$ y las condiciones $y(a) = \alpha$, $y(b) = \beta$. Si f , f'_y , $f'_{y'}$ son continuas en $D = \{(x, y, y') : a \leq x \leq b, -\infty < y < \infty, -\infty < y' < \infty\}$ y si $f'_{y'}(x, y, y') > 0$ para toda $(x, y, y') \in D$ y si existe una constante M tal que $|f'_{y'}(x, y, y')| \leq M$ para toda $(x, y, y') \in D$, entonces el problema con valor en la frontera tiene una solución única.

Corolario: sea el problema $y'' = p(x)y' + q(x)y + r(x)$ con $a \leq x \leq b$ y las condiciones iniciales $y(a) = \alpha$, $y(b) = \beta$. Si se satisface que

- $p(x), q(x), r(x)$ son continuas en $[a, b]$
- $q(x) > 0$ en $[a, b]$

entonces el problema tiene una solución única.

El método del disparo lineal tiene problemas de estabilidad. Se basa en la sustitución del problema de valor en la frontera por 2 problemas de valor inicial:

- (I) $y'' = p(x) \cdot y' + q(x) \cdot y + r(x)$ con $y(a) = \alpha$, $y'(a) = 0$
- (II) $y'' = p(x) \cdot y' + q(x) \cdot y$ con $y(a) = 0$, $y'(a) = 1$

Si llamamos $y_1(x)$ a la solución de (I) e $y_2(x)$ a la solución de (II), la solución final es:

$$y(x) = y_1(x) + \frac{\beta - y_1(b)}{y_2(b)} \cdot y_2(x)$$

9.3.2 Métodos de diferencias finitas para problemas lineales

El **método de diferencias finitas** es un método para resolver resoluciones de ecuaciones diferenciales de orden mayor o igual a dos. Consiste en reemplazar las derivadas por diferencias finitas mediante un cociente de diferencias. La aplicación de diferencias finitas generan un sistema de ecuaciones lineales del tipo $Ax = B$. Primero seleccionamos un entero $N > 0$ y dividimos el intervalo $[a, b]$ en $N + 1$ subintervalos iguales cuyos extremos son los puntos de malla $x_i = a + ih$ para $i = 0, 1, \dots, N + 1$, donde $h = \frac{b-a}{N+1}$. Al escoger h de este modo (no muy chico), se facilita la aplicación de un algoritmo matricial.

La **fórmula de las diferencias centradas** para $y''(x_i)$ es

$$y''(x_i) = \frac{1}{h^2} [y(x_{i+1}) - 2y(x_i) + y(x_{i-1}))] - \frac{h^2}{12} y^{(4)}(\xi_i)$$

El método de diferencias finitas con error de truncamiento de orden $O(h^2)$ consiste en resolver el siguiente sistema de ecuaciones:

$$\begin{bmatrix} 2 + h^2 q(x_1) & -1 + \frac{h}{2} p(x_1) & 0 & \cdots & 0 \\ -1 - \frac{h}{2} p(x_2) & 2 + h^2 q(x_2) & & & \vdots \\ 0 & & \ddots & & 0 \\ \vdots & & & \ddots & \\ 0 & \cdots & 0 & -1 - \frac{h}{2} p(x_N) & 2 + h^2 q(x_N) \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} -h^2 r(x_1) + (1 + \frac{h}{2} p(x_1)) w_0 \\ -h^2 r(x_2) \\ \vdots \\ -h^2 r(x_{N-1}) \\ -h^2 r(x_N) + (1 + \frac{h}{2} p(x_N)) w_{N+1} \end{bmatrix}$$

9.3.3 Método de los elementos finitos

El **método de los elementos finitos** es un método para aproximar soluciones de ecuaciones diferenciales muy utilizado en análisis estructural.