

METODI NUMERICI DELLE EQUAZIONI ALLE DERIVATE PARZIALI

Dalle lezioni del Prof. Paolo Zunino
per il corso di Ingegneria Matematica

Dispense di Simone Paloschi

Politecnico di Milano
A.A. 2022/2023

Indice

1	Differenze Finite Poisson	3
1.1	Approssimazione di u' e u'' con DF	3
1.2	DF per $-u''=f$	4
1.3	Condizioni al bordo	5
1.4	Proprietà algebriche della matrice	5
1.5	Analisi dell'errore	6
1.5.1	Consistenza	6
1.5.2	Stabilità	6
1.5.3	Convergenza	7
2	DF Calore	9
2.1	Derivazione dello schema	9
2.2	Proprietà della matrice caratteristica	11
2.3	Assoluta stabilità	11
2.4	Consistenza	12
2.5	Analisi dell'errore, Convergenza	13

3	DF trasporto	14
3.1	Introduzione	14
3.2	Forma generale degli schemi in avanti	15
3.3	Consistenza	16
3.4	Condizione CFL e convergenza	16
3.5	Condizione CFL e stabilità	17
3.6	Condizioni al bordo	19
4	FEM	20
4.1	Metodo di Galerkin	20
4.2	Formulazione matriciale del metodo di Galerkin	20
4.3	FEM di grado 1 (lineare)	21
4.4	FEM di grado 2	23
5	Analisi del metodo di Galerkin	26
5.1	Proprietà del metodo di Galerkin	26
5.2	Stabilità del MG	26
5.3	Consistenza del MG	27
5.4	Convergenza del MG	28
5.5	Analisi FEM	29
5.6	Stima dell'errore in L2	30
6	FEM Varie	31
6.1	Condizioni al bordo	31
6.2	FEM 2D	31
7	FEM diffusione e trasporto	33
7.1	Introduzione	33
7.2	Formulazione variazionale	34
7.3	Formulazione matriciale	35
7.4	FEM con diffusione artificiale	36
7.5	FEM stabilizzati	38

1 Differenze Finite Poisson

1.1 Approssimazione di u' e u'' con DF

Approssimiamo la derivata prima con le differenze finite

- DF in avanti: $u'(x_i) \simeq \frac{u(x_i + h) - u(x_i)}{h} \stackrel{def}{=} D^+ u(x_i)$ ordine=1
- DF all'indietro: $u'(x_i) \simeq \frac{u(x_i) - u(x_i - h)}{h} \stackrel{def}{=} D^- u(x_i)$ ordine=1
- DF centrate: $u'(x_i) \simeq \frac{u(x_i + h) - u(x_i - h)}{2h} \stackrel{def}{=} D^c u(x_i)$ ordine=2

Approssimazione derivata seconda con le differenze finite

- Schema a tre punti:

Equivale a uno schema centrato due volte con semipasso $\frac{h}{2}$

$$u''(x_i) \simeq \frac{1}{h^2} \left(u(x_i + h) - 2u(x_i) + u(x_i - h) \right) \stackrel{def}{=} D^2 u(x_i)$$

Vediamo l'analisi dell'errore del metodo a 3 punti come esempio.

In generale l'analisi dell'errore ha sempre gli stessi procedimenti:

- i) Scrivo lo sviluppo di Taylor
- ii) Sostituisco le derivate sviluppate nella formula dell'errore
- iii) Semplifico e per valutare l'ordine, guardo l'esponente più piccolo di h

Vediamo un esempio per lo schema a tre punti.

Sviluppi in serie di Taylor troncati:

$$u(x_i + h) = u(x_i) + hu'(x_i) + \frac{h^2}{2}u''(x_i) + \frac{h^3}{6}u'''(x_i) + \frac{h^4}{24}u''''(z_i^+)$$

$$u(x_i - h) = u(x_i) - hu'(x_i) + \frac{h^2}{2}u''(x_i) - \frac{h^3}{6}u'''(x_i) + \frac{h^4}{24}u''''(z_i^-)$$

Analisi dell'errore:

$$u''(x_i) - D_{xx}^2 u(x_i) = u''(x_i) - \frac{u(x_i + h) - 2u(x_i) + u(x_i - h)}{h^2} = \dots = \frac{h^2}{12}(u''''(z_i^+) + u''''(z_i^-)) = \mathcal{O}(h^2)$$

1.2 DF per $-u''=f$

Problema: data $f(x) : (0, 1) \rightarrow \mathbb{R}$, $f \in C^0(0, 1)$ trovare $u(x) : (0, 1) \rightarrow \mathbb{R}$, $u \in C^2(0, 1)$ tale che

$$-u'' = f, \quad x \in (0, 1); \quad u(0) = u(1) = 0.$$

Oss. Tale equazione non è un'EDP, lo diventa se si considera lo stesso modello in più dimensioni

Discretizzazione del problema con DF

1) Definiamo una griglia nell'intervallo

Siano $x_i = i \cdot h$ per $i = 0 \dots N$ e $h =$ passo di discretizzazione $= \frac{1}{N}$

2) Discretizzazione dell'equazione sulla griglia

$$\begin{array}{ccc} -u''(x_i) = f(x_i) & \longrightarrow & -D^2 u(x_i) = f(x_i) \quad i = 1 \dots N-1 \\ \text{continuo} & & \text{discreto} \end{array}$$

3) Discretizzazione del problema

Soluzione numerica = funzione discreta $U_i \simeq u(x_i)$

Troviamo un problema discreto lineare di dimensione $N-1$ nelle incognite U_i

$$\begin{cases} -\frac{1}{h^2}(U_{i+1} - 2U_i + U_{i-1}) = f(x_i) & i = 1 \dots N-1 \\ U_0 = U_N = 0 & \text{condizioni al bordo} \end{cases}$$

Scriviamo il problema lineare in forma matriciale

$$\frac{1}{h^2} \cdot \begin{bmatrix} 2 & -1 & 0 & \dots & \dots & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & \dots & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & \vdots \\ 0 & \dots & & & -1 & 2 & -1 & 0 \\ 0 & \dots & & & & -1 & 2 & -1 \\ 0 & \dots & & & & & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ u_{N-2} \\ u_{N-1} \end{bmatrix} = \begin{bmatrix} \frac{u_0}{h^2} + f(x_1) \\ f(x_2) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ f(x_{N-2}) \\ \frac{u_N}{h^2} + f(x_{N-1}) \end{bmatrix}$$

Oss. Abbiamo scritto la forma matriciale con le condizioni di Dirichlet non nulle

Il problema alle DF per $-u'' = f$ in forma matriciale è equivalente a trovare \underline{U} tc $A\underline{U} = \underline{F}$

1.3 Condizioni al bordo

- Dirichlet $u(0) = \varphi_0$

Devo semplicemente spostare la condizione al bordo dalla parte dei termini noti

Vediamo come cambia la condizione al nodo x_1 :

$$-\frac{\varphi_0 - 2U_1 + U_2}{h^2} = f(x_1) \rightarrow -\frac{-2U_1 + U_2}{h^2} = f(x_1) + \frac{\varphi_0}{h^2}$$

Otengo il problema matriciale visto sopra

- Neumann $-u'(0) = \varphi_0$

Dovrò discretizzare e approssimare l'equazione anche al nodo x_0

Potrei usare le DF in avanti, però inquinerei un'approssimazione del II ordine con un errore del I ordine

Uso il metodo del **nodo fantasma**:

Introduco il nodo x_{-1} e approssimo con DF centrare in x_0 , poi tratto U_{-1} come variabile ausiliare e la sostituisco nella condizione dell'equazione in x_0

$$\begin{aligned} -u'(x_0) = \varphi_0 &\rightarrow -D^c U(x_0) = \varphi_0 \rightarrow -\frac{1}{2h}(U_1 - U_{-1}) = \varphi_0 \rightarrow U_{-1} = U_1 + 2h\varphi_0 \\ -\frac{U_1 - 2U_0 + U_{-1}}{h^2} = f(x_0) &\rightarrow -\frac{2U_1 - 2U_0 + 2h\varphi_0}{h^2} = f(x_0) \rightarrow -\frac{U_1 - U_0}{h} = \varphi_0 + \frac{1}{2}hf(x_0) \end{aligned}$$

Interpreto questa approssimazione come una correzione di DF in avanti, rendendola del II ordine

Oss. Dovremo aggiungere una riga al sistema lineare

1.4 Proprietà algebriche della matrice

$$A_h \underline{U} = \underline{F}$$

La matrice A_h è tridiagonale e simmetrica definita positiva

\Rightarrow ho $N - 1$ autovalori distinti e positivi, inoltre posso valutare lo spettro di A_h :

$$\lambda_i = \frac{4}{h^2} \sin^2 \left(\frac{\pi}{2} i h \right) \quad i = 1 \dots N - 1$$

Questi autovalori sono ordinati al crescere di i , posso quindi valutare il massimo e il minimo

$$\lambda_{\max} = \frac{4}{h^2} \sin^2 \left(\frac{\pi}{2} \frac{N-1}{N} \right) \xrightarrow{N \rightarrow +\infty} \frac{4}{h^2} \cdot 1 \Rightarrow \lambda_{\max} \leq \frac{4}{h^2}$$

Dato che per x piccolo $\sin(x) \simeq x$, allora $\lambda_{\min} \simeq \pi^2$

Quindi λ_{\min} è indipendente da h ed è discosto da 0 $\Rightarrow A_h$ è sempre invertibile

Numero di condizionamento spettrale $= K_2(A_h) = \frac{\lambda_{\max}}{\lambda_{\min}} \leq \frac{4}{h^2} \cdot \frac{1}{\pi^2} = \mathcal{O}(h^{-2})$

1.5 Analisi dell'errore

DEFINIZIONE.

L'errore è la differenza tra la soluzione esatta e quella numerica

Errore locale: $e(x_i) = u(x_i) - U_i$

Errore globale: $E = \|\underline{e}\|$

Prima di iniziare l'analisi dell'errore, introduciamo qualche notazione:

Problema in forma astratta: $Lu = f$ dove L è un operatore, nel nostro caso $L = -\frac{d^2}{dx^2}$

Lo schema in forma astratta: $L_h u_h = f_h$, nel nostro caso $L_h = D^2$

Soluzione numerica: $\underline{U} \rightarrow u_h$

Termine forzante: $\underline{F} \rightarrow f_h$

1.5.1 Consistenza

La proprietà di consistenza misura se lo schema numerico è consistente con il problema continuo, per calcolare la consistenza, dovremo sostituire la soluzione continua nello schema e valutare il residuo $= L_h u - f_h$

DEFINIZIONE.

Lo schema $L_h u_h = f_h$ si dice asintoticamente **consistente** con $Lu = f$ se $\lim_{h \rightarrow 0} \|L_h u - f_h\| = 0$

Ovvero se il residuo tende a zero per $h \rightarrow 0$

Se uno schema è consistente allora $L_h \xrightarrow{h \rightarrow 0} L \quad f_h \xrightarrow{h \rightarrow 0} f$

Per schemi alle DF il residuo $L_h u - f_h$ si chiama errore di troncamento locale o LTE

Oss. Per lo schema a 3 punti $LTE(x_i) = \mathcal{O}(h^2) \quad \max_x |LTE(x)| \leq \frac{1}{12} h^2 \max_x |u^{iv}(x)|$

E quindi il metodo è consistente del secondo ordine

1.5.2 Stabilità

DEFINIZIONE.

Un metodo numerico è **stabile** se la soluzione numerica dipende con continuità dai dati

Ovvero se $\exists C_s$ indipendente da h t.c. $\|u_h\| \leq C_s \|f_h\|$

Prendiamo uno schema soggetto a perturbazione del dato: $L_h(u_h + \delta u_h) = f_h + \varepsilon$

Facciamo la differenza tra lo schema perturbato e quello non, otteniamo: $L_h \delta u_h = \varepsilon$

Possiamo dire che se lo schema è stabile allora $\|\delta u_h\| \leq C_s \|\varepsilon\|$

Oss. Per lo schema a 3 punti si ricava $\|u_h\|_{h,\infty} \leq \frac{1}{8} \|f_h\|_{h,\infty} \quad \max_i |U_i| \leq \frac{1}{8} \max_i |f(x_i)|$

Lo schema a 3 punti è stabile

1.5.3 Convergenza

DEFINIZIONE.

Lo schema è **convergente** se l'errore globale $\rightarrow 0$ per $h \rightarrow 0$

DEFINIZIONE.

Lo schema si dice di **ordine** p , se esiste $p \in \mathbb{R}^+$ t.c. l'errore globale $E = \mathcal{O}(h^p)$ ($\exists c > 0$ t.c. $E \leq ch^p$)

Per ricavare la convergenza vediamo l'analisi dell'errore per intero

1) Problema continuo $Lu = f$

2) Problema discreto $L_h U_h = f_h$

3) Schema perturbato $L_h u = f_h + L T E_h$

4) = 3) - 2) equazione dell'errore $e = u - u_h \quad L_h e = L T E_h$

Proseguiamo adesso nel caso dello schema a 3 punti

5) Consistenza $\implies L T E_h = \mathcal{O}(h^2)$ ovvero $\|L T E_h\|_\infty \leq \frac{1}{12} h^2 \max_x |u^{iv}(x)|$

6) Stabilità $\implies \|e\|_\infty \leq C_s \|L T E_h\|_\infty$ con $C_s = \frac{1}{8}$

7) Convergenza = consistenza + stabilità = 5) + 6)

$$E = \|e\|_\infty \leq \frac{1}{8} \|L T E_h\|_\infty \leq \frac{1}{96} h^2 \max_x |u^{iv}(x)|$$

Lo schema a 3 punti è convergente di ordine 2

Posso anche scrivere $\|u - u_h\|_{h,\infty} \leq \frac{h^2}{96} \|f''\|_\infty$

TEOREMA: Principio di equivalenza.

Per un metodo consistente, l'ordine di stabilità è uguale all'ordine di convergenza

Oss. Lo useremo principalmente per verificare la convergenza:

$$\text{Consistenza} + \text{stabilità} \iff \text{convergenza}$$

2 DF Calore

2.1 Derivazione dello schema

Vogliamo risolvere l'equazione del calore:

$$\begin{cases} u_t - D\Delta u = f & \text{in } Q_T \\ u(x, 0) = g(x) & \text{in } \overline{Q}_T \text{ condizione iniziale} \\ \text{Condizioni al bordo} & \text{su } \partial Q_T \end{cases}$$

In particolare analizzeremo $x \in (0, 1)$ $t \in (0, T)$, mentre per le condizioni al bordo valgono gli stessi discorsi visti per Poisson, ma considereremo $u(0, t) = u(1, t) = 0$

Siano $h = \Delta x = \frac{1}{N}$ e $\tau = \Delta t = \frac{T}{N}$ gli intervalli spaziali e temporali

La discretizzazione della griglia sarà con $x_i = i \cdot h$ e $t_n = n \cdot \tau$ dove $i, n = 1 \dots N$

Nell'equazione del calore ho derivate nel tempo e nello spazio, quindi procederò a discretizzare prima nello spazio e poi nel tempo

Con lo schema a 3 punti cerco un'approssimazione $U_i(t) \simeq u(t, x_i)$ dove il tempo è una variabile continua, ottengo un problema semi-discreto

$$\begin{cases} d_t U_i(t) - D^2 U_i(t) = f(t, x_i) & i = 1 \dots N-1 \quad t \in (0, T) \\ U_0(t) = U_N(t) = 0 & t \in (0, T) \quad B.C. \\ U_i(t=0) = g(x_i) & i = 1 \dots N-1 \quad I.C. \end{cases}$$

$$\text{Posti } \underline{U}(t) = \begin{bmatrix} U_1(t) \\ \vdots \\ U_i(t) \\ \vdots \\ U_{N-1}(t) \end{bmatrix} \quad \underline{F}(t) = \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_i) \\ \vdots \\ f(x_{N-1}) \end{bmatrix} \quad \text{e } \underline{G} = \begin{bmatrix} g(x_1) \\ \vdots \\ g(x_i) \\ \vdots \\ g(x_{N-1}) \end{bmatrix}$$

Posso riscrivere questo problema in formulazione matriciale:

$$\begin{cases} d_t \underline{U}(t) + A_h \underline{U}(t) = \underline{F}(t) \\ \underline{U}(0) = \underline{G} \end{cases}$$

Ho ottenuto un sistema di EDO, in particolare un problema di Cauchy del I ordine, per discretizzare nel tempo procederò con metodi numerici noti

Definiamo le seguenti approssimazioni: $\underline{U}_n \simeq \underline{U}(t_n)$ e $U_i^n \simeq U_i(t_n) \simeq u(t_n, x_i)$

- Eulero in Avanti $d_t \underline{U}(t_n) \simeq \frac{1}{\tau} (\underline{U}(t_{n+1}) - \underline{U}(t_n))$

Approssimazione dell'equazione del calore con EA al variare di $n = i \dots N$ in forma matriciale e in forma nodale:

$$\begin{aligned} \frac{1}{\tau} (\underline{U}_{n+1} - \underline{U}_n) + A_h \underline{U}_n &= \underline{F}(t_n) \implies \underline{U}^{n+1} = (I - \tau A_h) \underline{U}^n + \tau \underline{F}(t_n) \\ \begin{cases} \frac{1}{\tau} (U_i^{n+1} - U_i^n) - \frac{1}{h^2} (U_{i+1}^n - 2U_i^n + U_{i-1}^n) = f(t^n, x_i) & i = 1 \dots N \\ U_0^n = U_N^n = 0 & U_i^0 = g(x_i) \end{cases} \end{aligned}$$

- Eulero all'Indietro $d_t \underline{U}(t_{n+1}) \simeq \frac{1}{\tau} (\underline{U}(t_{n+1}) - \underline{U}(t_n))$

Approssimazione dell'equazione del calore con EI al variare di $n = i \dots N$ in forma matriciale e in forma nodale:

$$\begin{aligned} \frac{1}{\tau} (\underline{U}_{n+1} - \underline{U}_n) + A_h \underline{U}_{n+1} &= \underline{F}(t_{n+1}) \implies (I + \tau A_h) \underline{U}^{n+1} = \underline{U}^n + \tau \underline{F}(t_{n+1}) \\ \frac{1}{\tau} (U_i^{n+1} - U_i^n) - \frac{1}{h^2} (U_{i+1}^{n+1} - 2U_i^{n+1} + U_{i-1}^{n+1}) &= f(t^{n+1}, x_i) \quad i = 1 \dots N \end{aligned}$$

- Vediamo il metodo di Crank-Nicholson in forma matriciale

$$\begin{aligned} \frac{1}{\tau} (\underline{U}_{n+1} - \underline{U}_n) + \frac{1}{2} A_h \underline{U}_{n+1} + \frac{1}{2} A_h \underline{U}_n &= \frac{1}{2} \underline{F}(t_{n+1}) + \frac{1}{2} \underline{F}(t_n) \\ \implies (I + \frac{1}{2} \tau A_h) \underline{U}^{n+1} &= (I - \frac{1}{2} \tau A_h) \underline{U}^n + \frac{1}{2} \tau \underline{F}(t_{n+1}) + \frac{1}{2} \tau \underline{F}(t_n) \end{aligned}$$

Possiamo generalizzare questi metodi con il θ -metodo

$$\frac{1}{\tau} (\underline{U}_{n+1} - \underline{U}_n) + \theta A_h \underline{U}_{n+1} + (1 - \theta) A_h \underline{U}_n = \theta \underline{F}(t_{n+1}) + (1 - \theta) \underline{F}(t_n)$$

Definiamo un metodo esplicito se non richiede di risolvere un sistema

EA è esplicito, mentre EI e CN sono impliciti

Possiamo rappresentare i metodi visti attraverso la matrice caratteristica C_h^τ

$$\underline{U}^{n+1} = C_h^\tau \underline{U}^n + \tau \tilde{F}(t^*)$$

$$\begin{aligned} EA: \quad C_h^\tau &= (I - \tau A_h) \quad \tilde{F}(t^*) = \underline{F}(t^n) \\ EI: \quad C_h^\tau &= (I + \tau A_h)^{-1} \quad \tilde{F}(t^*) = (I + \tau A_h)^{-1} \underline{F}(t^{n+1}) \\ CN: \quad C_h^\tau &= (I + \frac{1}{2} \tau A_h)^{-1} (I - \frac{1}{2} \tau A_h) \end{aligned}$$

2.2 Proprietà della matrice caratteristica

Una matrice è positiva se tutti i coefficienti a_{ij} sono non negativi e almeno uno è positivo. Queste matrici conservano il segno tra $A * \underline{x}$ e \underline{x}

Quindi se C_h^τ è positiva, allora la soluzione del problema modello conserva il segno della condizione iniziale

$$\underline{U}_0 \geq 0 \implies \underline{U}_n \geq 0 \quad \forall n > 0$$

Questa proprietà è nota come il principio del massimo discreto

Per EA i coefficienti di C_h^τ sono $\frac{\tau}{h^2}$ e $1 - \frac{2\tau}{h^2}$

Quindi C_h^τ è una matrice positiva se e solo se $\tau \leq \frac{h^2}{2}$

Per EI invece sappiamo che $I + \tau A_h$ è una M-matrice, ovvero che ha inversa positiva e quindi soddisfa il principio del massimo discreto $\forall h, \tau$

2.3 Assoluta stabilità

Definiamo un problema modello per l'equazione del calore

$$\begin{cases} \partial_t u - \partial_{xx} u = 0 \\ u(0, t) = u(1, t) = 0 \\ f = 0 \end{cases}$$

DEFINIZIONE.

Un metodo si dice **A-stabile** se quando applicato al problema modello si ha $\lim_{n \rightarrow \infty} \|\underline{U}_n\|_\infty = 0 \quad \forall \underline{U}_0$

Dato che il problema modello ha $f = 0$ possiamo riscrivere i metodi $\underline{U}_{n+1} = (C_h^\tau) \underline{U}_0$

Quindi i metodi sono A-stabili se e solo se C_h^τ è convergente

Sappiamo che una matrice è convergente se e solo se il suo raggio spettrale $S(C_h^\tau) = \max_{1 \leq i \leq N-1} |\lambda_i(C_h^\tau)| < 1$

Inoltre gli autovalori di una matrice $I + \alpha M$ sono $\lambda_i(I + \alpha M) = 1 + \alpha \lambda_i(M)$, quindi possiamo legare gli autovalori di C_h^τ agli autovalori di A_h che conosciamo $\lambda_i(A_h) = \frac{4}{h^2} \sin^2\left(\frac{\pi i h}{2}\right) \quad i = 1 \dots N-1$

- EA ha autovalori $\lambda_i(C_h^\tau) = 1 - \tau\lambda_i(A_h)$

Da questa formula notiamo che l'autovalore con modulo massimo di C_h^τ corrisponde all'autovalore massimo di A_h quindi scriviamo la condizione sul raggio spettrale

$$|\lambda_{max}(C_h^\tau)| = |1 - \frac{4\tau}{h^2} \sin^2\left(\frac{\pi}{2}ih\right)| < |1 - \frac{4\tau}{h^2} \cdot 1| < 1$$

$$\text{EA è A-stabile} \iff \tau < \frac{h^2}{2}$$

Oss. Questa condizione è molto stringente, ma EA rimane comunque un metodo esplicito

- EI ha autovalori $\lambda_i((I + \tau A_h)^{-1}) = \frac{1}{1 + \tau\lambda_i(A_h)}$

$$\left| \frac{1}{1 + \underbrace{\frac{4\tau}{h^2} \sin^2\left(\frac{\pi}{2}ih\right)}_{>0}} \right| < 1$$

Questa condizione è sempre verificata \implies EI è incondizionatamente A-stabile

- Si può verificare che CN è incondizionatamente A-stabile

Relazione tra A-stabilità e stabilità

EA, EI, CN sono stabili se esiste una costante $C_s > 0$ indipendente da h e τ tale che $\|\underline{U}_{n+1}\| \leq C_s \|\underline{U}_0\|$

Applichiamo la formula di questi metodi nel caso del problema modello ($f = 0$), ovvero $\underline{U}_{n+1} = (C_h^\tau)^{n+1} \underline{U}_0$

$$\|\underline{U}_{n+1}\| = \|(C_h^\tau)^{n+1}\| \cdot \|\underline{U}_0\|$$

Stiamo quindi chiedendo che $\|(C_h^\tau)^{n+1}\| \leq C_s$ quindi l'assoluta stabilità, per cui $\|(C_h^\tau)^{n+1}\| \xrightarrow{n \rightarrow \infty} 0$, è una condizione sufficiente per la stabilità

2.4 Consistenza

Sappiamo che un metodo è consistente se $LTE \xrightarrow[h, \tau \rightarrow 0]{} 0$

L'errore di troncamento locale rappresenta il residuo che si genera sostituendo la soluzione del problema continuo nello schema, oppure si può anche interpretare come l'errore intrinseco di un passo temporale

Per valutare la soluzione del problema continuo nei nodi definiamo l'operatore:

$$\underline{Ru}(t) \stackrel{def}{=} \begin{bmatrix} u(t, x_1) \\ \vdots \\ u(t, x_i) \\ \vdots \\ u(t, x_{N-1}) \end{bmatrix}$$

Quindi possiamo scrivere LTE nel modo seguente

$$\tau LTE_n = Ru(t_{n+1}) - C_h^\tau Ru(t_n) - \tau \tilde{F}(t_*)$$

Esempio: valutazione di LTE per EA

$$LTE_i^n = \frac{1}{2} (u(t_{n+1}, x_i) - u(t_n, x_i)) - \frac{1}{h^2} (u(t_n, x_{i+1}) - 2u(t_n, x_i) + u(t_n, x_{i-1})) - f(t_n, x_i)$$

Sfruttando le approssimazioni di Taylor e il fatto che u risolva l'equazione del calore, si ottiene:

$$LTE = \partial_t u(t_n, x_i) + \mathcal{O}(\tau) - \partial_{xx} u(t_n, x_i) + \mathcal{O}(h^2) - f(t_n, x_i) = \mathcal{O}(\tau) + \mathcal{O}(h^2)$$

Si può procedere in modo analogo per EI che avrà $LTE(t_{n+1}, x_i) = \mathcal{O}(\tau) + \mathcal{O}(h^2)$

E per CN che avrà $LTE(t_n, x_i) = \mathcal{O}(\tau^2) + \mathcal{O}(h^2)$ (è del II ordine in spazio e tempo)

2.5 Analisi dell'errore, Convergenza

Sappiamo che l'errore locale è: $u(t_n, x_i) - U_i^n$

Mentre l'errore globale è: $\max_{1 \leq n \leq M} \|\underline{Ru}(t_n) - \underline{U}_n\|$

Un metodo si definisce convergente se l'errore globale $\rightarrow 0$ per $h, \tau \rightarrow 0$

Si dimostra, nel caso $f = 0$, che per gli schemi del tipo $U_{n+1} = C_h^\tau \underline{U}_n + \tau \tilde{F}(t_*)$ le proprietà di consistenza e stabilità implicano la convergenza con lo stesso ordine di LTE, ovvero:

$$\begin{aligned} \|\underline{Ru}(t_{n+1}) - \underline{U}_{n+1}\| &\xrightarrow{h, \tau \rightarrow 0} 0 \quad \forall n : 0 \leq n \leq M-1 \quad \text{e} \quad LTE = \mathcal{O}(\tau^p) + \mathcal{O}(h^q) \\ \implies \|\underline{Ru}(t_{n+1}) - \underline{U}_{n+1}\| &= \mathcal{O}(\tau^p) + \mathcal{O}(h^q) \end{aligned}$$

In particolare si ricava che $\|E(t_{n+1})\| \leq C_s t_{n+1} [\mathcal{O}(\tau^p) + \mathcal{O}(h^q)]$

Summary:

Proprietà	EA	EI	CN
Esplicita / implicita	E	I	I
Stabilità	$\tau < \frac{h^2}{2}$	Incondizionatamente	Incondizionatamente
Ordine	$\mathcal{O}(\tau) + \mathcal{O}(h^2)$	$\mathcal{O}(\tau) + \mathcal{O}(h^2)$	$\mathcal{O}(\tau^2) + \mathcal{O}(h^2)$

3 DF trasporto

3.1 Introduzione

$$\begin{cases} \partial_t u + a \partial_x u = f & \text{in } \mathbb{R} \\ u(0, x) = u_0(x) \end{cases}$$

Definiamo la griglia di calcolo con nodi equispaziati in spazio e tempo $h = \Delta x$ $\tau = \Delta t$

In tempo useremo le approssimazioni:

$$\text{In avanti } \partial_t u(t_n, x_i) \simeq \frac{1}{\tau} (u(t_{n+1}, x_i) - u(t_n, x_i))$$

$$\text{All'indietro } \partial_t u(t_{n+1}, x_i) \simeq \frac{1}{\tau} (u(t_n, x_i) - u(t_{n+1}, x_i))$$

Entrambe del primo ordine $e = \mathcal{O}(\tau)$

$$\text{In spazio: Centrato } a \partial_x u(t, x_i) \simeq \frac{a}{2h} (u(t, x_{i+1}) - u(t, x_{i-1}))$$

$$\text{Decentrato Downwind } a \partial_x u(t, x_i) \simeq \frac{a}{h} (u(t, x_{i+1}) - u(t, x_i))$$

$$\text{Decentrato Upwind } a \partial_x u(t, x_i) \simeq \frac{a}{h} (u(t, x_i) - u(t, x_{i-1}))$$

$$\text{Posti } \underline{U}_n = \begin{bmatrix} U_1^n \\ \vdots \\ U_i^n \\ \vdots \\ U_M^n \end{bmatrix} \quad \text{e} \quad \underline{F}_n = \begin{bmatrix} f(t_n, x_1) \\ \vdots \\ f(t_n, x_i) \\ \vdots \\ f(t_n, x_M) \end{bmatrix}$$

Posso scriviamo gli schemi in formulazione puntuale e matriciale:

- EAC, Eulero in avanti centrato:

$$\frac{1}{\tau} (U_i^{n+1} - U_i^n) + \frac{a}{2h} (U_{i+1}^n - U_{i-1}^n) = f(t_n, x_i)$$

$$\underline{U}_{n+1} = (\mathbb{1} - \frac{1}{2} a \lambda T) \underline{U}_n + \tau \underline{F}_n$$

- EIC, Eulero all'indietro centrato:

$$\frac{1}{\tau} (U_i^{n+1} - U_i^n) + \frac{a}{2h} (U_{i+1}^{n+1} - U_{i-1}^{n+1}) = f(t_{n+1}, x_i)$$

$$\underline{U}_{n+1} = (\mathbb{1} + \frac{1}{2} a \lambda T)^{-1} \underline{U}_n + \tau \underline{F}_{n+1}$$

- Upwind con $a > 0$, Eulero in avanti decentrato:

$$\frac{1}{\tau} (U_i^{n+1} - U_i^n) + \frac{a}{h} (U_i^n - U_{i-1}^n) = f(t_n, x_i)$$

$$\underline{U}_{n+1} = (\mathbb{1} - a\lambda T^-)\underline{U}_n + \tau F_n$$

- Upwind con $a < 0$, Eulero in avanti decentrato:

$$\frac{1}{\tau}(U_i^{n+1} - U_i^n) + \frac{a}{h}(U_{i+1}^n - U_i^n) = f(t_n, x_i)$$

$$\underline{U}_{n+1} = (\mathbb{1} - a\lambda T^+)\underline{U}_n + \tau F_n$$

$$\text{Dove } \lambda = \frac{\tau}{h} \quad T = \begin{bmatrix} 0 & +1 & \\ -1 & \ddots & \ddots \\ & \ddots & \ddots \end{bmatrix} \quad T^- = \begin{bmatrix} +1 & 0 & \\ -1 & \ddots & \ddots \\ & \ddots & \ddots \end{bmatrix} \quad T^+ = \begin{bmatrix} -1 & +1 & \\ 0 & \ddots & \ddots \\ & \ddots & \ddots \end{bmatrix}$$

Tutti questi schemi possono essere rappresentati nella forma $\underline{U}_{n+1} = C_h^r \underline{U}_n + \tau \tilde{F}(t_*)$

3.2 Forma generale degli schemi in avanti

Iniziamo ricavando una formula unica per il metodo Upwind

$$(a > 0) \quad \frac{1}{h}(U_i^n - U_{i-1}^n) = D_x^- U_i^n = D_x^C U_i^n - \frac{h}{2} D_{xx}^2 U_i^n$$

$$(a < 0) \quad \frac{1}{h}(U_{i+1}^n - U_i^n) = D_x^+ U_i^n = D_x^C U_i^n + \frac{h}{2} D_{xx}^2 U_i^n$$

Applichiamo queste formule al metodo upwind

$$\frac{1}{\tau}(U_i^{n+1} - U_i^n) + \begin{cases} a D_x^- U_i^n & (a > 0) \\ a D_x^+ U_i^n & (a < 0) \end{cases} = f(t_n, x_i) \implies \frac{1}{\tau}(U_i^{n+1} - U_i^n) + a D_x^C U_i^n \pm a \frac{h}{2} D_{xx}^2 U_i^n = f(t_n, x_i)$$

Forma generale del metodo upwind:

$$\frac{1}{\tau}(U_i^{n+1} - U_i^n) + a \frac{1}{2h}(U_{i+1}^n - U_{i-1}^n) - |a| \frac{h}{2} \frac{1}{h^2}(U_{i+1}^n - 2U_i^n + U_{i-1}^n) = f(t_n, x_i)$$

Può essere visto come EAC con una correzione che "stabilizza"

Partendo da questa forma otteniamo quella generale degli schemi espliciti:

$$U_i^{n+1} - U_i^n + \frac{1}{2} a \frac{\tau}{h}(U_{i+1}^n - U_{i-1}^n) - \frac{1}{2} K \frac{1}{h^2}(U_{i+1}^n - 2U_i^n + U_{i-1}^n) = \tau f(t_n, x_i)$$

EAC: $K = 0$ Upwind $K = |a|\tau h$ Lax-Friedrichs $K = h^2$ Lax-Wendroff $K = a^2 \tau^2$

3.3 Consistenza

Per studiare la consistenza facciamo l'analisi dell'errore di troncamento locale

In generale si riconoscono e si sostituiscono le approssimazioni delle derivate in spazio e tempo, successivamente si sfrutta il fatto che la soluzione $u(t, x)$ risolva l'equazione del trasporto

Vediamo l'esempio della forma generale di uno schema esplicito:

$$\tau LTE(t_n, x_i) = \underbrace{u(t_{n+1}, x_i) - u(t_n, x_i)}_{(1)} + \underbrace{\frac{1}{2}a\frac{\tau}{h}(u(t_n, x_{i+1}) - u(t_n, x_{i-1}))}_{(2)} - \underbrace{\frac{1}{2}K\frac{1}{h^2}(u(t_n, x_{i+1}) - 2u(t_n, x_i) + u(t_n, x_{i-1}))}_{(3)}$$

$$(1) = \tau [\partial_t u(t_n, x_i) + \mathcal{O}(\tau)]$$

$$(2) = a\tau [\partial_x u(t_n, x_i) + \mathcal{O}(h^2)]$$

$$(3) = -\frac{1}{2}K [\partial_{xx} u(t_n, x_i) + \mathcal{O}(h^2)]$$

$$\tau LTE(t_n, x_i) = \tau \underbrace{[\partial_t u(t_n, x_i) + a\partial_x u(t_n, x_i) - f(t_n, x_i)]}_0 - \frac{1}{2}K\partial_{xx} u(t_n, x_i) + \tau\mathcal{O}(\tau) + \tau\mathcal{O}(h^2)$$

$$LTE(t_n, x_i) = -\frac{1}{2}\frac{K}{\tau}\partial_{xx} u(t_n, x_i) + \mathcal{O}(\tau) + \mathcal{O}(h^2)$$

Sostituendo K nella formula otteniamo

$$\text{EAC: } \mathcal{O}(\tau) + \mathcal{O}(h^2) \quad \text{Upwind: } \mathcal{O}(\tau) + \mathcal{O}(h) \quad \text{LF: } \mathcal{O}(\tau) + \mathcal{O}(h)$$

Osservazioni:

- L'approssimazione in spazio ha $\mathcal{O}(h^2)$ quando $\Delta x = 2h$, ma con $\Delta x = h$ (come per Upwind) ha $\mathcal{O}(h)$
- In generale nel limite $\tau, h \rightarrow 0$, il rapporto λ è costante

- Questa analisi non funziona per LW, dove non dobbiamo sostituire la derivata in tempo

Quindi trasformeremo le derivate spaziali in temporali sfruttando l'equazione di trasporto: $\partial_t u = -a\partial_x u$

E derivando in tempo: $\partial_{tt} u = -a\partial_t \partial_x u = -a\partial_x \partial_t u = a^2 \partial_{xx} u$

Infine uso lo sviluppo di Taylor: $u(t_{n+1}, x_i) = u(t_n, x_i) + \tau \partial_t u(t_n, x_i) + \frac{1}{2}\tau^2 \partial_{tt}^2 u(t_n, x_i) + \mathcal{O}(\tau^3)$

Ottengo: $LTE(t_n, x_i) = \mathcal{O}(\tau^2) + \mathcal{O}(h^2)$

3.4 Condizione CFL e convergenza

La condizione CFL richiede che la velocità, ovvero la pendenza della linea caratteristica, dello schema sia maggiore o uguale di quella del modello

$$\frac{h}{\tau} \geq |a| \iff \tau \leq \frac{h}{|a|} \iff \lambda \cdot |a| \leq 1$$

Possiamo riformulare la condizione CFL, richiede che il dominio di dipendenza della soluzione numerica contenga quella del modello

Il dominio di dipendenza è l'insieme dei punti dello stato iniziale che influenzano la soluzione

Per l'equazione del trasporto c'è solo un punto che si trova risalendo la linea caratteristica del modello

Il dominio di dipendenza della soluzione $u(x, t)$ per il punto (x_i, t^{n+1}) è $D = x^* = x_i - at^{n+1}$

Invece il dominio di dipendenza della soluzione numerica $u_i^{n+1}(x, t)$ per il punto (x_i, t^{n+1}) è:

$$D_i^{n+1} = \{x \in \mathbb{R} : |x - x_i| \leq h(n+1)\}$$

$$CFL \iff D(t_{n+1}, x_i) \subseteq D_i^{n+1}$$

La condizione CFL è necessaria per la convergenza

Quindi no CFL \implies no convergenza

Per dimostrarlo, poniamo un dominio di dipendenza del modello non incluso nel dominio dello schema

Studiamo una perturbazione dello stato iniziale $\hat{u}_0(x) = u_0(x) + \varepsilon(x)$ tale che il supporto della perturbazione non si intersechi il dominio dello schema

A questo punto ho una perturbazione che si propaga nel tempo invariata $\implies \hat{u}(t_{n+1}, x_i) = u(t_{n+1}, x_i) + \varepsilon(x)$

Inoltre i nodi del dominio dello schema non vedono la perturbazione $\implies \hat{U}_i^{n+1} = U_i^{n+1}$

Qualunque sia il comportamento asintotico dello schema, l'errore del problema perturbato, rimane $\hat{e}(t_{n+1}, x_i) = e(t_{n+1}, x_i) + \underbrace{\varepsilon(x_i - at_{n+1})}_{\neq 0}$ che non converge a zero per $h, \tau \rightarrow 0$

In conclusione se CFL non è soddisfatta, lo schema numerico non converge

3.5 Condizione CFL e stabilità

Uno schema $\underline{U}_{n+1} = C_h^T \underline{U}_n$ è stabile se esiste $C_s > 0$ indipendente da h, τ tale che $\|\underline{U}_{n+1}\| \leq C_s \|\underline{U}_0\|$

Inoltre se $C_s \leq 1$ allora lo schema si dice **fortemente stabile**

Ripasso delle possibili norme:

$$\|\underline{U}\|_{h,1} = h \sum_i |U_i| \quad \|\underline{U}\|_{h,2} = \left(h \sum_i U_i^2 \right)^{\frac{1}{2}} \quad \|\underline{U}\|_{h,p} = \left(h \sum_i U_i^p \right)^{\frac{1}{p}}$$

Se $p \rightarrow \infty$ si ottiene la norma del massimo:

$$\|\underline{U}\|_{h,\infty} = \max_i |U_i|$$

Studiamo la relazione tra CFL e stabilità

Proprietà 1:

Se la condizione CFL è soddisfatta, allora gli schemi UW, LF e LW sono fortemente stabili nella norma 1

Proprietà 2:

Lo schema EIC è fortemente stabile nella norma 2, senza alcuna condizione sul passo τ

Proprietà 3:

Lo schema EAC non è mai fortemente stabile. Tuttavia è stabile se $\tau \leq \left(\frac{h}{a}\right)^2$

Vediamo la dimostrazione nel caso UW con $a > 0$

UW: $U_i^{n+1} = (1 - a\lambda)U_i^n + a\lambda U_{i-1}^n$ a cui ci applichiamo la norma

$$h \sum_i |U_i^{n+1}| = h \sum_i |(1 - a\lambda)U_i^n + a\lambda U_{i-1}^n| \leq h \sum_i |(1 - a\lambda)U_i^n| + h \sum_i |a\lambda U_{i-1}^n|$$

A questo punto CFL $\iff |a\lambda| \leq 1 \implies 1 - a\lambda \geq 0$

Quindi $|(1 - a\lambda)U_i^n| = (1 - a\lambda)|U_i^n|$ e $|a\lambda U_{i-1}^n| = a\lambda|U_{i-1}^n|$

Otteniamo $\|\underline{U}_{n+1}\|_{h,1} \leq h(1 - a\lambda) \sum_i |U_i^n| + ha\lambda \sum_i |U_{i-1}^n|$

Inoltre dato che lavoriamo su tutto \mathbb{R} si ha che $\sum_{i=-\infty}^{+\infty} \alpha_i = \sum_{i=-\infty}^{+\infty} \alpha_{i-1}$

$$\|\underline{U}_{n+1}\|_{h,1} \leq (1 - a\lambda + a\lambda)\|\underline{U}_n\|_{h,1} = \|\underline{U}_n\|_{h,1}$$

Per dimostrare la proprietà 2 si usa il metodo dell'energia, ovvero si moltiplica per la soluzione numerica

EIC: $U_i^{n+1} - U_i^n + \frac{1}{2}a\lambda(U_{i+1}^{n+1} - U_{i-1}^{n+1}) = 0$

Moltiplico per U_i^{n+1} e applico la norma

$$\underbrace{h \sum_i (U_i^{n+1} - U_i^n) U_i^{n+1}}_{(I)} + \underbrace{\frac{1}{2}a\lambda h \sum_i (U_{i+1}^{n+1} - U_{i-1}^{n+1}) U_i^{n+1}}_{(II)} = 0$$

Sulla (I) si applica l'identità algebrica $(B - A)B = \frac{1}{2}(B^2 - A^2 + (B - A)^2)$

$$\begin{aligned} h \sum_i (U_i^{n+1} - U_i^n) U_i^{n+1} &= \frac{h}{2} \sum_i ((U_i^{n+1})^2 - (U_i^n)^2 + (U_i^{n+1} - U_i^n)^2) \geq \\ &\geq \frac{h}{2} \sum_i ((U_i^{n+1})^2 - (U_i^n)^2) = \frac{1}{2} \|U_{n+1}\|_{h,2}^2 - \frac{1}{2} \|U_n\|_{h,2}^2 \end{aligned}$$

Invece riconosco (II) come una serie telescopica che quindi è nulla

Dalla formula iniziale otteniamo $\frac{1}{2} \|U_{n+1}\|_{h,2}^2 - \frac{1}{2} \|U_n\|_{h,2}^2 \leq 0 \implies \|U_{n+1}\|_{h,2} \leq \|U_n\|_{h,2}$ e quindi EIC è incondizionatamente fortemente stabile

La proprietà 3 si dimostra anch'essa con il metodo dell'energia, però è complicato e non lo vediamo

Oss. EAC è condizionatamente stabile è una richiesta più stringente di CFL $\tau \leq \frac{h^2}{a^2} \leq \frac{h}{a}$

3.6 Condizioni al bordo

In base alla direzione della velocità del modello \underline{a} si definiscono i bordi di inflow e outflow

$$\text{Inflow: } \{x \in \partial\Omega : \underline{a} \cdot \underline{n} < 0\}$$

Le condizioni al bordo per l'equazione di trasporto devono essere imposte solo su un bordo, se ($a > 0$) sulla frontiera inflow

Mentre i valori outflow sono determinati dall'equazione stessa

Il problema di trasporto con condizione al bordo e $a > 0$ è:

$$\begin{cases} \partial_t u + a \partial_x u = f & x \in (0, 1), t > 0 \\ u(t = 0, x) = u_0(x) & t = 0, x \in (0, 1) \quad (\text{condizioni iniziali}) \\ u(t, x = 0) = \Phi & x = 0, t > 0 \quad (\text{Inflow}) \end{cases}$$

Per discretizzare la condizione al bordo inflow dovrò semplicemente sostituire $U_0^n = \Phi$

Invece ho problemi a discretizzare al bordo outflow perché mi servirebbe U_{M+1}^n che non è definito, dovrò quindi approssimare questo valore con formule di estrapolazione

Lineare: $f(x+h) = 2f(x) - f(x-h) + \mathcal{O}(h^2)$ ovvero $U_{M+1}^n = 2U_M^n - U_{M-1}^n$

Però questo mi dà un'approssimazione solo del I ordine, per evitarlo posso usare quella quadratica:

$f(x+h) = 3f(x) - 3f(x-h) + f(x-2h) + \mathcal{O}(h^3)$ ovvero $U_{M+1}^n = 3U_M^n - 3U_{M-1}^n + U_{M-2}^n$

4 FEM

Metodi per la risoluzione delle EDP e proprietà fondamentali

- Differenze finite: - Formulazione classica
 - Discretizzazione regolare su nodi x_{ij}
- Elementi finiti: - Formulazione variazionale
 - Griglie generali, partizione di Ω in elementi
 - La soluzione è una funzione : $\Omega \rightarrow \mathbb{R}$

4.1 Metodo di Galerkin

Ripassiamo il PVA, problema variazionale astratto:

Dati V uno spazio di Hilbert, $F : V' \rightarrow \mathbb{R}$ funzionale lineare continuo e $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$ una forma bilineare continua e coerciva, trovare $u \in V$ t.c. $a(u, v) = F(v) \forall v \in V$

Il metodo di Galerkin, o MG, determina una soluzione approssimata e discreta del PVA attraverso la proiezione di u su un sottospazio a dimensione finita di V

Per ricavare la formulazione del MG supponiamo che $a(\cdot, \cdot)$ sia equivalente ad un prodotto scalare, poi varrà anche se quest'ipotesi non è rispettata.

Usiamo $a(\cdot, \cdot)$ per fare la proiezione sul sottospazio V_k

Avremmo $a(u_k, v_k) = a(u, v_k) \forall v_k \in V_k$ però per il PVA vale $a(u, v) = F(v) \forall v \in V$ e $V_k \subset V$, allora possiamo formulare il metodo di Galerkin nel seguente modo:

Dato $V_k \subset V$ un sottospazio con $\dim(V_k) = k < +\infty$. Trovare $u_k \in V_k$ funzione $u_k : \Omega \rightarrow \mathbb{R}$ t.c.

$$a(u_k, v_k) = F(v_k) \quad \forall v_k \in V_k$$

4.2 Formulazione matriciale del metodo di Galerkin

Posta $k = \dim(V_k)$ la dimensione di V_k

Le equazioni del metodo sono: $a(u_k, v_k) = F(v_k) \forall v_k \in V_k$

Ovvero ci sono tante equazioni quante sono le funzioni linearmente indipendenti di V_k , queste funzioni sono una base $\{\varphi_i(x)\}_{i=1}^k$ di V_k e sono $k = \dim(V_k)$

Quindi le equazioni sono equivalenti a: $a(u_k, \varphi_i) = F(\varphi_i) \forall i = 1 \dots k$

L'incognita del metodo è la funzione $u_k \in V_k$

Se conosco una base di V_k , allora $u_k(x) = \sum_{i=1}^k U_i \varphi_i(x)$

Quindi le incognite sono i coefficienti $\{U_i\}_{i=1}^k$

Riscriviamo le equazioni con le nuove incognite:

$$a\left(\sum_{j=1}^k U_j \varphi_j(x), \varphi_i(x)\right) = F(\varphi_i(x)) \quad \forall i = 1 \dots k$$

Dato che $a(\cdot, \cdot)$ è bilineare e U_j non dipendono da x

$$\sum_{j=1}^k U_j a(\varphi_j(x), \varphi_i(x)) = F(\varphi_i(x)) \quad \forall i = 1 \dots k$$

Il metodo di Galerkin è equivalente a un sistema lineare di k equazioni in k incognite: $A\underline{U} = \underline{F}$

Dove $A_{ij} = a(\varphi_j(x), \varphi_i(x))$ $(\underline{U})_i = U_i$ $(\underline{F})_i = F(\varphi_i(x))$ per $i = 1 \dots k$

Il metodo degli elementi finiti è un caso particolare del MG in cui il sottospazio V_k è definito come spazio di funzioni continue e polinomiali a tratti su una partizione del dominio Ω in "elementi"

Per i metodi FEM partendo dal parametro di discretizzazione h si userà il sottospazio discreto V_h con $\dim(V_h) = N_h$ dove N_h è proporzionale al numero di elementi della partizione

4.3 FEM di grado 1 (lineare)

Supponiamo $\Omega = (0, 1)$, la partizione sarà in sottointervalli (x_i, x_{i+1})

I punti $x_i = ih$ sono detti vertici e hanno passo di discretizzazione $h = \frac{1}{N}$

Lo spazio per le FEM di ordine 1 è:

$$V_h = X_h^1(0, 1) = \left\{ v_h \in C^0(0, 1) : v_h(x) \Big|_{x_i}^{x_{i+1}} \in \mathbb{P}^1(x_i, x_{i+1}) \quad i = 0 \dots N-1 \right\}$$

ovvero la restrizione $\Big|_{x_i}^{x_{i+1}}$ di v_h è un polinomio di grado 1

Questa definizione si può estendere ai gradi $r \geq 1$

Definiamo la base $\{\varphi_i(x)\}_{i=0}^N$ dello spazio FEM lineare

Gli elementi della base $\varphi_i(x) \in X_h^1(0, 1)$ valgono 1 nel nodo x_i e zero in tutti gli altri nodi, quindi

- $\varphi_i(x)$ hanno supporto compatto (x_{i-1}, x_{i+1}) , in questo modo sarà più facile calcolare $a(\varphi_i, \varphi_j)$

$$\bullet \varphi_i(x_j) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} = \delta_{ij} \text{ e quindi è una base lagrangiana}$$

Vale la seguente proprietà: $u_h(x_i) = U_i$

DIM.

$$u_h(x_i) = \sum_{j=0}^N U_j \varphi_j(x_i) = \sum_{j=0}^N U_j \delta_{ji} = U_i$$

□

Ovvero risolvendo il sistema lineare ottengo i valori nodali della funzione elementi finiti

$$\text{Conseguenza: } u_h(x) = \sum_{j=0}^N u_h(x_j) \varphi_j(x)$$

Elemento finito

Un elemento finito è definito attraverso 3 ingredienti:

- l'elemento geometrico K : in particolare per $1D$ avremo un intervallo, per $2D$ un triangolo/quadrilatero, $3D$ un tetraedro.
- lo spazio polinomiale su K .
- una base langragiana definita come

$$v_h(x) = \sum_j v_j \phi_j(x)$$

con $\phi(x_j) = \delta_{ij}$ e gradi di libertà $v_j = v_h(x_j)$

Cambiando uno di questi ingredienti cambia il FEM, per questo ci sono tantissimi FEM

Assemblaggio matrice FEM 1D

Formulazione variazionale di Poisson:

$$a(u, v) = (\nabla u, \nabla v) = \int_{\Omega} \nabla u(x) \cdot \nabla v(x) dx \quad F(v) = (f, v) = \int_{\Omega} f v dx$$

Formulazione matriciale:

$$A \underline{u} = \underline{f} \quad A_{ij} = a(\varphi_j, \varphi_i) \quad f_i = \int_{\Omega} f \varphi_i dx$$

Notiamo subito che A è tridiagonale per FEM \mathbb{P}^1 in $1D$

Infatti posta la base come vista sopra, il supporto di φ_i ha intersezione non nulla solo con φ_{i-1} e φ_{i+1}

Vogliamo automatizzare il calcolo dei coefficienti della matrice e del termine noto di moto da poter riscrivere il modello in un algoritmo

$$\begin{cases} A_{i,i-1} = \int_{x_{i-1}}^{x_i} \varphi'_{i-1} \varphi'_i \\ A_{i,i} = \int_{x_{i-1}}^{x_i} \varphi'_i \varphi'_i + \int_{x_i}^{x_{i+1}} \varphi'_i \varphi'_i \\ A_{i,i+1} = \int_{x_i}^{x_{i+1}} \varphi'_{i+1} \varphi'_i \end{cases}$$

L'idea è quella di costruire un elemento di riferimento e trasformare le basi su di esso

L'elemento di riferimento è l'intervallo $(0, 1)$.

Definiamo: La mappa $x = \Phi(\xi) = x_i + \xi(x_{i+1} - x_i)$ $\Phi : \xi \in (0, 1) \mapsto x \in (x_i, x_{i+1})$

E la mappa inversa $\xi = \Phi^{-1}(x) = \frac{x - x_i}{x_{i+1} - x_i}$ $\Phi^{-1} : x \mapsto \xi$

Definiamo le basi sull'intervallo di riferimento $\hat{\varphi}_0 = 1 - \xi$ $\hat{\varphi}_1 = \xi$

Elenchiamo dei risultati che saranno utili per calcolare gli integrali:

$$\begin{aligned} \varphi_i(x) \Big|_{x_i}^{x_{i+1}} &= \hat{\varphi}_0(\Phi^{-1}(x)) & \varphi_{i+1}(x) \Big|_{x_i}^{x_{i+1}} &= \hat{\varphi}_1(\Phi^{-1}(x)) \\ dx &= h d\xi & d\xi &= h^{-1} dx & \text{dove } h &= x_{i+1} - x_i \\ \frac{d\hat{\varphi}_0}{d\xi} &= -1 & \frac{d\hat{\varphi}_1}{d\xi} &= 1 \end{aligned}$$

Calcoliamo un elemento della matrice, gli altri due sono analoghi

$$\begin{aligned} A_{i,i+1} &= \int_{x_i}^{x_{i+1}} \varphi'_{i+1} \varphi'_i dx = \int_{x_i}^{x_{i+1}} \frac{d\varphi_{i+1}}{dx} \cdot \frac{d\varphi_i}{dx} dx = \int_{x_i}^{x_{i+1}} \frac{d\hat{\varphi}_1(\Phi^{-1}(x))}{dx} \frac{\hat{\varphi}_0(\Phi^{-1}(x))}{dx} dx = \\ &= \int_0^1 \frac{d\hat{\varphi}_1(\xi)}{d\xi} \frac{d\xi}{dx} \cdot \frac{d\hat{\varphi}_0(\xi)}{d\xi} \frac{d\xi}{dx} h d\xi = \int_0^1 -\frac{1}{h} d\xi = -\frac{1}{h} \end{aligned}$$

Si ottiene che $A = \frac{1}{h} \text{tridiag}[-1, 2, -1]$

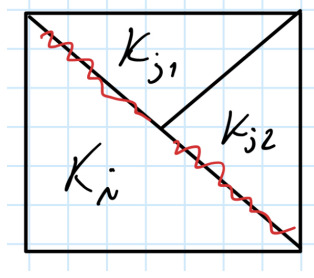
Invece per F, sappiamo che $f = \text{cost} \implies \int_{x_{i-1}}^{x_{i+1}} f \varphi_i(x) dx = hf \implies F = hf[1, 1...]$

4.4 FEM di grado 2

Dati due elementi diversi della partizione 2D, si richiede che:

$$K_i \neq K_j \implies K_i \cap K_j = \begin{cases} \emptyset & \text{insieme vuoto} \\ \text{un vertice} \\ \text{un intero lato di } K_i \text{ e } K_j \end{cases}$$

Controesempio non ammissibile perché K_i condivide lati non interi con altri elementi:



Definizione del parametro h :

Per ogni triangolo K definiamo $diam(K)$ come il diametro del cerchio circoscritto

Allora $h = \max_{K \in T_h} diam(K)$ dove T_h indica l'insieme degli elementi

Inoltre posto $\rho(K)$ il diametro del cerchio inscritto

La griglia si dice quasi-uniforme se $\exists c > 0$ indipendente dai triangoli t.c. $c < \frac{\rho(K)}{diam(K)} < 1 \quad \forall K \in T_h$

Questo garantisce che non ci siano triangoli estremamente difforni

Assemblaggio matrice FEM 2D

Per assemblare la matrice, bisogna:

1. scorrere tutti gli elementi
2. per ciascuno di essi calcolare la matrice locale, per cui sarà necessario calcolare B_K
3. mappare i coefficienti di A_{loc} in A usando la trasformazione da elemento corrente a elemento di riferimento

Prendiamo l'elemento di riferimento come il triangolo unitario con coordinate di riferimento ξ, η

Poniamo F_K la trasformazione dal riferimento al corrente, univocamente determinata dai vertici di K

Infine sia B_K la matrice jacobiana di questa trasformazione F_K

Dobbiamo calcolarci i coefficienti locali $A_{ij}^K = \int_K \nabla \varphi_j \cdot \nabla \varphi_i dx$

Sappiamo che $dx = \det(B_K) d\hat{x}$

Inoltre per trovare i gradienti "generalizziamo" il caso 1D: $\frac{d\varphi_i}{dx} = \frac{d\hat{\varphi}_i}{d\xi} \frac{d\xi}{dx}$ ovvero:

gradiente corrente = gradiente di riferimento \cdot inversa matrice jacobiana

$$\Rightarrow \nabla_{x,y} \varphi_i(x, y) = B_K^{-T} \cdot \hat{\nabla}_{\xi,\eta} \hat{\varphi}_i(\xi, \eta)$$

$$\begin{aligned}
A_{i,j}^K &= \int_K \nabla \varphi_j \cdot \nabla \varphi_i \, dx = \int_K \nabla \widehat{\varphi}_j(F_K^{-1}(x, y)) \cdot \nabla \widehat{\varphi}_i(F_K^{-1}(x, y)) \, dx \, dy = \\
&= \int_{\widehat{K}} \left(B_K^{-T} \widehat{\nabla}_{\xi, \eta} \widehat{\varphi}_j(\xi, \eta) \right) \cdot \left(B_K^{-T} \widehat{\nabla}_{\xi, \eta} \widehat{\varphi}_i(\xi, \eta) \right) \det(B_K) \, d\xi \, d\eta
\end{aligned}$$

Nel caso degli elementi finiti lineari B_K^{-T} e $\widehat{\nabla}_{\xi, \eta} \widehat{\varphi}_i$ sono costanti, quindi posso portare tutto fuori dall'integrale e mi rimane l'integrale di \widehat{K} che, essendo triangolo unitario, so essere $|\widehat{K}| = \frac{1}{2}$, ottengo:

$$A_{ij}^K = \frac{\det(B_K)}{2} \left(B_K^{-T} \widehat{\nabla}_{\xi, \eta} \widehat{\varphi}_j(\xi, \eta) \right) \cdot \left(B_K^{-T} \widehat{\nabla}_{\xi, \eta} \widehat{\varphi}_i(\xi, \eta) \right)$$

Osservo che la dipendenza da K si manifesta solo attraverso B_K

5 Analisi del metodo di Galerkin

5.1 Proprietà del metodo di Galerkin

Il MG è un caso particolare del PVA, di conseguenza:

V_h è uno spazio di Hilbert perché $V_h \subset V$

$a(\cdot, \cdot)$ è bilineare, continua e coerciva

$F(\cdot)$ è un funzionale lineare e continuo

Quindi per MG: trovare $u_h \in V_h$ tc $a(u_h, v_h) = F(v_h) \quad \forall v_h \in V_h$, vale il teorema di Lax-Milgram.

\implies Il problema è ben posto, $\exists!$ soluzione e vale la stima $\|u_h\|_V \leq \frac{1}{\alpha} \|F\|_{V'}$

Vediamo meglio queste proprietà:

TEOREMA.

La matrice A associata alla discretizzazione del problema ellittico $A\underline{u} = \underline{f}$ con forma bilineare $a(\cdot, \cdot)$ coerciva è definita positiva

DIM.

$$\begin{aligned} \forall \underline{V} \in \mathbb{R}^{N_h} \quad \underline{V}^T A \underline{V} &= \sum_{j=1}^{N_h} \sum_{i=1}^{N_h} V_j A_{ji} V_i = \sum_{i,j=1}^{N_h} V_j a(\varphi_j(x), \varphi_i(x)) V_i = \{a \text{ è bilineare}\} = \\ &= a\left(\sum_{j=1}^{N_h} V_j \varphi_j(x), \sum_{i=1}^{N_h} V_i \varphi_i(x)\right) = \left\{ \begin{array}{l} v(x) = \sum_{i=1}^{N_h} V_i \varphi_i(x) \\ v(x) \in V_h \end{array} \right\} = a(v(x), v(x)) \stackrel{\text{coerciva}}{\geq} \alpha \|v(x)\|_V^2 \geq 0 \end{aligned}$$

Ma essendo una norma vale anche che $\underline{V} \in \mathbb{R}^{N_h} \quad \underline{V}^T A \underline{V} = 0 \Leftrightarrow v(x) = 0 \Leftrightarrow \underline{V} = 0$

□

Proprietà: La matrice A è simmetrica se e solo se la forma bilineare $a(\cdot, \cdot)$ è simmetrica

Corollario: La soluzione del problema di Galerkin esiste ed è unica

5.2 Stabilità del MG

Corollario di Lax-Milgram: Il MG è stabile, uniformemente rispetto ad h , in quanto vale la maggiorazione

$\|u_h\|_V \leq \frac{1}{\alpha} \|F\|_{V'}$ dove α è la costante di coercività ed è indipendente da h

DIM.

Usiamo $v_h = u_h$ come funzione test nel MG

$$\alpha \|u_h\|_V^2 \stackrel{a \text{ coerciva}}{\leq} a(u_h, u_h) = F(u_h) \stackrel{F \text{ limitata}}{\leq} \|F\|_{V'} \cdot \|u_h\|_V$$

E poi divido per $\alpha \|u_h\|_V$

□

5.3 Consistenza del MG

Per analizzare la consistenza valutiamo il residuo che si ottiene sostituendo u (soluzione del PVA) nel MG

Sappiamo che, in generale, $u \in V$ ma $u \notin V_h$

Per definizione del PVA $a(u, v) = F(v) \forall v \in V$, ma $v_h \in V_h \subset V \implies a(u, v_h) = F(v_h)$

Di conseguenza sostituendo u nel MG: $a(u_h, v_h) = F(v_h)$ si ottiene che il residuo è nullo $a(u, v_h) - F(v_h) = 0$

Si dice che il MG è **fortemente consistente**, ovvero è consistente indipendentemente dal numero di basi e quindi dalla dimensione di V_h

Metodo di Galerkin come metodo di Proiezione:

Facendo la differenza tra il residuo $a(u, v_h) = F(v_h)$ e il MG $a(u_h, v_h) = F(v_h)$ si ottiene

$$a(u - u_h, v_h) = 0 \quad \forall v_h \in V_h \quad \text{dove } u - u_h \text{ è l'errore}$$

Questa proprietà è detta **ortogonalità di Galerkin** ed è equivalente alla consistenza

Quindi se la forma bilineare $a(\cdot, \cdot)$ è un prodotto scalare allora $(u - u_h) \perp v_h \in V_h$

Si può quindi vedere la soluzione di MG come la proiezione rispetto ad $a(\cdot, \cdot)$ di u su V_h

Ovvero u_h è la funzione di V_h più vicina a u nella distanza definita da $a(\cdot, \cdot)$

$$\|u - u_h\|_a = \inf_{v_h \in V_h} \|u - v_h\|_a$$

Possiamo concludere che sotto le ipotesi: $a(\cdot, \cdot)$ prodotto scalare e $\|\cdot\|_a \neq \|\cdot\|_V$ il metodo di Galerkin ha proprietà di approssimazione ottimale nello spazio V_h

5.4 Convergenza del MG

TEOREMA: Lemma di Ceà.

$$\|u - u_h\|_1 \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_1 \quad \text{dove } \|\cdot\|_1 \equiv \|\cdot\|_{H^1}$$

Notiamo che M e α sono, rispettivamente le costanti di continuità e coercività di $a(\cdot, \cdot)$, che $\frac{M}{\alpha}$ è indipendente da h e che, in generale, $\frac{M}{\alpha} \geq 1$

DIM.

Applico la coercività all'errore $u - u_h$:

$$a(u - u_h, u - u_h) \geq \alpha \|u - u_h\|_1^2$$

Inoltre vale:

$$a(u - u_h, u - u_h) = a(u - u_h, u - v_h + v_h - u_h) = a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h) \quad \forall v_h \in V_h$$

Per l'ortogonalità di Galerkin vale che $a(u - u_h, v_h - u_h) = 0$ perchè $v_h - u_h \in V_h$

Infine per la continuità di $a(\cdot, \cdot)$ $a(u - u_h, u - v_h) \leq M \|u - u_h\|_1 \cdot \|u - v_h\|_1$

Abbiamo quindi ottenuto che

$$\alpha \|u - u_h\|_1^2 \leq a(u - u_h, u - u_h) \leq M \|u - u_h\|_1 \|u - v_h\|_1 \quad \forall v_h \in V_h$$

Quindi divido per $\alpha \|u - u_h\|_1$ e al posto del per ogni pongo la condizione equivalente con inf

$$\|u - u_h\|_1 \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_1$$

□

Possiamo vedere questo lemma come un principio di equivalenza per MG, infatti la coercività implica la stabilità e l'ortogonalità equivale alla consistenza, quindi $\text{stabilità} + \text{consistenza} \implies \text{convergenza}$

Per valutare effettivamente la convergenza dobbiamo porre $\lim_{\substack{h \rightarrow 0 \\ N_h \rightarrow \infty}} \|u - u_h\|_1 = 0$ e per il lemma di CEA chiedi-

$$\lim_{\substack{h \rightarrow 0 \\ N_h \rightarrow \infty}} \inf_{v_h \in V_h} \|u - v_h\|_1 = 0$$

Questa proprietà è vera se $\lim_{\substack{h \rightarrow 0 \\ N_h \rightarrow \infty}} V_h$ è denso in V , ovvero se al limite per $h \rightarrow 0, N_h \rightarrow \infty$, data u , in ogni intorno di u esiste una funzione $v_h \in V_h$

Quindi per verificare la convergenza è necessario considerare una definizione specifica di V_h

5.5 Analisi FEM

Valutiamo $\inf_{v_h \in V_h} \|u - v_h\|_1$ nel caso di FEM 1D, dove $V_h \equiv X_N^r$

Per fare ciò valutiamo tre funzioni:

- la soluzione u del PVA
- la soluzione u_h di MG con FEM \mathbb{P}^r
- $\Pi_h^r u$ l'interpolante \mathbb{P}^r di u

Sappiamo che $\inf_{v_h \in V_h} \|u - v_h\|_1 \leq \|u - v_h\| \quad \forall v_h \in V_h \quad \text{ma } \Pi_h^r u \in V_h$

Possiamo quindi maggiorare l'errore FEM con l'errore di interpolazione:

$$\|u - u_h\|_1 \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_1 \leq \frac{M}{\alpha} \|u - \Pi_h^r u\|_1$$

Conosciamo l'errore di interpolazione con norma del massimo su funzioni C^{r+1} , vogliamo generalizzare per norma H^1 su funzioni $u \in H^1$, vale il seguente:

TEOREMA: Piecewise lagrangian interpolation.

Sia $v \in H^{r+1}$ per $r \geq 1$ e sia $\Pi_h^r v \in X_h^r$ la sua funzione interpolante, allora

$$|v - \Pi_h^r v|_{H^k(I)} \leq C_{k,r} h^{r+1-k} |v|_{H^{r+1}(I)} \quad \text{per } k = 0, 1$$

Dove le costanti $C_{k,r}$ sono indipendenti da v ed h

Ricordiamo che:

Norma H^s : $\|v\|_{H^s}^2 = \sum_{i=0}^s \|\partial_x^{(i)} v\|_{L^2}^2$

Seminorma H^s : $|v|_{H^s}^2 = \|\partial_x^{(s)} v\|_{L^2}^2$ ovvero ha solo la norma della derivata di ordine massimo

Oss. Quindi servirà $k = 0$ per stimare la norma L^2 e la somma di $k = 0$ e $k = 1$ per la norma H^1

Esempio: Stima dell'errore di interpolazione per FEM \mathbb{P}^1 in H^1 , presa $u \in H^2$

$$\|u - \Pi_h^1 u\|_{H^1}^2 = \|u - \Pi_h^1 u\|_{L^2}^2 + \|\partial_x(u - \Pi_h^1 u)\|_{L^2}^2 \leq C_{0,1}^2 h^4 |u|_{H^2}^2 + C_{1,1} h^2 |u|_{H^2}^2 \leq C h^2 |u|_{H^2}^2$$

$$\xRightarrow{\text{sotto radice}} \|u - \Pi_h^1 u\|_{H^1} \leq C h |u|_{H^2}$$

L'interpolante FEM \mathbb{P}^1 converge linearmente a u per $h \rightarrow 0$

$$\xRightarrow{\text{Per Cea}} \|u - u_h\|_{H^1} \leq \frac{M}{\alpha} \|u - \Pi_h^1 u\|_{H^1} \leq \frac{M}{\alpha} C h |u|_{H^2}$$

Quindi se $u \in H^2(\Omega)$ il metodo degli elementi finiti lineari in 1D converge linearmente in norma H^1

Generalizziamo al caso FEM \mathbb{P}^r con il seguente:

TEOREMA.

Siano $u \in V$ la soluzione esatta del PVA, u_h la sua soluzione approssimata con FEM di grado r , ovvero $u_h \in V_h = X_h^r \cap V$. Sia inoltre $u \in H^{p+1}$ per un $p \geq r$, allora

$$\|u - u_h\|_{H^1} \leq \frac{M}{\alpha} Ch^r |u|_{H^{r+1}} \quad \text{con } C \text{ indipendente da } u \text{ ed } h$$

Invece nel caso in cui $p < r$, vale il seguente:

TEOREMA.

Se $u \in H^{p+1}$, per un $p > 0$, allora esiste una costante C indipendente da u ed h tale che

$$\|u - u_h\|_{H^1} \leq \frac{M}{\alpha} Ch^s |u|_{H^{s+1}} \quad \text{con } s = \min\{r, p\}$$

Questo vuol dire che la regolarità di u è una barriera per l'ordine di convergenza di FEM

r	$u \in H^1(I)$	$u \in H^2(I)$	$u \in H^3(I)$	$u \in H^4(I)$	$u \in H^5(I)$
1	converge	h^1	h^1	h^1	h^1
2	converge	h^1	h^2	h^2	h^2
3	converge	h^1	h^2	h^3	h^3
4	converge	h^1	h^2	h^3	h^4

Si ha convergenza ottimale se $p \geq r$ e quindi $s = r$

Convergenza sotto ottimale se $p < r$ e quindi $s = p$

Invece nel caso $u \in H^1$ si può dire che il metodo FEM converge $\lim_{h \rightarrow 0} \|u - u_h\|_{H^1} = 0$ ma non è possibile caratterizzare un ordine di convergenza

Oss. I risultati di convergenza del metodo FEM non dipendono dalla dimensionalità del dominio

5.6 Stima dell'errore in L^2

Per la convergenza in norma L^2 si ha $\|u - u_h\|_{L^2} \leq \|u - u_h\|_{H^1} \leq \frac{M}{\alpha} Ch^r |u|_{H^{r+1}}$

Però questa disuguaglianza non è ottimale, si può ottenere un ordine superiore con una norma più debole

TEOREMA.

Siano $u \in V$ la soluzione esatta del PVA ed u_h la sua soluzione approssimata da FEM di grado r . Sia inoltre $u \in C^0(\bar{\Omega}) \cap H^{p+1}(\Omega)$ per un $p > 0$, allora

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch^{s+1} |u|_{H^{s+1}(\Omega)} \quad \text{con } s = \min\{r, p\} \text{ e } C \text{ indipendente da } u \text{ ed } h$$

6 FEM Varie

6.1 Condizioni al bordo

Si vuole analizzare un problema di Cauchy non omogeneo con FEM:

$$\begin{cases} -\nabla u = f & \text{in } \Omega \\ u = g & \text{su } \partial\Omega \end{cases}$$

In formulazione variazionale: Trovare $u \in \{v \in H^1(\Omega) : v|_{\partial\Omega} = g\}$ tc $a(u, v) = (f, v) \forall v \in H_0^1(\Omega)$

Non si può applicare L.M. perché lo spazio delle soluzioni è diverso dallo spazio delle funzioni test

Si applica il metodo del rilevamento della traccia per trasformare le condizioni al bordo

Si cerca una funzione $R(g) \in H^1(\Omega)$ che assume g su $\partial\Omega$

In questo modo si risolve il PVA con un'incognita ausiliaria $w \in H_0^1(0, 1)$ tc la soluzione sia $u = w + R(g)$

Dovremo risolvere: Trovare $w \in H_0^1(\Omega)$ tc $a(u, v) = (f, v) - a(R(g), v) \forall v \in H_0^1$

Per il rilevamento del dato al bordo con FEM, l'idea è quella di usare le basi che valgono 1 sul singolo vertice del bordo e 0 negli altri vertici

Si prende $R_h(g) = (I_h(g)) = \sum_{i=1}^{N_h^b} g(x_i) \varphi_i(x)$ dove N_h^b è il numero di nodi al bordo e $R_h(g) \in H^1(\Omega)$ perché è combinazione delle basi che sono funzioni $H^1(\Omega)$

6.2 FEM 2D

Per approssimare la soluzione con polinomi di secondo grado a tratti, avrò bisogno di più nodi perché una parabola è identificata univocamente con 3 punti. Di conseguenza si introducono nuovi nodi, ovvero i punti medi per ogni sotto intervallo

Sarà necessario definire le basi su tutti i nodi quindi sia sui vertici che sui punti medi. Si osserva che le basi dei punti medi, invece che essere definite su due sotto intervalli, sono definite su un solo sotto intervalli, per questo sono dette bolle

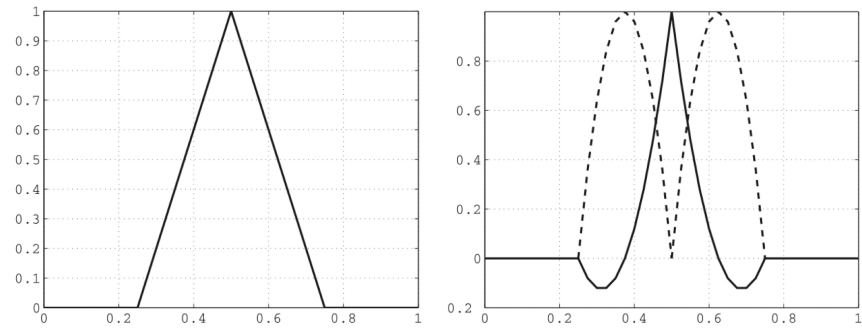


Figure 1: Esempio di base lineare e quadratica, che è nulla nei punti medi
Inoltre tratteggiate sono le due basi dei punti medi

Inoltre si nota che rispetto al caso 1D, dove la matrice locale era 2×2 , nel caso 2D la matrice locale è 3×3
La matrice globale sarà a banda con banda 5, aumentando la dimensione si peggiora la sparsità della matrice

7 FEM diffusione e trasporto

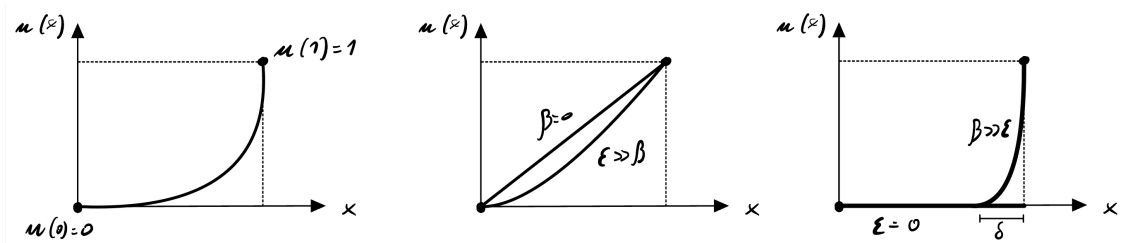
7.1 Introduzione

Analizziamo il problema omogeneo di diffusione e trasporto nel caso unidimensionale e stazionario

$$\begin{cases} -\varepsilon u'' + \beta u' = 0 & 0 < x < 1 \\ u(0) = 0; \quad u(1) = 1 \end{cases}$$

Questo problema essendo stazionario ha ε e β costanti nel tempo. La soluzione sarà unica e costante nel tempo

In questo particolare il problema equivale a una EDO lineare ai limiti con coefficienti costanti possiamo calcolarne la soluzione:



$$u(x) = \frac{e^{\frac{\beta x}{\varepsilon}} - 1}{e^{\frac{\beta}{\varepsilon}} - 1}$$

La soluzione del problema dipende dai parametri ε e β , per capirne l'effetto su u vediamo i casi limite:

- $\varepsilon \gg \beta$ diffusione dominante
$$\begin{cases} -\varepsilon u'' = 0 & 0 < x < 1 \\ u(0) = 0; \quad u(1) = 1 \end{cases}$$

avendo derivata seconda nulla, la soluzione è una retta

- $\beta \gg \varepsilon$ trasporto dominante
$$\begin{cases} -\beta u' = 0 & 0 < x < 1 \\ u(0) = 0 \end{cases}$$

avendo derivata prima nulla, la soluzione è costante

Si nota che lo stesso problema con anche $u(1) = 1$ non è ben posto

La regione dove la soluzione risente della diffusione si chiama strato limite e ha spessore $\delta \simeq \frac{\beta}{\varepsilon}$

7.2 Formulazione variazionale

Dato il problema:

$$\begin{cases} -\varepsilon u'' + \beta u' = f & \text{in } (0, 1) \\ u(0) = u(1) = 0 \end{cases}$$

Supponendo che $\varepsilon, \beta \in \mathbb{R}$ siano costanti e $\varepsilon > 0$

Si trova la formulazione variazionale moltiplicando per la funzione test e integrando per parti

PVA: trovare $u \in H_0^1(0, 1)$ tc

$$\varepsilon \int_0^1 u' v' dx + \beta \int_0^1 u' v dx = \int_0^1 f v dx$$

Per analizzare la forma bilineare la separo:

$$a(u, v) = a_\varepsilon(u, v) + a_\beta(u, v) \quad \forall u, v \quad \text{dove } a_\varepsilon = \varepsilon \int_0^1 u' v' dx \text{ e } a_\beta = \beta \int_0^1 u' v dx$$

Si vuol verificare la buona posizione del problema, per fare ciò serve verificare le ipotesi di Lax-Milgram, ovvero se $a(\cdot, \cdot)$ è coerciva e continua

Per la coercività si usa la disuguaglianza di Poincarè

$$\|v\|_{L^2} \leq C_p \|v'\|_{L^2} \quad \forall v \in H_0^1(0, 1) \implies \|v\|_{H^1}^2 \leq (1 + C_p^2) \|v'\|_{L^2}^2$$

$$a_\varepsilon(v, v) = \int_0^1 \varepsilon (v')^2 dx = \varepsilon \|v'\|_{L^2}^2 \geq \frac{\varepsilon}{1 + C_p^2} \|v\|_{H^1}^2$$

Quindi a_ε è coerciva con costante $\frac{\varepsilon}{1 + C_p^2}$

$$a_\beta(v, v) = \int_0^1 \beta v' v dx = - \int_0^1 \beta v v' dx \quad \forall v \in H_0^1$$

Ma essendo uguale al suo opposto $a_\beta(v, v) = 0$

In conclusione $a(v, v) = a_\varepsilon(v, v) + a_\beta(v, v) \geq \frac{\varepsilon}{1 + C_p^2} \|v\|_{H^1}^2$

Quindi $a(\cdot, \cdot)$ è coerciva con costante $\alpha = \frac{\varepsilon}{1 + C_p^2}$

Si osserva che $a_\beta(\cdot, \cdot)$ non è coerciva, quindi il problema con solo trasporto non è ben posto

Per la continuità si applica la disuguaglianza di Cauchy-Schwarz

$$a(u, v) = \int_0^1 (\varepsilon u' v' + \beta u' v) dx \leq \varepsilon \|u'\|_{L^2} \|v'\|_{L^2} + |\beta| \|u'\|_{L^2} \|v\|_{L^2} \leq (\varepsilon + |\beta|) \|u\|_{H^1} \|v\|_{H^1}$$

Quindi $a(\cdot, \cdot)$ è continua con costante $M = \varepsilon + |\beta|$

Grazie a Lax-Milgram sappiamo anche che per la stabilità vale $\|u\|_{H^1} \leq \frac{1}{\alpha} \|f\|_{L^2}$ dove $\alpha = \frac{1 + C_p^2}{\varepsilon} = C_s$

Ma per $\varepsilon \rightarrow 0$ abbiamo $C_s \rightarrow \infty$ questo potrebbe generare instabilità

Analizzando l'approssimazione FEM lineari $\|u - u_h\|_{H^1} \leq C \frac{M}{\alpha} h \|u\|_{H^1}$ dove $\frac{M}{\alpha} = \frac{\varepsilon + |\beta|}{\varepsilon} (1 + C_p^2)$

Quindi nei casi a trasporto dominante $|\beta| \gg \varepsilon$ si ha $\frac{M}{\alpha} \gg 1 \implies$ instabilità nel metodo FEM

7.3 Formulazione matriciale

Metodo di Galerkin - FEM \mathbb{P}^1 per la discretizzazione di $\begin{cases} -\varepsilon u'' + \beta u' = f \\ u(0) = u(1) = 0 \end{cases}$

Trovare la soluzione $u_h \in V_h = X_h^1(0, 1) \subset H_0^1(0, 1)$ tale che $a(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h$

Equivale a risolvere $\underline{A}\underline{U} = \underline{F}$ dove $A_{ij} = a(\varphi_j(x), \varphi_i(x)) = a_\varepsilon(\varphi_j, \varphi_i) + a_\beta(\varphi_j, \varphi_i)$

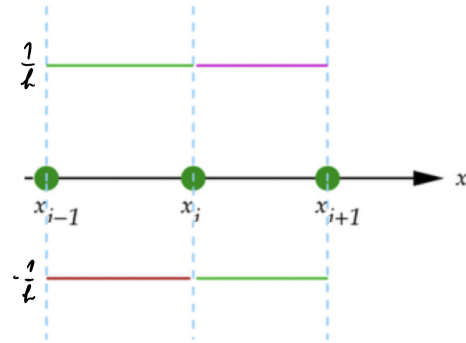
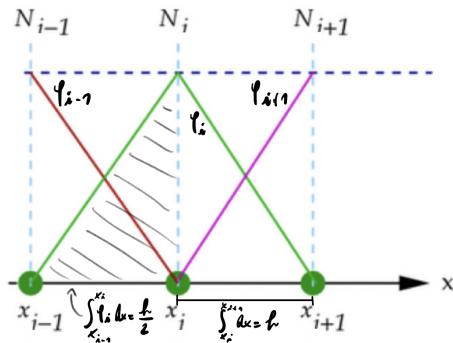
$$a_\varepsilon(\varphi_j, \varphi_i) = (A_\varepsilon)_{ij} = \varepsilon \int_0^1 \varphi_j' \varphi_i' dx \quad a_\beta(\varphi_j, \varphi_i) = (A_\beta)_{ij} = \beta \int_0^1 \varphi_j' \varphi_i dx$$

Fissata la riga i -esima $(\underline{A}\underline{U})_{i,\cdot} = \int_0^1 \varepsilon u_h' \varphi_i' dx + \int_0^1 \beta u_h' \varphi_i dx$

Cerchiamo la soluzione della forma $u_h(x) = \sum_{j=0}^n u_j \varphi_j(x)$ quindi si ottiene

$$\int_0^1 \varepsilon u_h' \varphi_i' dx = \varepsilon \left[u_{i-1} \int_{x_{i-1}}^{x_i} \varphi_{i-1}' \varphi_i' dx + u_i \int_{x_{i-1}}^{x_{i+1}} (\varphi_i')^2 dx + u_{i+1} \int_{x_i}^{x_{i+1}} \varphi_i' \varphi_{i+1}' dx \right]$$

$$\int_0^1 \beta u_h' \varphi_i dx = \beta \left[u_{i-1} \int_{x_{i-1}}^{x_i} \varphi_{i-1}' \varphi_i dx + u_i \int_{x_{i-1}}^{x_{i+1}} \varphi_i' \varphi_i dx + u_{i+1} \int_{x_i}^{x_{i+1}} \varphi_{i+1}' \varphi_i dx \right]$$



$$\int_{x_{i-1}}^{x_i} \varphi_{i-1}' \varphi_i' dx = -\frac{1}{h} \cdot \frac{1}{h} \cdot \int_{x_{i-1}}^{x_i} dx = -\frac{1}{h}$$

$$\text{Analogamente} \quad \int_{x_{i-1}}^{x_{i+1}} (\varphi'_i)^2 dx = 2h \cdot \frac{1}{h^2} \quad \int_{x_i}^{x_{i+1}} \varphi'_i \varphi'_{i+1} dx = -h \cdot \frac{1}{h} \frac{1}{h}$$

$$\int_{x_{i-1}}^{x_i} \varphi'_{i-1} \varphi_i dx = -\frac{1}{h} \int_{x_{i-1}}^{x_i} \varphi_i dx = -\frac{1}{h} \frac{h}{2} = -\frac{1}{2}$$

$$\text{Analogamente} \quad \int_{x_i}^{x_{i+1}} \varphi'_{i+1} \varphi_i dx = \frac{1}{2}$$

$$\int_{x_{i-1}}^{x_{i+1}} \varphi'_i \varphi_i dx = 0 \quad \text{per simmetria, dato che } \varphi_i \text{ è simmetrica e } \varphi'_i \text{ è antisimmetrica}$$

$$\text{Si ottiene la } i\text{-esima riga di } \underline{AU} : \quad -\frac{\varepsilon}{h}(u_{i+1} - 2u_i + u_{i-1}) + \frac{1}{2}\beta(u_{i+1} - u_{i-1}) = hf$$

Il termine noto vale hf perché $f = \text{cost}$ e quindi esce dall'integrale

Si ottiene uno schema equivalente alla discretizzazione DF centrate:

$$A = -\frac{\varepsilon}{h} \begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & \end{bmatrix} + \beta \begin{bmatrix} 0 & +1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & \end{bmatrix}$$

Se ci fosse stato anche il termine di reazione: $-\varepsilon u'' + \beta u' + \gamma u = f \quad \text{in } (0, 1)$

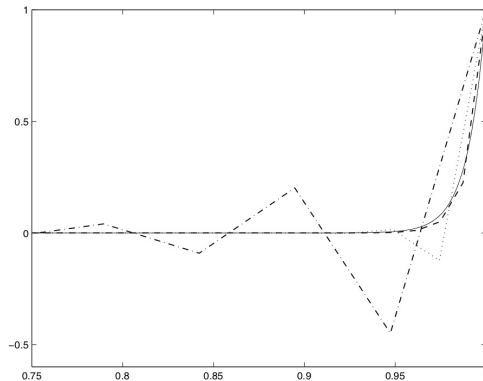
$$a_\gamma(u, v) = \gamma \int_0^1 uv dx \quad A_\gamma = \frac{1}{6}\gamma h \begin{bmatrix} 4 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & \end{bmatrix}$$

7.4 FEM con diffusione artificiale

Abbiamo visto che per il problema
$$\begin{cases} -\varepsilon u'' + \beta u' = 0 & \text{in } (0, 1) \\ u(0) = 0; u(1) = 1 \end{cases}$$

nel caso di trasporto dominante ci può essere instabilità

Infatti usando FEM-centrato, per alcuni valori di $\frac{\beta}{\varepsilon}$ la soluzione numerica oscilla:



A partire dal problema $-\frac{\varepsilon}{h^2}(u_{i+1} - 2u_i + u_{i-1}) + \frac{1}{2}\frac{\beta}{h}(u_{i+1} - u_{i-1}) = 0$; $u_0 = 0$; $u_N = 1$

Definiamo il numero di Peclet locale $Pe = \frac{|\beta|h}{2\varepsilon}$, allora lo schema "centrato" è equivalente a:

$$\begin{cases} (Pe - 1)u_{i+1} + 2u_i - (Pe + 1)u_{i-1} = 0 \\ u_0 = 0; u_N = 1 \end{cases}$$

Per ottenere una formula esplicita per la soluzione numerica di u_i si cercano le radici dell'equazione caratteristica e poi si impongono le condizioni ai limiti

$$u_i = \frac{\left(\frac{1+Pe}{1-Pe}\right)^i - 1}{\left(\frac{1+Pe}{1-Pe}\right)^N - 1}$$

Nel caso di diffusione dominante $\frac{|\beta|h}{2\varepsilon} < 1$ si ottiene $\left(\frac{1+Pe}{1-Pe}\right) > 1$ per cui la soluzione numerica ha andamento monotono crescente

Invece per trasporto dominante $\frac{|\beta|h}{2\varepsilon} > 1$ per cui $\left(\frac{1+Pe}{1-Pe}\right) < -1 \implies \begin{cases} > 0 & \text{se } i \text{ pari} \\ < 0 & \text{se } i \text{ dispari} \end{cases}$

Per risolvere questo problema si potrebbe prendere h sufficientemente piccolo:

$$h < h^* \quad \text{con} \quad \frac{|\beta|h^*}{2\varepsilon} = 1 \implies h^* = \frac{2\varepsilon}{|\beta|}$$

In questo modo $Pe < 1$ però per alcuni problemi reali si ottengono numeri irrealizzabili

L'alternativa per stabilizzare il metodo FEM per problemi a trasporto dominante è cercare uno schema decentrato, ovvero che se $\beta > 0$ allora la soluzione u_i non dipende da x_{i+1} e viceversa

Scriviamo lo schema decentrato upwind in funzione dello schema centrato

$$\begin{aligned} \beta > 0 \quad \frac{1}{h}(U_i - U_{i-1}) &= D_x^- U_i = D_x^c U_i - \frac{h}{2} D_{xx}^2 U_i = \frac{1}{2h}(U_{i+1} - U_{i-1}) - \frac{h}{2} \frac{1}{h^2}(U_{i+1} - 2U_i + U_{i-1}) \\ \beta < 0 \quad \frac{1}{h}(U_{i+1} - U_i) &= D_x^+ U_i = D_x^c U_i + \frac{h}{2} D_{xx}^2 U_i = \frac{1}{2h}(U_{i+1} - U_{i-1}) + \frac{h}{2} \frac{1}{h^2}(U_{i+1} - 2U_i + U_{i-1}) \\ \text{In generale} \quad \beta D_x^{UP} U_i &= \beta D_x^c U_i - \frac{h}{2} |\beta| D_{xx}^2 U_i \end{aligned}$$

Applichiamo questa formula a FEM usando il metodo della diffusione artificiale:

$$-\varepsilon D_{xx}^2 U_i + \beta D_x^{UP} U_i = \underbrace{-\varepsilon D_{xx}^2 U_i + \beta D_x^c U_i}_{\text{FEM } \mathbb{P}^1} - \frac{1}{2} h |\beta| D_{xx}^2 U_i$$

A livello variazionale il nuovo metodo FEM con diffusione artificiale è:

$$a_h(u_h, v_h) = a(u_h, v_h) + \frac{h}{2}|\beta| \int_0^1 u'_h v'_h dx = a_\varepsilon(u_h, v_h) + a_\beta(u_h, v_h) + b_h(u_h, v_h)$$

Per analizzare il metodo rappresentiamo la stessa forma bilineare con una perturbazione di ε

$$a_h(u_h, v_h) = a_{\varepsilon_h}(u_h, v_h) + a_\beta(u_h, v_h) \quad \text{con } \varepsilon_h = \varepsilon + \frac{1}{2}h|\beta| = \varepsilon(1 + Pe)$$

Il numero di Peclet del problema perturbato è $\widehat{Pe} = \frac{|\beta|h}{2\varepsilon_h} = \frac{Pe}{Pe+1} < 1$

Questa relazione garantisce che FEM con diffusione artificiale non presenti oscillazioni ed instabilità perché equivale a uno schema di tipo centrato con $\widehat{Pe} < 1$

7.5 FEM stabilizzati

FEM con diffusione artificiale:

$$a_h(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h \subset V$$

Si perde la consistenza del metodo di Galerkin, valutiamo la consistenza:

$$\text{Residuo} = a_h(u, v_h) - (f, v_h) = a(u, v_h) + b_h(u, v_h) - (f, v_h)$$

Per la consistenza forte del MG otteniamo $\text{Residuo} = b_h(u, v_h) = \frac{1}{2}h|\beta| \int_0^1 u' v'_h dx$

Si vuole studiare l'asintotica stabilità, per fare ciò dobbiamo togliere la dipendenza del residuo da v_h variabile

$$\text{Residuo normalizzato} = R_h(u) \sup_{v_h \in V_h} \frac{a_h(u, v_h) - (f, v_h)}{\|v_h\|_{H^1}} = \sup_{v_h \in V_h} \frac{b_h(u, v_h)}{\|v_h\|_{H^1}}$$

$b_h(\cdot, \cdot)$ è continua, infatti $b_h(u, v) = \frac{1}{2}h|\beta| \int_0^1 u' v' dx \leq \frac{1}{2}h|\beta| \|u'\|_{L^2} \|v'\|_{L^2} \leq \frac{1}{2}h|\beta| \|u'\|_{H^1} \|v'\|_{H^1}$

$$\implies R_h(u) = \sup_{v_h \in V_h} \frac{b_h(u, v_h)}{\|v_h\|_{H^1}} \leq \sup_{v_h \in V_h} \frac{1}{2}h|\beta| \frac{\|u\|_{H^1} \|v_h\|_{H^1}}{\|v_h\|_{H^1}}$$

$$R_h(u) \leq \frac{1}{2}h|\beta| \|u\|_{H^1} \implies \lim_{h \rightarrow 0} R_h(u) = 0 \quad \text{con } R_h(u) \simeq \mathcal{O}(h)$$

Quindi il metodo FEM-diffusione artificiale è asintoticamente consistente di ordine 1

Applicando Lax-Milgram al problema con trasporto dominante avevamo visto che la soluzione può presentare oscillazioni perché $C_s \rightarrow \infty$ per $\varepsilon \rightarrow 0$, si dice che il metodo non è robusto rispetto a ε

Applicando LM a FEM con diffusione artificiale otteniamo

$$\|\widehat{u}_h\|_{H^1} \leq \frac{1}{\alpha_h} \|f\|_{L^2} \quad \text{con } \alpha_h = \frac{\varepsilon_h}{1 + C_p^2} \implies C_s = \frac{1 + C_p^2}{\varepsilon_h} = \frac{1 + C_p^2}{\varepsilon + \frac{1}{2}h|\beta|}$$

C_s è limitata superiormente per $\varepsilon \rightarrow 0$ e quindi il metodo è robusto per $\varepsilon \rightarrow 0$

In questo caso non è valido il lemma di CEA, quindi per valutare la convergenza si usa il lemma di Strang:

$$\|u - u_h\|_1 \leq \left(1 + \frac{M_h}{\alpha_h}\right) \inf_{v_h \in V_h} \|u - v_h\|_1 + \frac{1}{\alpha_h} \sup_{v_h \in V_h} \frac{|a_h(u, v_h) - (f, v_h)|}{\|v_h\|_1}$$

La prima parte equivale al lemma di CEA e quindi è il contributo dell'errore dovuto alle proprietà di approssimazione dello spazio V_h

La seconda parte equivale al residuo normalizzato, ovvero il contributo della consistenza asintotica dell'errore globale. Questa parte sappiamo essere asintoticamente convergente con ordine 1: $R_h(u) = \mathcal{O}(h)$

In conclusione il metodo FEM - diffusione artificiale stabilizza le oscillazioni quando $Pe > 1$, ma pone un limite superiore all'ordine di convergenza per qualunque grado polinomiale