

Instituto Tecnológico de Costa Rica
Área académica de Ingeniería en Computadores
Curso: CE-2201 Análisis Numérico para Ingeniería



Catálogo

Profesor: Dr. Pablo Alvarado Moya
Estudiante: Juan Pablo Brenes Coto

10 de agosto de 2018

1. Lección 1: Análisis Numérico

1.1. Modelos matemáticos

Describen y simplifican el comportamiento de un fenómeno físico. Los modelos matemáticos comúnmente poseen la siguiente estructura:

$$\underline{s} = f(\underline{e}; \underline{p}; \underline{r}) \quad (1.1)$$

- \underline{s} : Vector de salidas
- \underline{e} : Vector de entradas (tiempo, espacio)
- \underline{p} : Vector de parámetros
- \underline{r} : Factores externos, ruido

1.2. Tipos de análisis

Solución de métodos numéricos por:

- Métodos analíticos
- Métodos numéricos: Reformulan el problema matemático para resolverlo mediante operaciones aritméticas.

2. Lección 2: Aproximaciones y Errores

2.1. Aproximaciones

Aproximación:

- Modelo matemático aproxima comportamiento real
- Método numérico aproxima solución analítica
- Aproximar error si se desconoce objeto

Cifras significativas:

- Indican confiabilidad de un valor numérico
- Igual a número de dígitos obtenidos con certeza, más uno estimado
- Métodos numéricos aproximan resultados: debe especificarse cuántas cifras significativas son validas

Exactitud y precisión:

- **Exactitud (Sesgo o bias)**: qué tan cercano está el valor medido o calculado al valor verdadero
- **Precisión (incertidumbre)**: qué tanto se dispersan las mediciones alrededor del valor medido

2.2. Errores

Error verdadero: Calculable solo si se cuenta con el valor verdadero, de otro modo debe aproximarse (error aproximado). Ignora orden de magnitud de valor estimado.

$$E_t = \text{valor verdadero} - \text{valor aproximado} \quad (2.1)$$

Error relativo verdadero: Error relativo fraccional verdadero considera orden del magnitud de valor estimado. Error relativo porcentual verdadero: $e_t = E_{rel} * 100 \%$

$$E_{rel} = \frac{E_t}{\text{valor verdadero}} = 1 - \frac{\text{valor aproximado}}{\text{valor verdadero}} \quad (2.2)$$

Error porcentual aproximado: Si no se cuenta con el valor verdadero, el error se normaliza con respecto al mismo valor aproximado.

$$e_a = \frac{\text{error aproximado}}{\text{valor aproximado}} * 100 \% \quad (2.3)$$

Si no se cuenta con el valor verdadero, para método iterativos se utiliza:

$$e_a = \frac{\text{aproximación actual} - \text{aproximación anterior}}{\text{aproximación actual}} * 100 \% \quad (2.4)$$

Se itera hasta $|e_a| > e_s$, donde e_s es el Umbral de Scarborough.

Umbral de Scarborough: Garantiza que el resultado será correcto en al menos n cifras significativas.

$$e_s = (0,5 \times 10^{2-n}) \% \quad (2.5)$$

2.3. Representaciones numéricas

2.3.1. Números codificados con coma fija

Las representaciones con coma fija son posicionales, donde el peso de cada bit en la representación es constante. $x = \sum_{n=0}^{N-1} b_n 2^n$

Coma fija sin signo: $x = \frac{1}{M} \sum_{n=0}^{N-1} b_n 2^n$. M es una constante de normalización, usualmente es elegida como 2^m

Complemento a dos: Para números enteros $x = -b_{N-1} 2^{N-1} + \sum_{n=0}^{N-2} b_n 2^n$ el ultimo bit b_{N-1} codifica el signo. Para números con signo $x = \frac{1}{M} (-b_{N-1} 2^{N-2} + \sum_{n=0}^{N-1} b_n 2^n)$

2.3.2. Números codificados con coma flotante

Según el estandar IEEE 754 (1985-2008): $x = (-1)^s x(1, m) x 2^{e-bias}$, donde **s** es el bit de signo, **e** es el exponente con E bits y **m** es la mantisa normalizada de M bits. **bias** = $2^{E-1} - 1$

La mantisa se completa con un 1 bit oculto (no se indica explícitamente), esta representa solo la parte fraccionaria.

Error de redondeo: Se produce al utilizar representaciones numéricas incapaces de representar todas las cifras significativas del número a representar.

Epsilon de formato: Sea Δx el intervalo entre presentaciones válidas alrededor de un valor x , si se utiliza corte, epsilon ε se define como $\varepsilon \geq \frac{|\Delta x|}{|x|}$. Si se utiliza redondeo: $\frac{\varepsilon}{2} \geq \frac{|\Delta x|}{|x|}$

3. Lección 3: Errores de truncamiento

3.1. Manipulaciones aritméticas

Suma y resta: Se toma el número con menor exponente, se alinea la coma decimal modificando la mantisa (igualar exponentes), se realiza la suma o resta. En caso necesario se renormaliza el número.

Multiplicación y división: Se suman (o restan) exponentes, se multiplican (o dividen) las mantisas de n dígitos, se normaliza el resultado y se redondea si es necesario.

Suma de números grandes y pequeños: La suma de un número grande y uno pequeño, con una diferencia en ordenes de magnitud mayor al número de cifras significativas, no produce ningún efecto (resultado es igual al número más grande).

En sumas de gran número de terminos, se debe sumar primero los terminos pequeños y por ultimo los grandes.

3.2. Series de Taylor

3.2.1. Teorema de Taylor

Si se tiene una serie infinita:

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + R_n \quad (3.1)$$

Donde el residuo R_n (en su forma integral) se calcula como:

$$R_n = \sum_{k=n+1}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = \int_{x_0}^x \frac{(x - \tau)^n}{n!} f^{(n+1)}(\tau) d\tau \quad (3.2)$$

Si omite a R_n , este corresponde al error de truncamiento.

3.2.2. Teorema del valor medio

Para integrales de funciones continuas, si $h(t)$ no cambia de signo en $[x_0, x]$

$$\int_{x_0}^x g(t)h(t)dt = g(\varepsilon) \int_{x_0}^x h(t)dt \quad (3.3)$$

Donde $\varepsilon \in [x_0, x]$

Aplicando el teorema del valor medio a la forma integral del residuo, se obtiene la forma de Lagrange del residuo:

$$R_n = \frac{f^{n+1}(\varepsilon)}{(n+1)!} (x - x_0)^{n+1} \quad (3.4)$$

4. Lección 4: Errores de truncamiento

4.1. Diferenciación numérica

Diferenciación numérica hacia adelante: $f'(x_i) = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} + O(x_{i+1} - x_i) = \frac{\Delta f_i}{h} + O(h)$. Donde Δf_i es la primera diferencia hacia adelante y h es el tamaño de paso o incremento.

Diferenciación numérica hacia atrás: $f'(x_i) = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} + O(x_i - x_{i-1})$. Donde $f(x_i) - f(x_{i-1})$ es la primera diferencia hacia atrás y $x_i - x_{i-1}$ es el tamaño de paso o incremento.

4.2. Propagación de errores

Si se desea estimar:

$$\Delta f(\tilde{x}) = |f(x) - f(\tilde{x})| \quad (4.1)$$

en función de la diferencia $\Delta \tilde{x} = x - \tilde{x}$. Si $x \approx \tilde{x}$ entonces con una aproximación de primer orden se obtiene:

$$f(x) - f(\tilde{x}) = \Delta f(\tilde{x}) \approx f'(\tilde{x})(x - \tilde{x}) = |f'(\tilde{x})| \Delta \tilde{x} \quad (4.2)$$

4.3. Estabilidad y condición

Error relativo de una función: $E_f = \frac{f(x) - f(\tilde{x})}{f(x)} \approx \frac{f'(\tilde{x})(x - \tilde{x})}{f(\tilde{x})}$. El error relativo de x es $E_x = \frac{x - \tilde{x}}{\tilde{x}}$

Número de condición: Es la relación entre los dos errores relativos e indica qué tanto una inexactitud de x se aumenta por $f(x)$ $N = \frac{E_f}{E_x} = \frac{\tilde{x}f'(\tilde{x})}{f(\tilde{x})}$

- $N=1$: Error relativo de la función idéntico al de x
- $N > 1$: Error relativo se amplifica. Si $N \gg 1$: función mal condicionada
- $N \leq 1$: Error relativo se atenúa

Un cálculo es **numéricamente inestable** si la inexactitud de los valores de entrada aumenta por el método numérico (la función es mal condicionada).

Error numérico total: Es la suma de los errores de redondeo y truncamiento.

Control de errores numéricos

- Evitar restas de números similares, pues se pierden cifras significativas
- Reordenar operaciones aritméticas e iniciar con números más pequeños
- Utilizar precisión extendida (precisión doble)

5. Lección 5: Raíces de ecuaciones

5.1. Métodos cerrados

- **Punto de partida:** Una función cambia de signo en la vecindad de una raíz
 - Solo se cumple para raíces de multiplicidad impar
- Buscan **una** raíz.
- Parten de un intervalo que encierra la raíz.
- Intervalo se reduce iterativamente para acorralar a la raíz (**Convergen**).
- Método de Bisección e interpolación lineal.

5.2. Método de Bisección

Se busca una raíz x_r , suponiendo que $x_r \in [x_l, x_u]$. Si el intervalo es suficientemente pequeño entonces los signos de $f(x_l)$ y $f(x_u)$ difieren: $f(x_l) * f(x_u) < 0$.

El algoritmo consiste en partir en cada iteración el intervalo en dos. La condición de parada se realiza cuando el error aproximado $e_a = \left| \frac{x_r^i - x_r^{i-1}}{x_r^i} \right| * 100\%$ es menor a un umbral.

Puesto que $x_r^i - x_r^{i-1} = \frac{x_u - x_l}{2}$ y $x_r^i = \frac{x_u + x_l}{2}$ entonces el error aproximado se puede expresar como $e_a = \left| \frac{x_u - x_l}{x_u + x_l} \right| * 100\%$. En cada iteración el error aproximado se reduce a la mitad y

$$E_a^n = \frac{\Delta x^{(0)}}{2^n} = \frac{x_u^{(0)} - x_l^{(0)}}{2^n}.$$

Para un error deseado E_d se despeja el número de iteraciones $n = \log_2\left(\frac{\Delta x^{(0)}}{E_d}\right)$

Otros conceptos de error: La definición de error e_a dada anteriormente es altamente riesgosa si la raíz es cero o cerca de cero, por lo que se sugiere utilizar como error aproximado $e_a = |x_r^{(i)} - x_r^{(i-1)}|$

5.3. Método de interpolación lineal

También denominado método de **de la falsa posición**. El método de interpolación lineal, en vez de dividir el intervalo en dos, asume una aproximación lineal de la función para encontrar la raíz.

Iteración:
$$x_i = x_u - \frac{f(x_u)(x_l - x_u)}{f(x_l) - f(x_u)}$$

Desventajas: El método parte de la suposición de que la raíz se encuentra siempre más cercana al extremo menor en magnitud, lo cual no es siempre cierto.

El problema anterior se puede solucionar simulando un valor menor de la función en un extremo (multiplicándolo por un factor 1/2), si este se cambia 2 veces de manera consecutiva.

6. Lección 6: Raíces de ecuaciones

6.1. Método abiertos

- Requieren un único valor inicial o un par de valores pero que no necesitan encerrar la raíz buscada.
- A veces son **divergentes**.
- Si convergen, lo hacen más rápido que los métodos cerrados.

6.2. Método de iteración de punto fijo

La iteración de punto fijo parte de reformular la ecuación $f(x) = 0$ en $x = g(x)$ y permitir que un proceso iterativo $x_{i+1} = g(x_i)$ converja a la raíz, con un error aproximado calculado como $e_a = \left| \frac{x_{i+1} - x_i}{x_{i+1}} \right| * 100\%$.

6.2.1. Teorema del valor medio de la derivada

Sean $g(x)$ y su derivada $g'(x)$ continuas en un intervalo $[a, b]$. El teorema establece que existe algún punto ξ sobre el que $g'(\xi)$ iguala a la pendiente de la recta trazada entre $g(a)$ y $g(b)$:

$$g'(\xi) = \frac{g(b) - g(a)}{b - a} \quad (6.1)$$

La ecuación iterativa de búsqueda de la raíz es $x_{i+1} = g(x_i)$ y para la raíz verdadera x_r se cumple $x_r = g(x_r)$. Restando ambas ecuaciones se obtiene $x_r - x_{i+1} = g(x_r) - g(x_i)$.

Con el teorema del valor medio se puede expresar con ξ en el intervalo entre x_i y x_r

$$g'(\xi) = \frac{g(x_r) - g(x_i)}{x_r - x_i} \quad (6.2)$$

Por tanto $g'(\xi)(x_r - x_{i+1}) = g(x_r) - g(x_i)$. Combinado esto con $x_r - x_{i+1} = g(x_r) - g(x_i)$ se obtiene $x_r - x_{i+1} = g'(\xi)(x_r - x_i)$

6.2.2. Convergencia del método de punto fijo

Si el error verdadero en la i -ésima iteración es $E_{t,i} = x_r - x_i$ entonces: $E_{t,i+1} = g'(\xi)E_{t,i}$. De donde se deriva que la magnitud $|g'(\xi)|$ debe ser menor que uno para asegurar **convergencia lineal**. Si $|g'(\xi)| > 1$ entonces este método diverge.

6.3. Método de Newton-Raphson

Encontrar un x_r tal que $0 = f(x_r)$. Se define la derivada como $f'(x_i) = \frac{f(x_i)}{x_i - x_{i+1}}$, de donde se despeja la iteración del método Newton-Raphson, la cual requiere tanto de $f(x)$ como de su derivada $f'(x)$

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} \quad (6.3)$$

6.3.1. Estimación del error

$E_{t,i+1} = -\frac{f''(\xi)}{2f'(x_i)}E_{t,i}^2$ Si el método converge, lo hace con orden **cuadrático**: el número de cifras significativas decimales se duplica (aproximadamente) en cada iteración. El método en general **converge** si $|\frac{f''(\xi)}{2f'(x_i)}| < 1$

6.3.2. Problemas del método Newton-Raphson

- Como en la iteración la derivada de f aparece en el denominador, el método se inestabiliza si pasos intermedios caen cerca de máximos y mínimos locales.
- El método puede oscilar sin converger cuando puntos de inflexión ($f''(x) = 0$) cerca de la raíz.
- No hay criterio de convergencia general.
- Puede converger a puntos que no son la raíz, por lo que al final de proceso iterativo debe comprobarse que $f(x_r) \approx 0$

6.4. Método de la Secante

Equivale al método de Newton-Raphson sustituyendo la derivada por su aproximación hacia atrás: $f'(x_i) \approx \frac{f(x_{i-1}) - f(x_i)}{x_{i-1} - x_i}$. La **iteración** se convierte entonces en:

$$x_{i+1} = x_i - \frac{f(x_i)(x_{i-1} - x_i)}{f(x_{i-1}) - f(x_i)} \quad (6.4)$$

La convergencia del método es subcuadrática, orden ≈ 1.618 , es decir $E_{t,i+1} = \text{const} * E_{t,i}^{1.618}$

6.5. Método de Brent

- Garantiza convergencia si intervalo inicial encierra una raíz.
- Combina:
 - Interpolación inversa cuadrática
 - Secante
 - Bisección

6.6. Raíces de polinomios

Forma estándar de polinomios: $f_n(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots a_nx^n$ requiere $n(n+1/2)$ multiplicaciones y n sumas. **Forma anidada (esquema de Horner):** $f_n(x) = a_0 + x(a_1 + x(a_2 + x(\dots(a_{n-1} + a_nx))))$ requiere n multiplicaciones y n sumas. Al requerir menos operaciones, la forma anidada produce menos errores de redondeo.

6.7. Deflación polinomial

Proceso de reducir grado de polinomio dividiéndolo por $(x - t_i)$, t_i es la raíz conocida del polinomio. Las raíces pueden ser algunas o todas iguales, reales o complejas. Si el polinomio es divisible por t entonces el residuo $r = 0$.

División polinomial: Necesaria para tratar el caso de pares de raíces complejas conjugadas, que producen polinomios de segundo orden de coeficientes reales.

Pulir: es el proceso, luego de la deflación y obtener nuevas raíces, utilizar éstas como puntos iniciales en la búsqueda de raíces con **con el polinomio original**.

6.8. Método de Müller

Similar a Secante, pero utiliza interpolación cuadrática para encontrar la siguiente estimación de la raíz.

La aproximación cuadrática centrada en x_i : $\tilde{f}(x) = a(x - x_i)^2 + b(x - x_i) + c$
a partir de tres puntos x_{i-2} , x_{i-1} y x_i se tiene

$$h_{i-2} = x_{i-1} - x_{i-2}$$

$$\delta_{i-2} = \frac{f(x_{i-1}) - f(x_{i-2})}{x_{i-1} - x_{i-2}}$$

$$h_{i-1} = x_i - x_{i-1}$$

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

Coeficientes dados por:

$$a = \frac{\delta_{i-1} - \delta_{i-2}}{h_{i-1} - h_{i-2}}$$

$$b = ah_{i-1} + \delta_{i-1}$$

$$c = f(x_i)$$

La estimación de la raíz se realiza con: $x_{i+1} = x_i - \frac{2c}{b \pm \sqrt{b^2 - 4ac}}$. Las raíces pueden ser complejas.

Problemas del método: Produce dos raíces. Establece elegir signo que coincida con el signo de b para obtener la raíz estimada más cercana.

7. Lección 7: Optimización unidimensional

7.1. Optimización

- Proceso de encontrar extremos de una función.
- Se distinguen extremos **locales y globales**.
- Extremos son **mínimos o máximos**.
- Optimizadores unidimensionales son cerrados o abiertos.

7.2. Método cerrado: Sección dorada

Similar a bisección. Recibe x_l : límite inferior y x_u : límite superior.

Se necesitan otros dos puntos dentro del intervalo $[x_l, x_u]$, el algoritmo selecciona si el punto extremo esta en primeros o segundos 3 puntos.

Elección de puntos intermedios es crítica para eficiencia: **Meta:** minimizar número de evaluaciones de función.

$$\text{Razón aurea: } R = \frac{\sqrt{5} - 1}{2} \approx 0,61803$$

$$d = R(x_u - x_l) \quad x_1 = x_l + d \quad x_2 = x_u - d$$

- Si $f(x_1) > f(x_2)$ entonces máximo en $[x_2, x_u]$
- Si $f(x_2) > f(x_1)$ entonces máximo en $[x_l, x_1]$

Error se aproxima con tamaño de intervalo, normalizado por el óptimo estimado hasta el momento: $e_a = (1 - R) \left| \frac{x_u - x_l}{x_{opt}} \right| * 100\%$

7.3. Método cerrado: Interpolación parabólica

Con los 3 puntos disponibles interpolar parábola y usar máximo de esa parábola como siguiente aproximación.

Resolviendo resulta en $(x_1, x_2$ y x_3 triada inicial):

$$x_4 = \frac{1}{2} * \frac{f(x_1)(x_2^2 - x_3^2) + f(x_2)(x_3^2 - x_1^2) + f(x_3)(x_1^2 - x_2^2)}{f(x_1)(x_2 - x_3) + f(x_2)(x_3 - x_1) + f(x_3)(x_1 - x_2)}$$

Estrategias: **(1)** Utilizar los puntos en orden secuencial (los últimos 3 siempre). **(2)** Usar los dos más cercanos al nuevo punto.

7.4. Método abierto: Método de Newton

Si x_{i+1} es extremo, entonces $f'(x_{i+1}) = 0$ y por tanto: $x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$

- Idéntido a Newton-Raphson pero aplicado a $f'(x)$
- Puede diverger
- Debe verificarse signo de segunda derivada para asegurar extremo correcto

7.5. Método híbrido: Método de Brent

- **Idea:** Combinar sección dorada (segura pero lenta) con interpolación parabólica.
- **Función se comporta mal:** Usar sección dorada
- **Función es suave:** Usar interpolación parabólica
- Combinación de métodos debe:
 - Evitar evaluaciones innecesarias cuando se cambia de método
 - Cuidado en fase final, con error estimado (cerca de precisión numérica)
 - Debe tener detección robusta de tipo de función para seleccionar el método apropiado

8. Lección 8: Optimización multidimensional

8.1. Métodos sin gradiente

8.1.1. Búsqueda aleatoria

- Genere números aleatorios con alguna distribución probabilística en el rango de definición y evalúe la función en dichos puntos.
- Selecciones valor máximo
- **Ventajas:**
 - Funciona con discontinuidades y funciones sin derivadas
 - No se atasca en extremos locales.
- **Desventajas:**
 - Ineficiente
 - Excesivo número de evaluaciones de la función.

8.1.2. Búsqueda univariable y búsquedas patrón

- Realice búsquedas lineales de máximos alternando los ejes.
- Emplea métodos ya vistos (Brent).
- Direcciones patrón apuntan en dirección del máximo.

Método de Powell: Explota direcciones patrón para encontrar máximo.

8.1.3. Simplex de bajada

- También conocido como "descenso de ameba" o algoritmo de Nelder-Mead.
- Un **simplex** es la generalización de un triángulo en espacio d-dimensional tiene $d + 1$ vértices.
- Se inicia con cualquier simplex no degenerado (con volumen no nulo).
- El proceso paso a paso mueve un vértice del simplex a la vez para acercarse al extremo.

Modificaciones del simplex:

- Vértice con el peor valor asociado se **refleja**, pasándolo al otro lado del simplex pero manteniendo su volumen.
- Si la reflexión lleva a un mejor valor, hay expansión, y sino, a compresión del simplex.

8.2. Métodos con gradiente

8.2.1. Gradiente

El gradiente $\nabla f(\underline{x})$ de una función indica la dirección de mayor cambio de la función:
$$\nabla f(\underline{x}) = \left[\frac{\partial f(\underline{x})}{\partial x_1} \frac{\partial f(\underline{x})}{\partial x_2} \dots \frac{\partial f(\underline{x})}{\partial x_n} \right]^T \text{ con } \underline{x} = [x_1, x_2, \dots, x_n]^T.$$

Los extremos ocurren en puntos donde no hay cambio ($\nabla f(\underline{x}) = 0$).

8.2.2. Matriz Hessiana

Con la segunda derivada en el caso unidimensional se determina si un extremo es máximo o mínimo.

Por ejemplo, si $f'(t_0) = 0$ entonces:

- $f(t_0)$ es máximo si $f''(t_0) < 0$
- $f(t_0)$ es mínimo si $f''(t_0) > 0$

El equivalente de la segunda derivada multidimensional es la **matriz Hessiana** $H(\underline{x})$ (o hessiano).

Para los extremos de $f(\underline{x})$, estos se encuentran donde $\nabla f(\underline{x}) = 0$ y

- Si $|H| > 0$ y $\frac{\partial^2 f}{\partial x^2} > 0$, entonces $f(x, y)$ tiene un mínimo local.

- Si $|H| > 0$ y $\frac{\partial^2 f}{\partial x^2} < 0$, entonces $f(x,y)$ tiene un máximo local.
- Si $|H| < 0$, $f(x,y)$ tiene un punto de silla.

8.2.3. Máxima inclinación

Estrategia: Seguir dirección del gradiente para maximizar (opuesta para minimizar).