Modellistica Numerica

Università di Pavia

2 luglio 2019 — ore 10:00 — aula E9

Punti totali: 33 Durata dell'esame: 2 ore.

Problema 1. Differenze finite per il problema periodico [20 punti]

Consideriamo l'equazione di diffusione–reazione con condizioni al bordo periodiche:

$$-u'' + qu = f \text{ in } (a,b), \qquad u(a) = u(b), \quad u'(a) = u'(b) \tag{1}$$

dove $(a,b) \subset \mathbb{R}$, $q, f \in C^0([a,b])$. Assumiamo che $q^* := \min_{x\in[a,b]} q(x) > 0$.

Dato $n \in \mathbb{N}$, definiamo $h := \frac{b-a}{n+1}$ e fissiamo i nodi equispaziati $x_j := a + jh$, per $j = 0, \dots, n+1$.

(a) Scrivere il sistema lineare $\underline{\underline{\mathbf{A}}}\vec{\mathbf{U}} = \vec{\mathbf{f}}$ corrispondente al metodo delle differenze finite centrate per il problema al bordo (1).

Attenzione: la matrice $\underline{\underline{\mathbf{A}}}$ non è tridiagonale.

(b) Mostrare che la matrice del metodo delle differenze finite è invertibile.

Suggerimento: dimostrare che $\underline{\underline{\mathbf{A}}}$ appartiene ad una classe di matrici descritte a lezione.

(c) Dimostrare che tutti gli autovalori di $\underline{\underline{\mathbf{A}}}$ sono reali, positivi e maggiori o uguali a q_*.

Suggerimento: usare il teorema dei cerchi di Gershgorin.

(d) Dedurre una maggiorazione per la norma dell'inversa $\left\|\underline{\underline{\mathbf{A}}}^{-1}\right\|_2$.

Suggerimento: ricordare la relazione tra la norma di una matrice simmetrica ed i suoi autovalori.

(e) Assumiamo che la soluzione u del problema (1) sia di classe C^4. Definire il vettore $\vec{\mathbf{T}}$ dell'errore di troncamento e stimare la sua dipendenza dal passo h.

(f) Dedurre dai passi precedenti una stima dell'errore $|u(x_j) - U_j|$ del metodo delle differenze finite.

Misurare l'errore in una norma vettoriale opportuna.

(g) Scrivere $\underline{\underline{\mathbf{A}}}$ come somma di due matrici $\underline{\underline{\mathbf{A}}} = \underline{\underline{\mathbf{T}}} + \underline{\underline{\mathbf{P}}}$, dove $\underline{\underline{\mathbf{T}}}$ è tridiagonale e $\underline{\underline{\mathbf{P}}}$ ha rango uno.

Cosa implica questo fatto per la risoluzione del sistema lineare corrispondente?

(h) Usare il metodo dell'energia per dimostrare che il problema al bordo (1) ammette al più una soluzione $u \in C^2([a,b])$.

Soluzione: (a) Con $f_j := f(x_j)$, $q_j := q(x_j)$,

$$\underline{\underline{\mathbf{A}}} = \frac{1}{h^2}\begin{pmatrix} 2+q_1h^2 & -1 & & & & -1 \\ -1 & 2+q_2h^2 & -1 & & & \\ & -1 & 2+q_3h^2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2+q_nh^2 & -1 \\ -1 & & & & -1 & 2+q_{n+1}h^2 \end{pmatrix}, \quad \vec{\mathbf{b}} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_n \\ f_{n+1} \end{pmatrix}, \vec{\mathbf{U}} = \begin{pmatrix} U_1 \\ U_2 \\ U_3 \\ \vdots \\ U_n \\ U_{n+1} \end{pmatrix}.$$

(b) Per ogni $j = 1, \dots, n+1$ vale

$$A_{j,j} = \frac{2}{h^2} + q_j \geq \frac{2}{h^2} + q_* > \frac{2}{h^2} = |A_{j,j-1}| + |A_{j,j+1}| = \sum_{k\neq j} |A_{j,k}|,$$

cioè la matrice è a predominanza diagonale stretta e quindi invertibile (qui abbiamo usato la convenzione degli indici periodici $A_{1,0} = A_{1,n+1}$ e $A_{n+1,n+2} = A_{n+1,1}$).

(c) Gli autovalori sono reali poiché la matrice è simmetrica. Inoltre appartengono all'unione dei cerchi di centro $A_{j,j} = \frac{2}{h^2} + q_j$ e raggio $|A_{j,j-1}| + |A_{j,j+1}| = \frac{2}{h^2}$, quindi $\min\{\lambda \text{ autovalore di } \underline{\underline{\mathbf{A}}}\} \geq \min_j(\frac{2}{h^2} + q_j - \frac{2}{h^2}) = \min_j q_j \geq q_*$.

(d) Usando ancora la simmetria di $\underline{\underline{\mathbf{A}}}$:

$$\left\|\underline{\underline{\mathbf{A}}}^{-1}\right\|_2 = \rho(\underline{\underline{\mathbf{A}}}^{-1}) = (\max\{\mu \text{ autovalore di } \underline{\underline{\mathbf{A}}}^{-1}\}) = (\min\{\lambda \text{ autovalore di } \underline{\underline{\mathbf{A}}}\})^{-1} \leq q_*^{-1}.$$

(e) Definiamo il vettore $\vec{\mathbf{u}} \in \mathbb{R}^{n+1}$ del valore della soluzione esatta nei nodi $((\vec{\mathbf{u}})_j = u(x_j))$ e il vettore dell'errore di troncamento $\vec{\mathbf{T}} = \underline{\underline{\mathbf{A}}}\vec{\mathbf{u}} - \vec{\mathbf{b}}$. Allora, usando l'errore di troncamento di D_h^{2C} e l'equazione differenziale soddisfatta da u, vale:

$$T_j = -D_h^{2C} u(x_j) + q(x_j)u(x_j) - f(x_j) = -u''(x_j) - \frac{h^2}{12}u^{(iv)}(\xi_j) + q(x_j)u(x_j) - f(x_j) = -\frac{h^2}{12}u^{(iv)}(\xi_j).$$

(f) L'errore è legato al troncamento da $\vec{\mathbf{e}} := \vec{\mathbf{u}} - \vec{\mathbf{U}} = \underline{\underline{\mathbf{A}}}^{-1}\vec{\mathbf{T}}$. Quindi la sua norma è controllata come

$$\|\vec{\mathbf{e}}\|_2 \quad \leq \quad \left\|\underline{\underline{\mathbf{A}}}^{-1}\right\|_2 \left\|\vec{\mathbf{T}}\right\|_2 \quad \leq \quad \left\|\underline{\underline{\mathbf{A}}}^{-1}\right\|_2 \sqrt{n+1}\left\|\vec{\mathbf{T}}\right\|_\infty \quad \leq \quad q_*^{-1}\sqrt{n+1}\frac{1}{12}h^2 \left\|u^{(iv)}\right\|_{L^\infty(a,b)}.$$

(g) Ad esempio $\underline{\underline{\mathbf{A}}} = \underline{\underline{\mathbf{T}}} + \vec{\mathbf{u}}\vec{\mathbf{w}}^\top$ con $\vec{\mathbf{u}} = (-\frac{1}{h^2}, 0, \ldots, 0, -\frac{1}{h^2})^\top$, $\vec{\mathbf{w}} = (1, 0, \ldots, 0, 1)^\top \in \mathbb{R}^{n+1}$, mentre $\underline{\underline{\mathbf{T}}}$ ha valori $-\frac{1}{h^2}$ nella sopra- e sotto-diagonale e $\mathrm{diag}\underline{\underline{\mathbf{T}}} = (\frac{3}{h^2} + q_1,\ \frac{2}{h^2} + q_2,\ \ldots,\ \frac{2}{h^2} + q_n,\ \frac{3}{h^2} + q_{n+1})^\top$.

Questo permette di risolvere il sistema diagonale con complessità lineare in n, risolvendo successivamente due sistemi lineari per $\underline{\underline{\mathbf{T}}}$.

(h) Siano u_1, u_2 due soluzioni di (1). Detta $w = u_1 - u_2$, allora $-w'' + qw = -u_1'' - u_2'' + qu_1 + qu_2 = f - f = 0$ quindi, moltiplicando questa equazione per w stessa e integrando sull'intervallo,

$$0 = \int_a^b (-w'' + qw)w \,\mathrm{d}x = \int_a^b (w'w' + qww) \,\mathrm{d}x - w'(b)w(b) + \underbrace{w'(a)}_{=w'(b)}\underbrace{w(a)}_{=w(b)} = \int_a^b \left((w')^2 + qw^2\right) \mathrm{d}x \geq q^* \|w\|^2_{L^2(a,b)},$$

quindi $w = 0$ e $u_1 = u_2$.

Problema 2. Metodo di shooting [13 punti]

(a) Descrivere brevemente il metodo di *shooting* per il problema al bordo non lineare

$$u''(x) = F\big(x, u(x), u'(x)\big) \quad x \in (a,b), \qquad u(a) = \alpha, \quad u(b) = \beta.$$

(b) Fissiamo $p, q, r \in C^0([a,b])$ con $q \geq 0$. Consideriamo il problema lineare

$$u''(x) = p(x)u'(x) + q(x)u(x) + r(x) \quad x \in (a,b), \qquad u(a) = \alpha, \quad u(b) = \beta.$$

Scrivere una iterazione del metodo di shooting accoppiato con il metodo di *Newton*, scegliendo come primo valore del dato iniziale $u_1'(a) = s_1 = 0$.

Assumendo che il problema ai valori iniziali corrispondente venga risolto esattamente, mostrare che una sola iterazione del metodo di Newton permette di trovare il valore esatto di $u'(a)$.

(c) Calcolare due iterazioni del metodo di shooting combinato al metodo di *bisezione* applicato al problema

$$u''(x) = 2\frac{\big(u'(x)\big)^2}{u(x)} \quad x \in (0,1), \qquad u(0) = 1, \quad u(1) = \frac{1}{10}.$$

Come valori iniziali per $u'(0)$ scegliere $s_0 = -20$ e $s_1 = 0$.

Al posto della risoluzione numerica dei problemi ai valori iniziali, usare il fatto che la soluzione del problema ai valori iniziali per l'equazione differenziale considerata con $u(0) = 1$ e $u'(0) = -C$ è $u(x) = 1/(1 + Cx)$, per qualsiasi $C \geq 0$.

Soluzione: (a) Sia $U(x, s)$ la soluzione del problema ai valori iniziali

$$U''(x; s) = f\big(x, U(x; s), U'(x; s)\big) \quad x \in (a,b), \qquad U(a; s) = \alpha, \quad U'(a; s) = s,$$

al variare di un parametro $s \in \mathbb{R}$. Definiamo la funzione reale $\varphi(s) := U(b, s) - \beta$. Avremo che $U(\cdot, s)$ coincide con la soluzione u del problema al bordo se $\varphi(s) = 0$. Il metodo di shooting consiste nella ricerca di uno zero di φ attraverso un metodo numerico. Ogni valutazione di φ richiede la soluzione di un problema ai valori iniziali, che viene calcolata per mezzo di un metodo numerico di avanzamento in tempo.

(b) Sia $u_1(x) = U(x, s_1) = U(x, 0)$ soluzione di

$$u_1''(x) = p(x)u_1'(x) + q(x)u_1(x) + r(x) \quad x \in (a,b), \qquad u_1(a) = \alpha, \quad u_1'(a) = s = 0.$$

Allora $z_1(x) = \frac{\partial U}{\partial s}(x, 0)$ soddisfa

$$z_1''(x) = p(x)z_1'(x) + q(x)z_1(x) \quad x \in (a,b), \qquad z_1(a) = 0, \quad z_1'(a) = 1.$$

L'iterata di Newton è $s_2 = s_1 - \frac{\varphi(s_1)-\beta}{\varphi'(s_1)} = -\frac{u_1(b)-\beta}{z_1(b)}$. La soluzione corrispondente è $u_2(x) = U(x, s_2)$. Usando la linearità dell'equazione differenziale, si vede che $u_2(x) = u_1(x) + s_2 z_1(x)$. Per definizione u_2 soddisfa l'equazione differenziale desiderata e la condizione al bordo $u_2(a) = \alpha$. Inoltre vale $u_2(b) = u_1(b) - \frac{u_1(b)-\beta}{z_1(b)} z_1(b) = \beta$, cioè anche la seconda condizione al bordo è verificata. Quindi u_2, calcolata con un'unica iterata di Newton, è la soluzione del problema al bordo da cui siamo partiti.

(c) Denotiamo $U(x, s)$ la soluzione di

$$U''(x, s) = \frac{\big(U'(x, s)\big)^2}{U(x, s)}, \quad x \in (0,1), \qquad u(0) = 1, \quad u'(0) = s,$$

cioè $U(x, s) = 1/(1 - sx)$. Sia $\varphi(s) = U(1, s) - \beta = \frac{1}{1-s} - \frac{1}{10}$. Allora $\varphi(s_0) = \varphi(-20) = \frac{1}{21} - \frac{1}{10} \approx -0.05238$ e $\varphi(s_1) = \varphi(0) = 0.1$. Poiché $\varphi(s_0)\varphi(s_1) < 0$ possiamo applicare il metodo di bisezione.

$$s_m = \frac{s_0 + s_1}{2} = -10, \quad \varphi(s_m) = \frac{1}{11} - \frac{1}{10} \approx -0.00909, \quad \varphi(s_m)\varphi(s_1) < 0 \quad \Rightarrow s_0^{(1)} = s_m = -10, \quad s_1^{(1)} = s_1 = 0,$$

$$s_m^{(1)} = \frac{s_0^{(1)} + s_1^{(1)}}{2} = -5, \quad \varphi(s_m) \approx 0.06667, \quad \varphi(s_m^{(1)})\varphi(s_0^{(1)}) < 0 \quad \Rightarrow s_0^{(2)} = s_0^{(1)} = -10, \quad s_1^{(2)} = s_m^{(1)} = -5,$$

$$s_m^{(2)} = \frac{s_0^{(2)} + s_1^{(2)}}{2} = -7.5, \quad \varphi(s_m) \approx 0.017647, \quad \varphi(s_m^{(2)})\varphi(s_0^{(2)}) < 0 \quad \Rightarrow s_0^{(3)} = s_0^{(2)} = -10, \quad s_1^{(3)} = s_m^{(2)} = -7.5.$$

Quindi la soluzione $s = u'(0)$ si trova nell'intervallo $(-10, -7.5)$.
