

ELEMENTI FINITI PER L'EQUAZIONE DI POISSON

Abbiamo visto come il problema

$$-\Delta u = f$$

può essere risolto in generale trovati metodi iterativi in quanto la soluzione diretta del sistema di equazioni risultante è in generale troppo costosa.

In particolare abbiamo osservato come il metodo iterativo di Jacobi sia equivalente ad un metodo esplicito applicato all'equazione del calore bi-dimensionale

$$u_t = \Delta u + f$$

con la scelta del massimo Δt ammissibile dalla condizione di stabilità ossia $\Delta t = \frac{\Delta x^2}{4}$ con $\Delta x = \Delta y$.

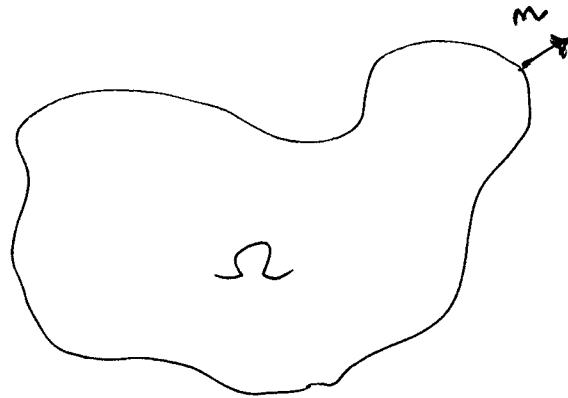
Abbiamo infatti che per $t \rightarrow \infty$ l'equazione del calore raggiunge una soluzione stazionaria u_t non si lavora più variando $u_t = 0$. Ne consegue che

$$-\Delta u = f$$

ossia la soluzione cercata.

Lo svantaggio principale del metodo visto basato sulle differenze finite, sta nella difficoltà di

trattare problemi con geometrie complicate, quel che
pratico si presenterà



TH. DIVERGENZA : Dato un vettore $\sigma = (v_1, v_2)$ definito
in Ω si ha

$$\int_{\Omega} \operatorname{div} \sigma \, dX = \int_{\partial\Omega} \sigma \cdot n \, d\delta$$

$n = (n_1, n_2)$ vettore unitario normale a $\partial\Omega$, dX elemento di
area in \mathbb{R}^2 , $d\delta$ elemento di lunghezza dell'arco lungo $\partial\Omega$.

NOTA: $\operatorname{div} \sigma = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y}$.

Applicandola ora ai vettori $\sigma = \left(w \frac{\partial u}{\partial x}, 0 \right)$, $\sigma = \left(0, w \frac{\partial u}{\partial y} \right)$
si ha

$$\left\{ \begin{array}{l} \int_{\Omega} \frac{\partial w}{\partial x} \frac{\partial u}{\partial x} \, dX + \int_{\Omega} w \frac{\partial^2 u}{\partial x^2} \, dX = \int_{\partial\Omega} w \frac{\partial u}{\partial x} n_1 \, d\delta \\ \int_{\Omega} \frac{\partial w}{\partial y} \frac{\partial u}{\partial y} \, dX + \int_{\Omega} w \frac{\partial^2 u}{\partial y^2} \, dX = \int_{\partial\Omega} w \frac{\partial u}{\partial y} n_2 \, d\delta \end{array} \right.$$

Sommando le due precedenti espressioni otteniamo

$$\int_{\Omega} \text{grad } w \cdot \text{grad } u \, dX = \int_{\partial\Omega} w \frac{\partial u}{\partial n} \, d\sigma - \int_{\Omega} w \Delta u \, dX$$

con $\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x} n_1 + \frac{\partial u}{\partial y} n_2$, nota come formule

d. Green.

Considero ora il problema

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ u = 0 & \text{in } \partial\Omega \end{cases}$$

moltiplicando per w
e integrando rispetto ad Ω otteniamo la formulazione
variazionale del problema, trovare $u \in V$ t.c.

$$-\int_{\Omega} \Delta u w \, dX = \int_{\Omega} f w \, dX \quad \forall w \in V$$

con V spazio ... funzionale opportuno.

Applicando ora il precedente risultato abbiamo

$$+\int_{\Omega} \text{grad } w \cdot \text{grad } u \, dX - \int_{\partial\Omega} w \frac{\partial u}{\partial n} \, d\sigma = \int_{\Omega} f w \, dX$$

ossia $a(u, w) = (f, w) \quad \forall w \in V$

con $V = \left\{ w / w \in C^0(\bar{\Omega}), \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y} \in L^2(\Omega), w = 0 \text{ su } \partial\Omega \right\}$
e continue a pezzi in Ω 93

e (u, w) prodotto scalare -

Si dimostra che $u \in V$ soddisfa al precedente problema
 se e solo se u è soluzione del problema di minimizzare

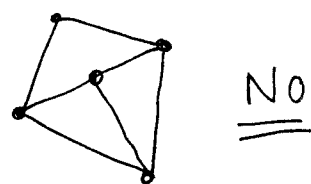
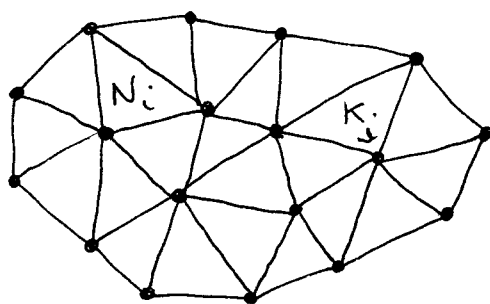
$$J(u) \leq J(w) \quad \forall w \in V$$

$$J(w) = \frac{1}{2} a(w, w) - (f, w).$$

DEFINIZIONE DI UNO SPAZIO DISCRETO

Assumiamo per semplicità' ora che $\partial\Omega$ sia una poligonale
 e introduciamo in Ω una triangolazione T_h

$$\Omega = \bigcup_{K_j \in T_h}^m K_j = K_1 \cup K_2 \cup \dots \cup K_m$$



$$h = \max_{K_j \in T_h} \text{diam}(K_j), \quad \text{diam}(K_j) = \text{diametro di } K_j$$

$$V_h = \left\{ w_h / w_h \in C^0(\bar{\Omega}), w_h \text{ lineare in } K \in T_h, (w_h)_{\partial\Omega} = 0 \right\}$$

tutte le funzioni continue lineari in ogni triangolo K e che
 si annullano su $\partial\Omega$.

Si osserva che $V_h \subset V$.

Per descrivere $w_h \in V$ utilizziamo i nodi N_i ,
 $i = 1, \dots, M$ interni al dominio Ω .

Definiamo ora le funzioni $\phi_j \in V_h$ t.c.

$$\phi_j(N_i) = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \quad i, j = 1, \dots, M$$

e definiamo

$$w_h(x) = \sum_{j=1}^M \lambda_j \phi_j(x) \quad , \quad \lambda_j = w_h(N_j)$$

per $x \in \Omega$. (vedi figura successiva).

Abbiamo quindi il problema discreto (elementi finiti)

trovare $u_h \in V_h$ t.c. $a(u_h, w_h) = (f, w_h) \quad \forall w_h \in V_h$.

Il che equivale a risolvere un insieme lineare delle form

$$A \xi = b$$

con $A = (a_{ij})$, $a_{ij} = a(\phi_i, \phi_j)$, $b = (b_i)$

$b_i = (f, \phi_i)$, $\xi = (\xi_i)$, $\xi_i = u_h(N_i)$.

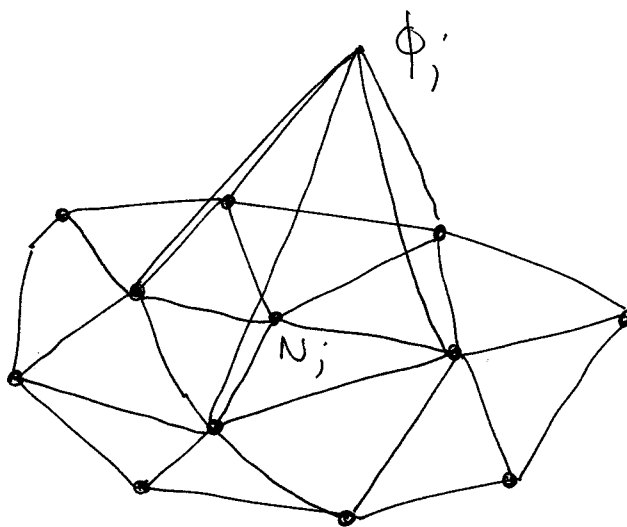
A detta matrice di stiffness ed è simmetrica e definita
positiva, inoltre A è sparsa in quanto $a_{ij} = 0$
a meno che N_i e N_j non siano nodi nello stesso
triangolo.

Gli elementi a_{ij}

$$a(\phi_i, \phi_j) = \sum_{K \in T_h} a_K(\phi_i, \phi_j)$$

$$= \sum_{K \in T_h} \int_K \text{grad } \phi_i \cdot \text{grad } \phi_j \, dX$$

sono dati dalla somma dei contributi dei diversi triangoli.



Il metodo appena descritto è detto metodo di Galerkin agli elementi finiti.

EQUAZIONE DI POISSON: ELEMENTI FINITI

Abbiamo visto che il problema

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ u = 0 & \text{su } \partial\Omega \end{cases}$$

($\Omega \subset \mathbb{R}^2$ dominio limitato con frontiera $\partial\Omega$)

in forma debole ha la forma

$$+ \int_{\Omega} \text{grad } w \cdot \text{grad } u \, dX - \int_{\partial\Omega} w \frac{\partial u}{\partial n} \, d\sigma = \int_{\Omega} f w \, dX$$

per $w \in V$ spazio funzionale opportuno.

Scegliendo funzioni test w nulle al bordo otteniamo

$$+ \int_{\Omega} \text{grad } w \cdot \text{grad } u \, dX = \int_{\Omega} f w \, dX$$

che in forma compatta diventa

$$a(u, w) = (f, w) \quad \forall w \in V,$$

con $a(u, w)$ forma bilineare.

Per la disuguaglianza di Cauchy-Schwartz

$$\left| \int_{\Omega} fg \, d\Omega \right| \leq \|f\|_{L^2} \|g\|_{L^2}$$

$$\text{con } \|\cdot\|_{L^2} = \sqrt{\int_{\Omega} |\cdot| \, d\Omega}.$$

Dunque $a(u, w)$ ha senso se ∇u e $\nabla w \in L^2$

e questo è garantito se $u, w \in H^1(\Omega)$ spazio di Hilbert

$$H^1(\Omega) = \left\{ v: \Omega \rightarrow \mathbb{R} \mid v \in L^2(\Omega), \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y} \in L^2(\Omega) \right\}.$$

In particolare il termine (f, w) sarà ben definito se

$$f \in L^2(\Omega).$$

Vali il seguente risultato

Lemma (Lax-Milgram): Sia V uno spazio di

Hilbert, $a(\cdot, \cdot): V \times V \rightarrow \mathbb{R}$ una forma bilineare

continua e coerciva. Sia $F: V \rightarrow \mathbb{R}$ un funzionale

lineare e continuo. Allora esiste unica la soluzione
del problema:

$$\text{cercare } u \in V : a(u, w) = (f, w) \quad \forall w \in V.$$

Oss: Una forma bilineare è coerciva se

$$a(u, u) \geq \alpha \|u\|_V^2 \quad \forall u \in V$$

Oss: Se la forma bilineare è simmetrica, ossia

$a(u, w) = a(w, u) \quad \forall u, w \in V$ allora il problema è equivalente a:

cercare $u \in V$: $J(u) = \min_{w \in V} J(w)$

con

$$J(w) = \frac{1}{2} a(w, w) - (f, w)$$

Infatti per $v \in V$

$$\begin{aligned} J(u+v) &= \frac{1}{2} [a(u, u) + 2a(u, v) + a(v, v)] - \\ &\quad - [(f, u) + (f, v)] = J(u) + [a(u, v) - (f, v)] \\ &\quad + \frac{1}{2} a(v, v) = J(u) + \frac{1}{2} a(v, v) \end{aligned}$$

Dunque

$$J(u+v) \geq J(u) + \frac{\alpha}{2} \|v\|_V^2$$

Posto $w = u+v$ per l'arbitrarietà di v si ha

$$J(w) > J(u) \quad \forall w \in V, \quad w \neq u.$$

ANALISI METODO DI GALERKIN

Abbiamo visto che il metodo di Galerkin corrisponde al problema approssimato

$$(1) \quad a(u_h, w_h) = (f, w_h) \quad \forall w_h \in V_h$$

con $V_h \subset V$ spazio di approssimazione opportuno.

Esempio: V_h spazio funzioni lineari caratterizzate dalle base φ_j , $j = 1, 2, \dots, M$ t.c.

$$\varphi_j(N_i) = \delta_{ij}$$

con N_i , $i = 1, 2, \dots, M$ nodi della mesh computazionale.

In genere data φ_j una base di V_h avremo

$$a(u_h, \varphi_i) = (f, \varphi_i) \quad i = 1, 2, \dots, M$$

Ora poiché $u_h \in V_h$ avremo

$$u_h(x) = \sum_{j=1}^M u_j \varphi_j(x)$$

con u_j coefficienti incogniti ed il problema (1)

diventa

$$\sum_{j=1}^M u_j a(\varphi_j, \varphi_i) = (f, \varphi_i) \quad i = 1, \dots, M$$

Introdotta la matrice di stiffness $a_{ij} = a(\varphi_j, \varphi_i)$

abbiamo

$$(2) \quad A u = b.$$

Teorema: La matrice A è definita positiva e simmetrica.

Dim: La simmetria è conseguenza immediata della definizione di $a(\cdot, \cdot)$.

Vogliamo dimostrare che dato $u_h \in V_h$ $u_h^T A u_h \geq 0$ e
vale $u_h^T A u_h = 0 \Leftrightarrow u_h = 0$. Nota che $u = (u_1, \dots, u_n)$.

$$\begin{aligned} u_h &= \sum_j u_j \varphi_j \Rightarrow u_h^T A u_h = \sum_j \sum_i u_i a_{ij} u_j = \\ &= \sum_j \sum_i u_i a(\varphi_j, \varphi_i) u_j = \sum_j \sum_i a(u_j \varphi_j, u_i \varphi_i) \\ &= a\left(\sum_j u_j \varphi_j, \sum_i u_i \varphi_i\right) = a(u_h, u_h) \geq \alpha \|u_h\|_V^2 \geq 0. \end{aligned}$$

oss: Dunque la risoluzione può essere effettuata in modo efficiente tramite decomposizione di Cholesky o metodi iterativi come il gradient conjugate.

Direttamente dal lemma di Lax-Nilgran si ottiene che

Corollario I: La soluzione di (1) esiste ed è unica.

Osserviamo che poiché (1) è equivalente al sistema lineare (2) ed A è definita positiva allora tale sistema ammette una unica soluzione. Questo è un altro modo di dimostrare il corollario precedente.

Corollario II: Il problema (1) è stabile nel senso che vale

$$\|u_h\|_V \leq \frac{1}{\alpha} G \|f\|_{L^2}$$

dim

Dalla coercività di $a(\cdot, \cdot)$ discende

$$\alpha \|u_h\|_V^2 \leq a(u_h, u_h) = (f, u_h)$$

$$(f, u_h) \leq \|f\|_{L^2} G \|u_h\|_V.$$

□

Si dimostra inoltre che il metodo è consistente ossia:

Lemma: Il metodo di Galerkin è fortemente consistente

$$a(u - u_h, w_h) = 0 \quad \forall w_h \in V_h.$$

da
 $V_h \subset V$ vale

$$a(u, w_h) = (f, w_h) \quad \forall w_h \in V_h$$

na avendo

$$a(u_h, w_h) = (f, w_h)$$

sottraendo dalla bilinearità si ottiene la tesi. \square

Oss. Il risultato precedente esprime il fatto che u_h è la miglior approssimazione di u in V_h . Se $a(\cdot, \cdot)$ fosse il prodotto scalare euclideo, u e u_h vettori allora l'espressione trovata esprimerebbe il fatto che l'errore $u - u_h$ è ortogonale rispetto al sottospazio V_h di V .

Convergenza del metodo di Galerkin

Consideriamo ora

$$\begin{aligned} a(u - u_h, u - u_h) &= a(u - u_h, u - w_h) + \overbrace{a(u - u_h, w_h - u_h)}^0 \\ &= a(u - u_h, u - w_h) \end{aligned}$$

dove $a(u - u_h, w_h - u_h) = 0$ poiché $w_h - u_h \in V_h$.

$$|a(u - u_h, u - u_h)| \leq M \|u - u_h\|_V \|u - w_h\|_V$$

in quanto la forma bilineare è continua.

Per la coercività si ha inoltre

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

da cui

$$\|u - u_h\|_V \leq \frac{\beta}{\alpha} \|u - w_h\|_V \quad \forall w_h \in V_h$$

In particolare

$$\|u - u_h\|_V \leq \frac{\beta}{\alpha} \inf_{w_h \in V_h} \|u - w_h\|_V$$

Dunque se per $h \rightarrow 0$ lo spazio V_h coincide con V allora

$$\lim_{h \rightarrow 0} \|u - u_h\|_V \leq \frac{\beta}{\alpha} \lim_{h \rightarrow 0} \inf_{w_h \in V_h} \|u - w_h\|_V = 0$$

ed il metodo è convergente.

La scelta dello spazio V_h è dunque fondamentale e deve garantire tale convergenza. La velocità di convergenza dipenderà dalla scelta di V_h e dalla regolarità di u .

UNA DEFINIZIONE DI ELEMENTO FINITO

Un elemento dello spazio degli elementi finiti è caratterizzato da tre ingredienti

1. Il dominio di definizione K dell'elemento (ad esempio un triangolo, un quadrilatero nel caso b. dimensionale)
2. Lo spazio dei polinomi P_r definito su K
3. Un insieme finito di punti di K (detti gradi di libertà), $\{a_j\}$, $j = 1, \dots, m$ con $m \geq r$.

Inoltre l'insieme $\Sigma = \{a_j\}_{j=1, \dots, m}$ deve essere univoco rispetto a P_r ossia il problema dell'interpolazione deve essere ben posto.

Una condizione necessaria è che $\dim(P_r) = \text{card}(\Sigma)$ poi basta fornire le funzioni base $\varphi_i \in P_r$ tali che

$$\varphi_i(a_j) = \delta_{ij}, \quad \forall i, j.$$

Esempio

In due dimensioni possiamo considerare

$$\mathbb{P}_1 = \{ f(x,y) = a + bx + cy, \quad a, b, c \in \mathbb{R} \}$$

$$\mathbb{P}_2 = \{ f(x,y) = a + bx + cy + dxy + ex^2 + gy^2, \quad a, b, c, d, e, g \in \mathbb{R} \}$$

\vdots

$$\mathbb{P}_r = \left\{ f(x,y) = \sum_{i+j \leq r} a_{ij} x^i y^j \quad \text{con } a_{ij} \in \mathbb{R} \right\}$$

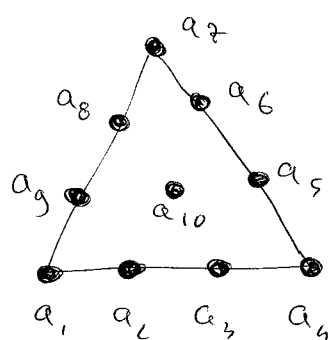
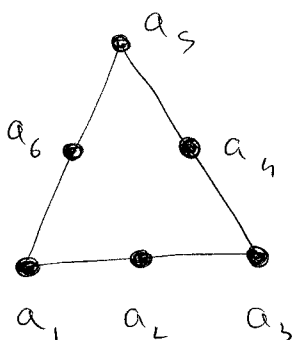
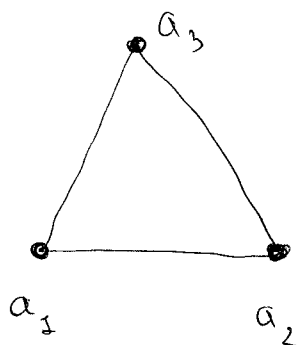
Lo spazio generato dagli element. punti h è P_h

$$X_h^r = \{ \sigma_h \in C^0(\bar{\Omega}) : \sigma_h|_K \in \mathbb{P}_r, \quad \forall K \in \mathcal{T}_h \}$$

si nota che vale

$$\dim(\mathbb{P}_r) = \frac{(r+1)(r+2)}{2}$$

Dunque $\dim(\mathbb{P}_1) = 3$, $\dim(\mathbb{P}_2) = 6$, $\dim(\mathbb{P}_3) = 10$
e in una triangolazione abbiamo bisogno rispettivamente di
3, 6 e 10 nodi opportunamente scelti.



OSS: Si può dimostrare che la matrice A associata al metodo di Galerkin è malcondizionata per valori piccoli del parametro h che caratterizza la discretizzazione del dominio computazionale.

In altre parole $h = \max_k h_k$, $h_k = \max_{x,y \in K} |x-y|$

allora il numero di condizionamento $\mu(A)$ della matrice di ordine 2 è dato da

$$\mu(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$$

con λ_{\max} , λ_{\min} rispettivamente massimo e minimo autovalore e' tale che $\mu(A) = \frac{C}{h^2}$ con C costante opportuna dipendente dal grado degli elementi finiti utilizzati.

OSS: Da quanto visto e' chiaro che il metodo di Galerkin può anche essere applicato nel caso unidimensionale. In tale situazione gli elementi finiti sono intervalli e gli spazi coincidono sostanzialmente con le funzioni splines anche se non imponiamo condizioni esterne sulle derivate ma determineremo le polinomiali a tratti utilizzando un numero adeguato di nodi in ogni intervallo.

ELEMENTI FINITI PER EQUAZIONI PARABOLICHE

Consideriamo ora il problema

$$u_t = \Delta u + f \quad x \in \Omega \subset \mathbb{R}^2$$

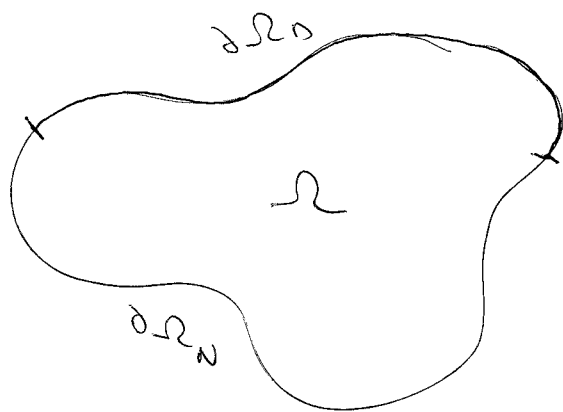
$f = f(x, y, t)$ funzione assegnata. Con le condizioni iniziali e al contorno

$$u(x, y, 0) = u_0(x, y) \quad \text{in } \Omega$$

$$u|_{\partial\Omega_0} = \varphi(x, y) \quad \text{in } \partial\Omega_0$$

in alternative possiamo avere condizioni di tipo Neumann

$$\frac{\partial u}{\partial n} \Big|_{\partial\Omega_N} = \psi(x, y) \quad \text{in } \partial\Omega_N$$



Avremo condizioni miste quando $\partial\Omega = \partial\Omega_0 \cup \partial\Omega_N$,
condizioni di Dirichlet per $\partial\Omega = \partial\Omega_0$ e di Neumann
per $\partial\Omega = \partial\Omega_N$.

Anche in questo caso si procede ad una formulazione debole del problema e si sfruttano i risultati visti per il caso ellittico

$$\int_{\Omega} \frac{\partial u}{\partial t} w \, dx + a(u, w) = (f, w) \quad \forall w \in V$$

dove $a(\cdot, \cdot)$ e (f, \cdot) sono la forma bilineare e il funzionale associati all'operatore di Poisson.

Per semplicità abbiamo supposto $\varphi = 0$ e $\psi = 0$ al bordo.

Possiamo passare direttamente all'approssimazione di Galerkin

$$(1) \quad \int_{\Omega} \frac{\partial u_h}{\partial t} w_h \, dx + a(u_h, w_h) = (f, w_h) \quad \forall w_h \in V_h$$

dove $V_h \subset V$ spazio opportuno a dimensione finita.

A tale problema associamo il dato iniziale

$$u_h(0) = u_{0h}.$$

Il problema (1) è detto problema semi-discreto in quanto la variabile temporale t è continua. Si noti che

rappresentare in diverso modo di applicare il metodo delle linee.

Come in procedure esprimiamo

$$u_h = \sum_{j=1}^M u_j(t) \varphi_j(x, y)$$

con $\{\varphi_j\}$ base di V_h e $u_j(t)$ incognite del nostro problema. Per sostituzione in (1) otteniamo

$$\int_{\Omega} \sum_{j=1}^M u_j'(t) \varphi_j(x, y) \varphi_i(x, y) dX +$$

$$a \left(\sum_{j=1}^M u_j \varphi_j, \varphi_i \right) = (f, \varphi_i) \quad i=1, \dots, M$$

dove si è scelto $\omega_h = \varphi_i$ in quanto ω_h è esprimibile come combinazione lineare della base $\{\varphi_j\}$ e gli operatori coinvolti sono bilineari e lineari.

Ossia

$$\sum_{j=1}^M u_j' \int_{\Omega} \varphi_j \varphi_i dX + \sum_{j=1}^M u_j a(\varphi_j, \varphi_i) = (f, \varphi_i)$$

$$i = 1, \dots, M$$

Poniamo ora

$$m_{ij} = \int_{\Omega} \varphi_j \varphi_i \, dx, \quad M = (m_{ij}) \text{ matrice di masse}$$

$$a_{ij} = a(\varphi_j, \varphi_i), \quad A = (a_{ij}) \text{ matrice di stiffness}$$

$$f_i = (f, \varphi_i)$$

abbiamo il problema in forma matriciale

$$M u'(t) + A u(t) = b(t)$$

$$\text{con } u = (u_1, \dots, u_n), \quad b = (f_1, \dots, f_n).$$

È facile dimostrare che la matrice M è definita positiva e dunque invertibile (verifica per esercizio) abbiamo

$$u'(t) = -M^{-1} A u(t) + M^{-1} b(t)$$

Dunque un sistema di EDO che può essere risolto con metodi che abbiamo visto nelle prime parti del corso, ad esempio con il metodo di Crank-Nicolson che garantisce un'accuratezza di ordine 2 nel tempo.

$$M \frac{u^{n+1} - u^n}{\Delta t} + \frac{1}{2} A (u^{n+1} + u^n) = \frac{1}{2} (b^{n+1} + b^n)$$

$$\left(\frac{M}{\Delta t} + \frac{1}{2} A \right) u^{n+1} = \left(\frac{M}{\Delta t} - \frac{1}{2} A \right) u^n + \frac{1}{2} (b^{n+1} + b^n)$$

$$K u^{n+1} = G$$

con $K = M/\Delta t + \frac{1}{2} A$ matrice associata al nodo

e $G = \left(\frac{M}{\Delta t} - \frac{1}{2} A \right) u^n + \frac{1}{2} (b^{n+1} + b^n)$ vettore dei

termini noti.

Si noti che se la mesh è fissa nel tempo allora anche K è una matrice costante e predefinita.

Dunque può essere fattorizzata all'inizio del processo.

Inoltre sia M che A sono simmetriche, dunque anche

K è simmetrica e può essere fattorizzata con Cholesky

nella forma $K = H H^T$ con H triangolo superiore.

Ad ogni passo temporale si risolvono dunque due sistemi triangolari in M incognite ognuno dei quali richiede $M^2/2$ operazioni.

$$\begin{cases} H y = G \\ H^T u^{n+1} = y \end{cases}$$

OSS: Per quanto riguarda la stabilità è possibile estendere gli stessi risultati vlti. nel caso delle differenze finite. In particolare il metodo di Eulero esplicito e il metodo di Crank-Nicolson risultano incondizionatamente stabili, ossia stabili $\forall \Delta t$.

Nel caso di Eulero esplicito invece si ottiene una condizione del tipo $\Delta t \leq 2/\lambda_{\max}$, con λ_{\max} autovalore massimo della matrice di stiffness. In particolare si ha che $\lambda_{\max} \sim G/h^2$ da cui una condizione di stabilità analoga a quella vista con le differenze finite.