

INSTITUTO DE INGENIERÍA MATEMÁTICA Y COMPUTACIONAL

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE



IMT3410: Métodos para ecuaciones diferenciales

Clase 11

Manuel A. Sánchez 2024.09.11

Métodos de diferencias finitas

Métodos de diferencias finitas

Un método numérico clásico para resolver problemas de valores de frontera es el método de diferencias finitas, el cual reemplaza de forma directa los operadores diferenciales, o derivadas, por expresiones de **diferencias finitas** e impone una version discreta del problema en algunos puntos o grilla del dominio. Esta discretización da origen a un sistema lineal o no lineal de ecuaciones con solución la aproximación en los puntos de la grilla.

Vamos a considerar esquemas de una grilla o malla uniforme de un intervalo finito (a, b), esto es:

$$\{x_n\}_{n=0}^{N+1}: \qquad a = x_0 < x_1 < \ldots < x_N < x_{N+1}, \quad x_n = a + nh; \ h = \frac{b-a}{N+1}$$

Definimos funciones de grilla $u \in \Gamma_h[a, b]$ como $u : \{x_n\} \mapsto \mathbb{R}^{N+1}$.

Ecuaciones lineales de segundo orden

Consideraremos el problema de Sturm-Liouville

$$L(y) = r(x), \quad a \le x \le b, \qquad L(y) := -y'' + p(x)y' + q(x)y$$

con condiciones de frontera simples

$$y(a) = \alpha, \ y(b) = \beta.$$

Asumimos que existen constantes positivas \bar{p} , q y \bar{q} tales que

$$|p(x)| \leq \bar{p}, \quad 0 < q \leq q(x) \leq \bar{q}, \quad a \leq x \leq b.$$

Operador de diferencias finitas: Definimos para una función de grilla $u \in \Gamma_h[a, b]$

$$(L_h u)_n = -\frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} + p(x_n) \frac{u_{n+1} - u_{n-1}}{2h} + q(x_n)u_n$$
, para $n = 1, 2, ..., N$.

Error de truncación

Definición

Para toda función suficientemente suave v definida sobre [a,b] y operador diferencial L y operador de diferencias finitas L_h asociado a la grilla $\{x_n\}$ definimos el **error de truncación** T_h por

$$(T_h v)_n = (L_h v)_n - (L v)(x_n), \quad n = 1, 2, ..., N$$

Tenemos que, para el operador asociado al problema de Sturm Liouville tenemos, para $v \in C^4([a,b])$

$$(T_h v)_n = -\frac{h^2}{12} \left(v^{(4)} (\xi_1 - 2p(x_n) v^{(3)}(\xi_2)), \quad \xi_1, \xi_2 \in [x_n - h, x_n + h], \right)$$

o si $v \in C^6([a,b])$

$$(T_h v)_n = -\frac{h^2}{12} \left(v^{(4)}(x_n) - 2p(x_n)v^{(3)}(x_n) \right) + \mathcal{O}(h^4), \quad h \to 0.$$

Estabilidad

Definición

Decimos que un operador de diferencias L_h es **estable** si existe una constante M, independiente de h, tal que para h suficientemente pequeño, tenemos para toda función de grilla $v = \{v_n\}$

$$||v||_{\infty} \leq M\left(\max\{|v_0|,|v_{N+1}|+||L_h(v)||_{\infty}\right)$$

Donde

$$\|v\|_{\infty} = \max_{0 \le n \le N+1} |v_n|, \quad \|L_h(v)\|_{\infty} = \max_{1 \le n \le N} |(L_h v)_n|$$

Es el operador de nuestro operador de diferencias estable?

Teorema, operador estable

Teorema

Si $h\bar{p} \le 2$, entonces el operador L_h (aprox. S-L) es estable. En efecto, la constante de estabilidad $M=\max\{1,1/q\}$

Ver ejemplo numérico de cuando las condiciones de estabilidad no se satisface

Demostración

Reescribirmos el operador de diferencias por

$$\frac{h^2}{2}(L_h v)_n = a_n v_{n-1} + b_n v_n + c_n v_{n+1}, \quad n = 1, 2, ..., N$$

donde

$$a_n = -rac{1}{2}\left(1 + rac{1}{2}hp(x_n)
ight), \quad b_n = 1 + rac{1}{2}h^2q(x_n), \quad c_n = -rac{1}{2}\left(1 - rac{1}{2}hp(x_n)
ight).$$

Por el supuesto del teorema, tenemos que $\frac{1}{2}h|p(x_n)| \leq \frac{1}{2}h\bar{p} \leq 1$, y así $a_n \leq 0$, $c_n \leq 0$ y

$$|a_n| + |c_n| = \frac{1}{2} \left(1 + \frac{1}{2} h p(x_n) \right) + \frac{1}{2} \left(1 - \frac{1}{2} h p(x_n) \right) = 1$$

Demostración continuación

Además, tenemos que $b_n \geq 1 + \frac{1}{2}h^2q$. Observemos que

$$b_n v_n = -a_n v_{n-1} - c_n v_{n+1} + \frac{1}{2} h^2 (L_h v)_n,$$

Entonces, se sigue que

$$\left(1 + \frac{1}{2}h^2\underline{q}\right)|v_n| \le \|v\|_{\infty} + \frac{1}{2}h^2\|L_hv\|_{\infty}, \quad n = 1, 2, ..., N$$

Observe que $||v||_{\infty} = \max\{|v_0|, |v_{N+1}|\}$, entonces M = 1.

Por otro lado, si v alcanza el máximo en el interior, esto es, $||v||_{\infty} = |v_{n_0}|$, $1 \le n_0 \le N$, se sigue

$$\left(1 + \frac{1}{2}h^2\underline{q}\right)|v_{n_0}| \le |v_{n_0}| + \frac{1}{2}h^2||L_hv||_{\infty}, \quad n = 1, 2, ..., N$$

lo que implica que

$$||v||_{\infty} = |v_{n_0}| \leq \frac{1}{q} ||L_h v||_{\infty}.$$

Métodos de diferencias finitas

Una aproximación por diferencias finitas del problema de Sturm Liouville es la función de grilla $u = \{u_n\}$ solución del problema:

$$(L_h u)_n = r(x_n), \quad n = 1, ..., N, \quad u_0 = \alpha, \ u_{N+1} = \beta$$

Sistema lineal asociado: (demuestre que la matriz es estrictamente diagonal dominante)

$$\begin{bmatrix} b_{1} & c_{1} & 0 & \dots & 0 \\ a_{2} & b_{2} & c_{2} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & a_{N-1} & b_{N-1} & c_{N-1} \\ 0 & \dots & 0 & a_{N} & b_{N} \end{bmatrix} \begin{bmatrix} u_{1} \\ u_{2} \\ \vdots \\ u_{N-1} \\ u_{N} \end{bmatrix} = \frac{1}{2}h^{2} \begin{bmatrix} r(x_{1}) \\ r(x_{2}) \\ \vdots \\ r(x_{N-1}) \\ r(x_{N}) \end{bmatrix} - \begin{bmatrix} a_{1}\alpha \\ 0 \\ \vdots \\ 0 \\ c_{N}\beta \end{bmatrix}$$

Métodos de diferencias finitas

Teorema

Asuma que L_h es estable. Entonces, el problema de diferencias finitas tiene una única solución o equivalentemente la matriz de diferencias finitas es no singular.

Demostración. Observe que el problema homogenero asociado, esto es, r(x) = 0 y $\alpha = \beta = 0$ pueden sólo tener la solución trivial dado que $L_h u = 0$ y $u_0 = u_{N+1} = 0$, lo que implica, por la condición de estabilidad que $u_n = 0$, n = 0, 1, ..., N + 1.

Aproximación

Teorema

Si $h\bar{p} \leq 2$, entonces

$$||u - y||_{\infty} \le M ||T_h y||_{\infty}, \quad M = \max\{1, 1/q\}$$

donde $u = \{u_n\}$ es la solución de diferencias finitas, $y = \{y_n = y(x_n)\}$ es la función de grilla inducida por la solución exacta y(x) del problema y T_h el error de truncación. Si $y \in C^4([a,b])$, entonces

$$||u-y||_{\infty} \le \frac{1}{12} h^2 M \left(||y^{(4)}||_{\infty} + 2\bar{p} ||y^{(3)}||_{\infty} \right)$$

Demostración

Manuel A. Sánchez 14/18

Teorema de aproximación

Teorema

Sean $p, q \in C^2([a, b])$, $y \in C^6([a, b])$, $y h\bar{p} \leq 2$. Entonces,

$$u_n - y(x_n) = h^2 e(x_n) + \mathcal{O}(h^4), \quad n = 0, 1, ..., N + 1,$$

donde e(x) es la solución de

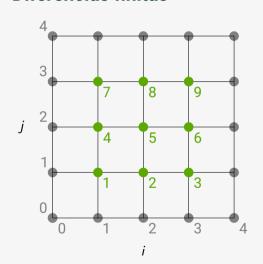
$$Le = \theta(x), \quad a \le x \le b; \quad e(a) = 0, \ e(b) = 0,$$

con
$$\theta(x) = \frac{1}{12}(y^{(4)}(x) - 2p(x)y^{(3)}(x).$$

Demostración

Manuel A. Sánchez 16/18

Diferencias finitas



$$\begin{array}{c} (1,1) \longrightarrow 1 \\ (2,1) \longrightarrow 2 \\ (3,1) \longrightarrow 3 \\ (1,2) \longrightarrow 4 \\ \dots \\ (3,3) \longrightarrow 9. \end{array}$$

En una dimensión tenemos que la ecuación de Poisson toma la siguiente forma

$$\begin{cases} -u''(x) = f(x), x \in (0,1) \\ u(0) = \alpha \\ u(1) = \beta \end{cases}$$

Por lo cual, usando diferencias finitas vemos que

$$\begin{cases} -\frac{1}{h^2} \left(u_{j-1} - 2u_j + u_{j+1} \right) = f\left(x_j \right) \\ u_0 = \alpha \\ u_n = \beta \end{cases}$$

para $1 \le jn$. El cual, puede ser reescrito de la forma $A_{1D}u = F$, donde $u = (u_i)_{i=1}^n$,

$$A_{1D} = rac{1}{h^2} \left[egin{array}{cccccc} 2 & -1 & 0 & \cdots & 0 \ -1 & 2 & -1 & \ddots & dots \ 0 & \ddots & \ddots & \ddots & 0 \ dots & \ddots & -1 & 2 & -1 \ 0 & \cdots & 0 & -1 & 2 \end{array}
ight]$$

$$F = \begin{bmatrix} f(x_1) - \alpha/h^2 \\ f(x_2) \\ \vdots \\ f(x_{n-2}) \\ f(x_{n-1}) - \beta/h^2 \end{bmatrix}$$

Para este caso, tenemos que el error de truncación viene dado por la siguiente expresión

$$T_{j} = -\frac{1}{h^{2}} \left(u\left(x_{j-1} \right) - 2u\left(x_{j} \right) + u\left(x_{j+1} \right) \right) - f\left(x_{j} \right)$$

y usando series de Taylor vemos que

$$egin{aligned} T_{j} &= \left(-u''\left(x_{j}
ight) - rac{1}{12}h^{2}u^{\left(4
ight)}\left(x_{j}
ight) + O\left(h^{4}
ight)
ight) - f\left(x_{j}
ight) \ &= -rac{1}{12}h^{2}u^{\left(4
ight)}\left(x_{j}
ight) + O\left(h^{4}
ight) = O\left(h^{2}
ight). \end{aligned}$$

cuando h o 0, por lo cual obtenemos un error de orden dos. En este caso, si

$$e_{j} = u_{j} - u(x_{j}) : A_{1D} = -T$$
, entonces

$$e(x) \approx +\frac{1}{12}h^2u'' - \frac{1}{12}h^2(u''(0) + x(u''(1) - u''(0))$$

 $y e_0 = e_n = 0$. Por lo cual, si $T(x) \approx -\frac{1}{12}h^2u^{(4)}(x)$ entonces

 $\sim O(h^2)$ Para trabajar la estabilidad denotemos por $A = A^h$; $e = e^h$; $T = T^h$, el súper

 $-\frac{1}{-2}(e_{j-1}-2e_j+e_{j+1})=-T_j, 1\leqslant j\leqslant n-1$

índice, denota una dependencia de h. De esta forma.

$$egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$$

Entonces para $\|e^h\| \sim O(h^2)$ necesitamos $\|(A^h)^{-1}\| \leqslant C$.

$$\mathsf{Imos} \, ig\| ig(A^h ig) \quad ig\| \leqslant 0$$

Definition

Suponga que un método de diferencias finitas para un problema de valores de frontera lineal tiene una forma matricial $A^h u^h = F^h$, para h tamaño de malla.

Décimo que el método es estable si $\left(A^h\right)^{-1}$ existe para todo h suficientemente pequeño y existe una constante C, independiente de h, tal que

$$||(A^n)^{-1}|| \leqslant C \quad \forall h < h_0.$$

Decimos que el método es consistente con la ecuación diferencial y las condiciones de frontera si

$$||T^h|| \longrightarrow 0$$
 cuando $h \longrightarrow 0$.

Decimos que el método es convergente si

$$||e^h|| \longrightarrow 0$$
 cuando $h \longrightarrow 0$.

De esta forma, presentamos el siguiente Teorema

Theorem

Teorema fundamental de métodos de diferencias finitas Si el problema es es consistente y estable entonces es convergente.

Recordemos la definición de la norma 2

$$||A||_2 = \rho(A) := \max_{1 \leqslant j \leqslant n-1} |\lambda_j|$$

y para la inversa tenemos lo siguiente

$$\left\|A^{-1}\right\|_{2} = \rho\left(A^{-1}\right) = \max_{1 \leq j \leqslant n-1} \left|\lambda_{j}^{-1}\right| = \left(\min_{1 \leqslant j \leqslant n-1} \left|\lambda_{j}\right|\right)^{-1}.$$

Necesitamos mostrar que los valores propios de A están acotados por abajo cuando $h \to 0$.

 $\lambda_j = \frac{2}{h^2}(\cos(j\pi \cdot h) - 1), j = 1, \dots, n-1$

 $u_i^{j_i} = \sin(j\pi i h).$

y los vectores propios
$$u^j$$
 asociados a λ_j son

Así, se satisface que $A^h u^j = \lambda_j u^j$, j = 1, ..., n - 1El menor valor propio en magnitud de A^h es

$$egin{align} \lambda_1 &= rac{\pi}{h^2}(\cos(\pi h) - 1) \ &= rac{2}{h^2} \left(-rac{1}{2!} \pi^2 h^2 + rac{1}{4!} \pi^4 h^4 + O\left(h^6
ight)
ight) \ &= -\pi^2 + O\left(h^2
ight) \ \end{aligned}$$

Ejercicios: Para $h = \frac{1}{2} \log n - 1$ valores propios de *A* son

el cual esta acotado lejos de cero cuando $h \to 0$. Por lo tanto el método es estable

$$\left\|e^h
ight\|_2 \sim rac{1}{\pi^2} \left\|T^h
ight\|_2.$$

Además, observe que el vector propio u^j es cercano a la función propia del operador. Para el sistema,

$$\begin{cases} u'' = \mu u \\ u(0) = 0 \\ u(1) = 0 \end{cases}$$

Entonces las funciones propias y valores propios son

$$u^j=\sin(j\pi x); \mu_j=-j^2\pi^2, \quad j\in \mathbb{Z}.$$

Los valores propios λ_i no son exactamente u_i , pero tenemos

$$\lambda_j = rac{2}{h^2} \left(-rac{1}{2} j^2 \pi^2 h^2 + rac{1}{4!} j^4 \pi^4 h^4 + ..
ight) = \mu_j + O(h^2).$$

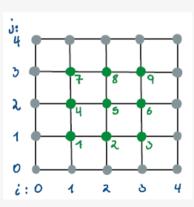
En este caso, la ecuación es

$$\begin{cases} \Delta u = f(x), x \in \Omega \\ u = g \in \partial \Omega \end{cases}$$

Luego, usando diferencias finitas vemos que

$$\frac{u_{i-1j} + u_{i+1j} - 4u_{ij} + u_{ij-1} + u_{ij+1}}{h^2} = f_{ij} \quad 1 \leqslant i, j \leqslant n-1$$

con $u_{ij} = g_{ij}$ si $(i = 0 \lor i = n) \lor (j = 0 \lor j = n)$. Entonces, gráficamente tenemos que el esquema es



de donde obtenemos la siguiente relación

$$egin{array}{l} (2,1) \longrightarrow 2 \ (3,1) \longrightarrow 3 \ & \ldots \ (3,3) \longrightarrow 9. \end{array}$$

Para este caso, la matriz A_{2D} es

$$A_{2D} = \frac{1}{h^2} \begin{bmatrix} -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -4 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -4 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -4 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -4 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 & 0 \end{bmatrix}$$

De esta forma, vemos que la matriz puede ser vista en 9 bloques, esto es

$$A_{2D} = rac{1}{h^2} \left[egin{array}{cccc} T & I & 0 \ I & T & I \ 0 & I & T \end{array}
ight]; T = \left[egin{array}{cccc} -4 & 1 & 0 \ 1 & -4 & 1 \ 0 & 1 & -4 \end{array}
ight]; A_{1D} = \left[egin{array}{cccc} -2 & 1 & 0 \ 1 & -2 & 1 \ 0 & 1 & -2 \end{array}
ight]$$

Más aún.

$$A_{2D} = rac{1}{h^2} \left(\left[egin{array}{ccc} A_{1D} & 0 & 0 \ 0 & A_{1D} & 0 \ 0 & 0 & A_{1D} \end{array}
ight] + \left[egin{array}{ccc} -2I & I & 0 \ I & -2I & I \ 0 & I & -2I \end{array}
ight]
ight)$$

$$A_{2D} = \frac{1}{h^2} \left(\begin{bmatrix} A_{1D} & 0 & 0 \\ 0 & A_{1D} & 0 \\ 0 & 0 & A_{1D} \end{bmatrix} + \begin{bmatrix} -2I & I & 0 \\ I & -2I & I \\ 0 & I & -2I \end{bmatrix} \right)$$
$$= \frac{1}{h^2} \left(I \otimes A_{1D} + A_{1D} \otimes I \right).$$

Donde ⊗ denota el producto kronocker

$$\otimes: \mathbb{R}^{m_1 \times n_1} \times \mathbb{R}^{m_2 \times n_2} \longrightarrow \mathbb{R}^{(m_1 m_2) \times (n_1 n_2)}$$

$$B, C \longrightarrow B \otimes C$$

 $B \otimes C$ es una matriz por bloques de $m_1 \times n_1$ bloques con cada bloque $b_{ij} C \in \mathbb{R}^{m_2 \times n_2}$.

Ejercicios: Muestre que para el problema de Poisson en \mathbb{R}^3

$$-\Delta u(x) = f(x), \quad x = (x_1, x_2, x_3) \in (a, b) \subset \mathbb{R}^3,$$

el método resultante de aplicar el stencil de 3 puntos en cada dirección tiene como matriz

$$A_{3D} = A_{1D} \otimes I \otimes I + I \otimes A_{1D} \otimes I + I \otimes I \otimes A_{1D}.$$

El error de truncación local viene dado por

$$T_{ij} = -\frac{1}{h^2} \left(u(x_{i-1}, y_j) + u(x_{i+1}, y_j) - 4u(x_i, y_j) + u(x_i, y_{j-1}) + u(x_i, y_{j+1}) \right) - f(x_i, y_j)$$

$$= -\frac{1}{12} h^2 \left(\frac{\partial^4 u}{\partial x^4} + \frac{\partial^4 u}{\partial y^4} \right) + o(h^4)$$

El error global viene dado por, para $1 \le i, j \le n-1$

$$e_{ii} = u_{ii} - u(x_i, y_i) \Rightarrow A^h e^h = -T^h$$
.

En este caso los valores propios son

$$\lambda_{jk} = \frac{2}{h^2}((\cos(j\pi h) - 1) + (\cos(k\pi h) - 1)),$$

por ende los vectores propios son

$$u_{il}^{jk} = \sin(j\pi i h) \sin(k\pi l h).$$

Por lo tanto, el menor valor propio en modulo es

$$\lambda_{11} = -2\pi^2 + o(h^2).$$

Así,

$$\rho\left(\left(\mathcal{A}^h\right)^{-1}\right) = \frac{1}{\lambda_{11}} \approx \frac{1}{2\pi^2}$$

y el método es estable. Calculemos tambien el número de condición de A^h

$$k_2(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{|\lambda_{n-1}|_{n-1}}{|\lambda_{n-1}|} \approx \frac{8/h^2}{2\pi^2} = \frac{4}{\pi^2 h^2}$$

donde vemos que la matriz se vuelve mal condicionada cuando $h o 0$.					



INSTITUTO DE INGENIERÍA Matemática y computacional

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE