

▷ Lezione X

Equazioni alle derivate parziali

Equazioni alle derivate parziali

In questa parte consideriamo metodi numerici per risolvere equazioni differenziali alle derivate parziali con dati al bordo. Dato un dominio sufficientemente regolare Ω , ci focalizziamo su due importanti equazioni. La prima è l'equazione di Poisson che è data da

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega \subset \mathbb{R}^d \\ + \text{condizioni al bordo} &&& \text{su } \partial\Omega \end{aligned}$$

dove d è la dimensione spaziale e $\partial\Omega$ rappresenta il bordo di Ω , dove Δ è l'operatore di Laplace definito come $\sum_i \partial_{x_i^2}$, ad esempio in 2D abbiamo $\partial_{x^2} + \partial_{y^2}$. Per $d = 1$ la soluzione u dipende solo da una variabile e quindi l'equazione di Poisson viene semplicemente scritta come

$$\begin{aligned} -u'' &= f && \text{in } \Omega = (a, b) \\ + \text{condizioni al bordo} &&& \text{in } a \text{ e } b \end{aligned},$$

dove u'' può essere usato al posto di $\partial_{x^2}u$. Tale equazione è detta anche equazione ellittica. Una seconda equazione differenziale di cui studieremo la discretizzazione è l'equazione del calore, che, a differenza della precedente, dipende dal tempo ed è data da

$$\begin{aligned} \partial_t u - D\Delta u &= f && \text{in } \Omega \times (0, T] \\ + \text{condizioni al bordo} &&& \text{su } \partial\Omega \times (0, T], \\ + \text{condizioni iniziale} &&& \text{in } \Omega \end{aligned}$$

dove T è l'istante finale e le condizioni al bordo vengono assegnate per ogni istante di tempo in $(0, T]$, tale equazione è detta parabolica. Nel caso monodimensionale il problema viene scritto come

$$\begin{aligned} \partial_t u - D\partial_{x^2}u &= f && \text{in } \Omega = (a, b) \\ + \text{condizioni al bordo} &&& \text{in } a \text{ e } b \text{ per ogni } t \in (0, T]. \\ + \text{condizioni iniziale} &&& \text{in } \Omega \end{aligned}$$

Nota: la notazione $\partial_t \cdot$ corrisponde a $\frac{\partial \cdot}{\partial t}$.

Definizione della griglia

Dato un dominio monodimensionale $\Omega = (a, b)$ prima di procedere con la discretizzazione definiamo una *griglia computazionale* o *mesh*. Suddividiamo Ω in $N + 1$ sotto-intervalli, per semplicità uniformi, tali che $x_0 = a$ e $x_{N+1} = b$ e che il passo di griglia o la spaziatura sia

$$h = \frac{b - a}{N + 1}.$$

È possibile anche definire griglie con spaziatura non uniforme, ossia con $h_i = x_{i+1} - x_i$, tuttavia ci limitiamo al caso uniforme in cui il generico nodo è $x_i = ih + a$. Dettagli sulla definizione della griglia in domini 2D o 3D saranno forniti in seguito.

Metodo di Galerkin

Il metodo di Galerkin prende il nome dal matematico sovietico Boris Grigoryevich Galerkin ed è utilizzato per la risoluzione numerica di problemi differenziali. Il metodo si basa sulla formulazione debole del problema. Partiamo dal caso monodimensionale considerando l'equazione di Poisson con dati di Dirichlet nulli al bordo (per semplicità). Abbiamo, per $\Omega = (a, b)$

$$\begin{aligned} -u'' &= f & \text{in } \Omega \\ u(a) &= 0 \\ u(b) &= 0 \end{aligned}.$$

Per ricavare la formulazione debole si effettuano i seguenti tre passaggi: *i)* moltiplico tutti i termini per un'opportuna funzione test v , *ii)* integro sul dominio Ω e successivamente *iii)* integro per parti il termine con la derivata seconda su u ("scaricando" la derivata sulla funzione test). Otteniamo quindi

$$-\int_a^b u'' v dx = \int_a^b f v dx \quad \rightarrow \quad \int_a^b u' v' dx - [u' v]_a^b = \int_a^b f v dx.$$

A questo punto sfruttiamo le condizioni al bordo del problema richiedendo che anche v , come u , sia nulla in a e b e otteniamo la seguente formulazione debole associata al problema

$$\int_a^b u' v' dx = \int_a^b f v dx.$$

Essendo v arbitraria la precedente condizione deve essere valida per tutte le funzioni v nello spazio

$$V = \{v \in H^1(\Omega) : v(a) = 0 \text{ e } v(b) = 0\}.$$

Ricordiamo che $H^1(\Omega)$ indica lo spazio di Sobolev delle funzioni $v \in L^2(\Omega)$ e il cui gradiente, in questo caso la cui derivata essendo in 1D, appartiene anch'essa a $L^2(\Omega)$. Introducendo la forma bilineare a e il funzionale F definiti come

$$a(u, v) = \int_a^b u' v' dx \quad \text{e} \quad F(v) = \int_a^b f v dx$$

possiamo quindi scrivere la formulazione debole in maniera astratta. Il problema risulta ora: trovare $u \in V$ tale che

$$a(u, v) = F(v) \quad \forall v \in V.$$

È possibile utilizzare il lemma di Lax-Milgram per mostrare che il problema precedente è ben posto, ossia che ha una sola soluzione che dipende con continuità dai dati.

Lemma 10.1 - Lax-Milgram

Se V è uno spazio di Hilbert, a è una forma bilineare continua e coerciva, F è un funzionale lineare e continuo allora il problema della forma: trovare $u \in V$ tale che

$$a(u, v) = F(v) \quad \forall v \in V$$

ammette un'unica soluzione, inoltre sussiste la seguente stima

$$\|u\|_V \leq c \|F\|_*.$$

(dove $\|\cdot\|_*$ è un'opportuna norma).

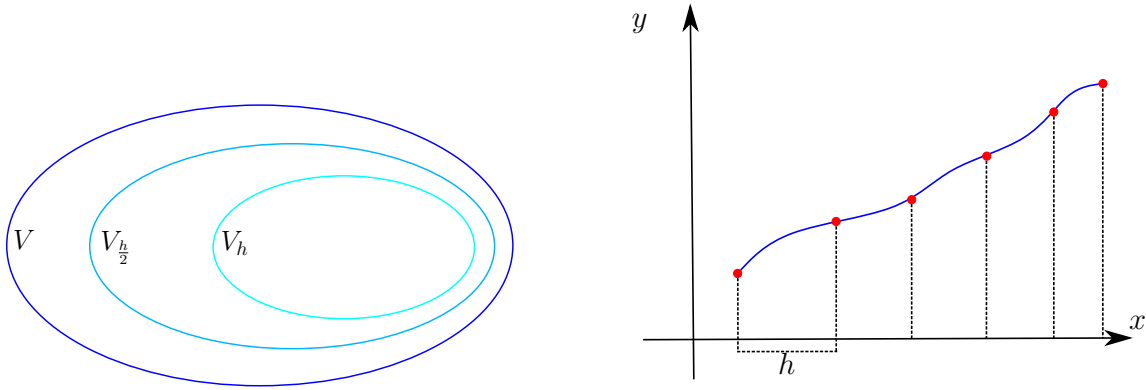
Ricordiamo per comodità le seguenti definizioni: La forma bilieare è *continua* se esiste $M > 0$ tale che

$$|a(u, v)| \leq M \|u\|_V \|v\|_V \quad \forall u, v \in V.$$

Inoltre è *coerciva* se esiste $\alpha > 0$ tale che

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

Il metodo di Galerkin costruisce un sottospazio V_h finito dimensionale di V tale per cui per $h \rightarrow 0$ allora $V_h \rightarrow V$, dove appunto h è il passo di griglia o comunque un parametro legato alla discretizzazione.



Ad esempio, V_h può essere scelto come lo spazio delle funzioni lineari a tratti sulla griglia definita precedentemente. Nota: cosa significa *spazio finito dimensionale*? Pensiamo agli spazi vettoriali \mathbb{R}^n , come $\mathbb{R}^2, \mathbb{R}^3$: essi sono n -dimensionali, in quanto tutti i loro elementi possono essere espressi con una base formata da n vettori. Lo stesso concetto, come vedremo, si può applicare agli spazi di funzioni.

Scelto quindi uno spazio V_h , introduciamo ora la formulazione di Galerkin del problema visto precedentemente, che consiste nel trovare $u_h \in V_h$ tale che

$$a(u_h, v_h) = F(v_h) \quad \forall v_h \in V_h \quad (1.1)$$

dove la forma bilineare a e il funzionale F non sono variate ma solo gli argomenti su cui operano, infatti abbiamo in questo caso

$$a(u_h, v_h) = \int_a^b u_h' v_h' dx \quad \text{e} \quad F(v_h) = \int_a^b f v_h dx.$$

Essendo V_h un sottospazio finito dimensionale di uno spazio di Hilbert, risulta anch'esso uno spazio di Hilbert ed è quindi possibile utilizzare il lemma di Lax-Milgram anche al problema discreto ed ottenere l'esistenza e l'unicità di u_h . Infatti consideriamo la stessa forma bilineare e lo stesso funzionale del caso continuo, che mantengono le proprietà essendo $V_h \subset V$. Nonostante V_h sia finito dimensionale la formulazione di Galerkin (1.1) deve essere ancora verificata per ogni v_h , ossia per infinite funzioni test. Dato che V_h è finito dimensionale possiamo introdurre delle funzioni $\{\phi_i\}$ che formino la base di V_h , con $i = 1, \dots, N_h$, dove $N_h = \dim(V_h)$. Essendo v_h e u_h elementi di V_h allora possiamo espanderli sulla base introdotta e scriverli come

$$v_h(x) = \sum_{i=1}^{N_h} \alpha_i \phi_i(x) \quad \text{e} \quad u_h(x) = \sum_{j=1}^{N_h} u_j \phi_j(x),$$

dove $\alpha_i \in \mathbb{R}$ e $u_j \in \mathbb{R}$ per ogni i e $j = 1, \dots, N_h$. Sostituendo l'espressione di v_h nell'equazione vista precedentemente e sfruttando la linearità nel secondo argomento di a e la linearità di F otteniamo

$$\begin{aligned} a(u_h, v_h) = F(v_h) &\rightarrow a\left(u_h, \sum_{i=1}^{N_h} \alpha_i \phi_i\right) = F\left(\sum_{i=1}^{N_h} \alpha_i \phi_i\right) \\ \sum_{i=1}^{N_h} \alpha_i a(u_h, \phi_i) &= \sum_{i=1}^{N_h} \alpha_i F(\phi_i), \end{aligned}$$

ma, dato che tale relazione deve valere per tutti i $v_h \in V_h$, questo implica che deve valere

$$a(u_h, \phi_i) = F(\phi_i) \quad \forall i = 1, \dots, N_h.$$

Il sistema è ora formato da N_h equazioni e non più un numero infinito come visto precedentemente. Inoltre, come detto, anche la soluzione discreta u_h può essere scritta data la base $\{\phi_i\}$ nel seguente modo

$$u_h(x) = \sum_{j=1}^{N_h} u_j \phi_j(x)$$

dove u_j sono ora coefficienti da determinare. A questo punto infatti determinare la soluzione numerica u_h equivale a determinare i valori $\{u_j\}$, detti anche *gradi di libertà*, dato che le ϕ_j sono note (saranno definite più precisamente in seguito)

$$u_h(x) = \sum_{i=1}^{N_h} u_i \phi_i(x) \quad \xrightarrow{\text{è definita dal vettore}} \quad \mathbf{u} = [u_1, \dots, u_{N_h}]^\top.$$

Sostituiamo nell'espressione precedente l'espansione di u_h sulla base di V_h per ottenere

$$a(u_h, \phi_i) = F(\phi_i) \quad \forall i = 1, \dots, N_h \quad \rightarrow \quad a\left(\sum_{j=1}^{N_h} u_j \phi_j, \phi_i\right) = F(\phi_i) \quad \forall i = 1, \dots, N_h$$

e dalla bi-linearità della forma a otteniamo che

$$\sum_{j=1}^{N_h} u_j a(\phi_j, \phi_i) = F(\phi_i) \quad \forall i = 1, \dots, N_h.$$

A questo punto, possiamo introdurre la matrice $A \in \mathbb{R}^{N_h \times N_h}$ e il vettore $\mathbf{f} \in \mathbb{R}^{N_h}$ con elementi così definiti

$$\begin{aligned} A &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1,N_h} \\ a_{21} & a_{22} & \dots & a_{2,N_h} \\ \dots & & & \\ a_{N_h,1} & a_{N_h,2} & \dots & a_{N_h,N_h} \end{bmatrix} \quad \text{dove} \quad a_{ij} = a(\phi_j, \phi_i) \\ \mathbf{f} &= [f_1, f_2, \dots, f_{N_h}]^\top \quad \text{dove} \quad f_i = F(\phi_i). \end{aligned}$$

Il problema discreto precedente viene scritto ora come la risoluzione di un problema lineare, detto problema algebrico equivalente, ovvero trovare \mathbf{u} tale che

$$A\mathbf{u} = \mathbf{f}.$$

Una volta risolto il sistema lineare possiamo ricostruire la soluzione u_h sfruttando ancora le funzioni di base $\{\phi_i\}$ e la seguente relazione

$$u_h(x) = \sum_{i=1}^{N_h} u_i \phi_i(x).$$

Ricordiamo che la forma a è coerciva, ossia è tale che esiste $\alpha > 0$ che soddisfa

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

risulta inoltre continua, ovvero esiste $M > 0$ tale che

$$|a(u, v)| \leq M \|u\|_V \|v\|_V \quad \forall u, v \in V.$$

Abbiamo il seguente importante risultato come conseguenza del Lemma di Lax-Milgram.

Teorema 10.1

Il problema di Poisson approssimato con il metodo di Galerkin è ben posto, ovvero ammette una sola soluzione u_h che è stabile rispetto ai dati

$$\|u_h\|_V \leq \frac{1}{\alpha} \|F\|_{V'},$$

dove $\|\cdot\|_{V'}$ è un'opportuna norma.

Nel risultato precedente la costante α è la costante di coercività. Una importante proprietà dal punto di vista numerico è che se la forma bilineare a è simmetrica allora anche la matrice A risulta simmetrica infatti

$$a_{ij} = a(\phi_j, \phi_i) = a(\phi_i, \phi_j) = a_{ji}.$$

Inoltre se la forma a è coerciva allora la matrice A risulta definita positiva, cioè ha tutti gli autovalori positivi. Considerando la generica coppia (λ, \mathbf{v}) di autovalore e associato autovettore a norma unitaria abbiamo infatti

$$A\mathbf{v} = \lambda\mathbf{v} \quad \Rightarrow \quad \mathbf{v}^\top A\mathbf{v} = \lambda \|\mathbf{v}\|^2 \quad \Rightarrow \quad \mathbf{v}^\top A\mathbf{v} = \lambda.$$

Una generica funzione v può essere espressa sulla base come

$$v(x) = \sum_{i=1}^{N_h} v_i \phi_i(x)$$

dove v_i è la componente i -esima del vettore \mathbf{v} . Dato che a è coerciva abbiamo che per una qualsiasi funzione v , corrispondente al vettore \mathbf{v} ,

$$\mathbf{v}^\top A\mathbf{v} = a(v, v) \geq \alpha \|v\|_V^2 > 0 \quad \forall v \neq 0$$

implicando quindi che $\lambda > 0$, ovvero che A è definita positiva.

Riportiamo ora il seguente risultato.

Lemma 10.2 - ortogonalità di Galerkin

La soluzione del problema di Galerkin u_h è tale che soddisfa la seguente proprietà

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

dove u è la soluzione esatta.

Proof. Essendo V_h sotto-spazio di V (se $v_h \in V_h$ allora $v_h \in V$) allora vale la seguente relazione

$$a(u, v) = F(v) \quad \forall v \in V \quad \xrightarrow{V_h \subset V} \quad a(u, v_h) = F(v_h) \quad \forall v_h \in V_h.$$

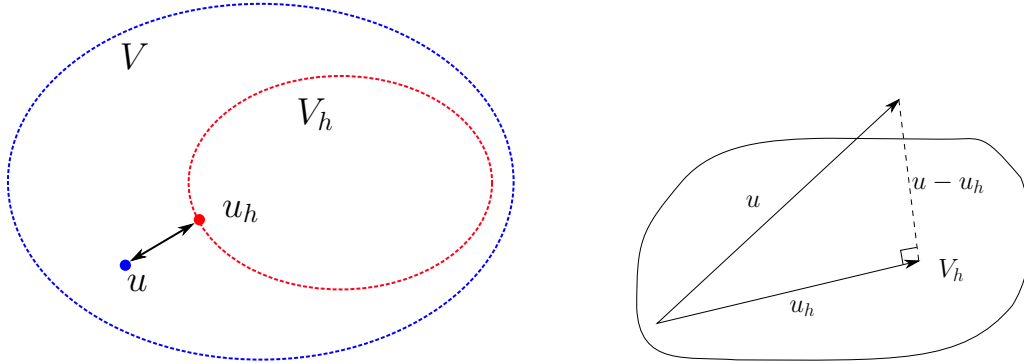
Considerando il problema di Galerkin e sfruttando la linearità di a otteniamo

$$\begin{cases} a(u, v_h) = F(v_h) \\ a(u_h, v_h) = F(v_h) \end{cases} \quad \forall v_h \in V_h \quad \xrightarrow{\text{sottraggo}} \quad a(u, v_h) - a(u_h, v_h) = 0 \quad \forall v_h \in V_h$$

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

□

Un'interpretazione grafica di tale lemma si trova nelle figure seguenti.



La forma a , se simmetrica, definisce un prodotto scalare che induce una norma data da

$$\|v\|_A = \sqrt{a(v, v)},$$

quindi avere che $a(u - u_h, v_h) = 0$ per tutti i v_h vuole dire che $u - u_h$, l'errore commesso, è ortogonale a tutti gli elementi di V_h rispetto al prodotto scalare indotto da a .

Abbiamo il seguente risultato di convergenza.

Teorema 10.2

La soluzione discreta u_h del problema di Galerkin converge alla soluzione u del problema in forma debole, ovvero

$$u_h \xrightarrow{h \rightarrow 0} u \quad \text{in } V.$$

Proof. Per dimostrare la convergenza, ossia che

$$\lim_{h \rightarrow 0} \|u - u_h\|_V = 0$$

posso utilizzare le proprietà della forma bilineare a , del funzionale F , e del problema di Galerkin. Infatti abbiamo

$$\begin{aligned} |a(u - u_h, u - u_h)| &= |a(u - u_h, u - u_h + v_h - v_h)| \leq |a(u - u_h, u - v_h)| + |a(u - u_h, v_h - u_h)| \\ &\leq M \|u - u_h\|_V \|u - v_h\|_V \quad \forall v_h \in V_h \end{aligned}$$

dove $a(u - u_h, v_h - u_h) = 0$ data l'ortogonalità di Galerkin e del fatto che $v_h - u_h \in V_h$, nell'ultimo passaggio abbiamo sfruttato la continuità della forma a . Usando ora la coercività di a abbiamo

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

otteniamo quindi la seguente stima

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \|u - v_h\|_V \quad \forall v_h \in V_h \quad \xrightarrow{\text{dalla generalità di } v_h} \quad \|u - u_h\|_V \leq C \inf_{v_h \in V_h} \|u - v_h\|_V.$$

Dato che il termine a destra va a zero per $V_h \rightarrow V$ (v_h approssima sempre meglio un elemento di V) allora otteniamo la convergenza del metodo di Galerkin. \square

Il metodo degli elementi finiti

La questione chiave ricade quindi nella scelta dello spazio discreto V_h perché appunto deve convergere a V per $h \rightarrow 0$. Per questo introduciamo il metodo agli elementi finiti, consideriamo anzitutto la versione lineare. In questo caso lo spazio V_h è definito come lo spazio delle funzioni lineari e continue a tratti sulla griglia definita dai nodi x_j , in formula abbiamo

$$V_h = X_h^1 = \{v \in C^0(a, b) : v|_{(x_j, x_{j+1})} \in \mathbb{P}^1(x_j, x_{j+1})\}$$

nel caso in cui dobbiamo considerare le condizioni al bordo direttamente nello spazio, come le condizioni di Dirichlet, introduciamo il sottospazio

$$V_{h,0} = \{v \in C^0(a, b) : v|_{(x_j, x_{j+1})} \in \mathbb{P}^1(x_j, x_{j+1}) \text{ } v(a) = 0, v(b) = 0\}$$

Per questo spazio introdotto abbiamo che per una qualunque funzione $u \in H_0^1((a, b))$ vale la seguente proprietà

$$\lim_{V_h \rightarrow V} \|u - u_h\|_V = 0, \quad \text{ovvero} \quad \lim_{h \rightarrow 0} \|u - u_h\|_V = 0,$$

che deriva appunto dalla proprietà di convergenza dell'interpolatore lineare. Ricordiamo che la norma di V (che nel nostro caso è H_0^1) è indotta dal suo prodotto scalare e, date due funzioni $v, w \in V$, abbiamo come

$$(v, w)_V = (v, w)_{L^2(\Omega)} + (\nabla v, \nabla w)_{L^2(\Omega)} \quad \text{che induce} \quad \|v\|_V = \sqrt{(v, v)_V}.$$

Questa definizione si basa sul prodotto scalare in L^2 che ricordiamo essere definito come

$$(v, w)_{L^2(\Omega)} = \int_{\Omega} vw \quad \text{che induce} \quad \|v\|_{L^2(\Omega)} = \sqrt{(v, v)_{L^2(\Omega)}}.$$

Abbiamo già mostrato che vale la seguente stima dell'errore per il metodo di Galerkin

$$\|u - u_h\|_V \leq C \inf_{v_h \in V_h} \|u - v_h\|_V$$

È possibile dimostrare che per per gli elementi finiti lineari, scegliendo $V_h = X_h^1 \subset V$, otteniamo la seguente stima dell'errore in termini di h

$$\|u - u_h\|_{L^2(\Omega)} \leq c_1 h^2 \quad \text{e anche che} \quad \|u - u_h\|_V \leq c_2 h$$

dove c_1 e c_2 sono due costanti. Notiamo che se misuriamo l'errore in norma V “perdiamo” un ordine di h rispetto all'errore misurato tramite la norma L^2 , infatti la norma in V richiede di considerare anche la norma del gradiente che converge come h e risulta quindi dominante per h piccoli.

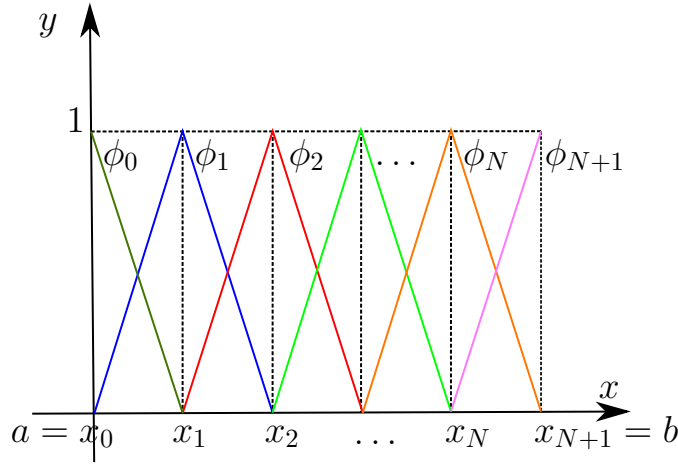
Abbiamo visto precedentemente la procedura per ricavare il sistema lineare associato, ovvero considerando $u_h, v_h \in V_h$ abbiamo

$$a(u_h, v_h) = F(v_h) \quad \xrightarrow{\text{sviluppando sulla base di } V_h} \quad \mathbf{A}\mathbf{u} = \mathbf{f} \quad \text{con} \quad u_h(x) = \sum_{j=1}^{N_h} u_j \phi_j(x)$$

dobbiamo ora scegliere una base per V_h , ovvero dobbiamo definire gli elementi ϕ_j nell'espressione precedente. Per il caso lineare, ovvero per $V_h = X_h^1$, una possibile scelta che possiamo fare sono le cosiddette funzioni a capanna (*hat functions*), tali che

$$\phi_i(x_j) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases},$$

dove gli x_i sono i nodi della griglia V_h . Graficamente la base di $V_h = X_h^1$, spazio di elementi finiti lineari, è rappresentabile come



Abbiamo quindi che $\{\phi_i\}$ composto dalle funzioni a capanna è una base di V_h . Eliminando ϕ_0 e ϕ_{N+1} ottengo invece una base per $V_{h,0}$, infatti il numero di basi è N_h che equivale alla dimensione dello spazio $V_{h,0}$. Inoltre le $\{\phi_i\}$ sono linearmente indipendenti tra loro. In dettaglio abbiamo che i vari ϕ_i sono definiti da

$$\phi_i(x) = \begin{cases} 0 & x < x_{i-1} \\ \frac{x - x_{i-1}}{x_i - x_{i-1}} & x_{i-1} \leq x \leq x_i \\ \frac{x_{i+1} - x}{x_{i+1} - x_i} & x_i \leq x \leq x_{i+1} \\ 0 & x > x_{i+1} \end{cases} \rightarrow \phi'_i(x) = \begin{cases} 0 & x < x_{i-1} \\ \frac{1}{x_i - x_{i-1}} & x_{i-1} \leq x \leq x_i \\ -\frac{1}{x_{i+1} - x_i} & x_i \leq x \leq x_{i+1} \\ 0 & x > x_{i+1} \end{cases}.$$



Ricordiamo che $x_{i+1} - x_i = h$ nel caso di griglia uniforme, per cui le espressioni precedenti si semplificano in

$$\phi_i(x) = \begin{cases} 0 & x < x_{i-1} \\ \frac{x - x_{i-1}}{h} & x_{i-1} \leq x \leq x_i \\ \frac{x_{i+1} - x}{h} & x_i \leq x \leq x_{i+1} \\ 0 & x > x_{i+1} \end{cases} \rightarrow \phi'_i(x) = \begin{cases} 0 & x < x_{i-1} \\ \frac{1}{h} & x_{i-1} \leq x \leq x_i \\ -\frac{1}{h} & x_i \leq x \leq x_{i+1} \\ 0 & x > x_{i+1} \end{cases}.$$

Una volta risolto il problema $Au = f$, la soluzione u_h viene quindi rappresentata come

$$u_h(x) = \sum_{j=1}^{N_h} u_j \phi_j(x) \quad \xrightarrow{\text{graficamente risulta}} \quad \begin{array}{c} v_h \\ \text{---} \\ x_0 \quad x_1 \quad x_{i-1} \quad x_i \quad x_{i+1} \quad x_{N+1} \end{array}$$

Vediamo quindi come calcolare gli elementi della matrice A , dalla formulazione precedente abbiamo

$$a_{ij} = a(\phi_j, \phi_i) \quad \text{con} \quad a(u, v) = \int_a^b u' v' dx \quad \Rightarrow \quad a_{ij} = \int_a^b \phi_j' \phi_i' dx$$

data la definizione delle funzioni di base ϕ_i e ϕ_j , l'ultima espressione risulta non nulla solamente per $i-1 \leq j \leq i+1$ ottenendo quindi, nel caso semplificato in cui la spaziatura della griglia è uniforme ovvero h costante,

$$a_{ii} = \int_{x_{i-1}}^{x_{i+1}} \phi_i' \phi_i' dx = \int_{x_{i-1}}^{x_i} \phi_i' \phi_i' dx + \int_{x_i}^{x_{i+1}} \phi_i' \phi_i' dx = \int_{x_{i-1}}^{x_i} \frac{1}{h^2} dx + \int_{x_i}^{x_{i+1}} \left(-\frac{1}{h}\right)^2 dx = \frac{2}{h}$$

gli altri due termini risultano

$$\begin{aligned} a_{i,i+1} &= \int_{x_{i-1}}^{x_{i+1}} \phi_i' \phi_{i+1}' dx = \int_{x_i}^{x_{i+1}} \phi_i' \phi_{i+1}' dx = \int_{x_i}^{x_{i+1}} \left(-\frac{1}{h}\right) \frac{1}{h} dx = -\frac{1}{h} \\ a_{i,i-1} &= \int_{x_{i-1}}^{x_{i+1}} \phi_i' \phi_{i-1}' dx = \int_{x_{i-1}}^{x_i} \phi_i' \phi_{i-1}' dx = \int_{x_{i-1}}^{x_i} \frac{1}{h} \left(-\frac{1}{h}\right) dx = -\frac{1}{h} \end{aligned}$$

la matrice A , detta di rigidità, risulta quindi sparsa e in particolare tridiagonale

$$A = \frac{1}{h} \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & \dots & 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

Inoltre, A è simmetrica e definita positiva, con condizionamento dato da $\text{cond}(A) \approx \mathcal{O}(h^{-2})$.

Per quanto riguarda il termine noto, il generico elemento f_i di \mathbf{f} si ottiene calcolando l'integrale

$$f_i = \int_{x_{i-1}}^{x_{i+1}} f(x) \phi_i dx.$$

Nel caso di $f(x) = c$ ossia costante otteniamo in particolare che $f_i = h$ per tutti gli $i = 1, \dots, N$.

Condizioni al bordo di Dirichlet non omogenee

Consideriamo il caso più generale, un problema con dati non omogenei di Dirichlet dato da

$$\begin{cases} -u'' = f & \text{in } \Omega = (a, b) \\ u(a) = \alpha \\ u(b) = \beta \end{cases}$$

Vogliamo quindi trovare l'approssimazione numerica $u_h \in V_h$ di $u \in V$ tale che $u_h(a) = \alpha$ e $u_h(b) = \beta$ e che

$$a(u_h, v_h) = F(v_h) \quad \forall v_h \in V_{h,0}.$$

Volendo avere sia la soluzione u_h che la funzione test v_h nello stesso spazio $V_{h,0}$ possiamo procedere nel seguente modo: la soluzione numerica u_h su tutti i nodi della griglia può essere scritta sfruttando una scomposizione, come

$$u_h(x) = \sum_{j=0}^{N+1} u_j \phi_j(x) = \sum_{j=1}^N u_j \phi_j(x) + \alpha \phi_0(x) + \beta \phi_{N+1}(x),$$

dove abbiamo messo in evidenza il dato di Dirichlet al primo e all'ultimo nodo. I valori della matrice A sono calcolati come fatto in precedenza ed otteniamo che la prima, la generica e l'ultima riga del sistema lineare sono rispettivamente date da

$$\begin{aligned} \frac{2u_1 - u_2}{h} &= \int_{x_0}^{x_1} f(x) \phi_1(x) dx + \frac{\alpha}{h} \\ \frac{-u_{j-1} + 2u_j - u_{j+1}}{h} &= \int_{x_{j-1}}^{x_{j+1}} f(x) \phi_j(x) dx \\ \frac{-u_{N-1} + 2u_N}{h} &= \int_{x_N}^{x_{N+1}} f(x) \phi_N(x) dx + \frac{\beta}{h} \end{aligned}$$

dove $-\int_{x_0}^{x_1} \phi'_0 \phi'_1 \alpha = \frac{\alpha}{h}$ e $-\int_{x_N}^{x_{N+1}} \phi'_N \phi'_{N+1} \beta = \frac{\beta}{h}$. Dobbiamo quindi modificare opportunamente il termine noto, la struttura matrice A rimane inalterata.

Problemi ellittici con coefficienti

Il problema di Poisson discusso finora può essere generalizzato con l'introduzione di un coefficiente γ che per semplicità assumiamo essere C^1 :

$$\begin{cases} -(\gamma u')' = f & \text{in } \Omega = (a, b) \\ u(a) = 0 \\ u(b) = 0 \end{cases}$$

Nel caso in cui il coefficiente di γ è costante l'equazione può essere riscritta semplicemente come

$$-\gamma u'' = f$$

e la sua discretizzazione porta semplicemente ad una matrice A simile a quella presentata in precedenza, premoltiplicata per γ . Se γ è variabile invece otteniamo che $(\gamma u')' = \gamma' u' + \gamma u''$. Questo ha conseguenze nell'interpretazione delle condizioni al contorno come discusso nella sezione successiva.

Condizioni al bordo di Neumann

Analizziamo ora come approssimare il problema nel caso in cui consideriamo le condizioni di Neumann, ovvero il seguente problema

$$\begin{cases} -(\gamma u')' = f & \text{in } \Omega = (a, b) \\ u(a) = 0 \\ -\gamma u'(b) = g \end{cases}$$

L'equazione precedente è scritta in forma conservativa, infatti la quantità che viene conservata è il flusso totale definito da $-\gamma u'$. Consideriamo ad esempio il caso del flusso di calore definito,

secondo Fourier, dal prodotto $-k\partial_x T$, dove T è la temperatura e k la conducibilità termica. Se consideriamo due materiali in contatto, il primo isolante, il secondo più conduttivo, osserveremo all'interfaccia fra i due una variazione improvvisa della derivata della temperatura per garantire la continuità del flusso termico.

Tale scrittura è quindi da preferirsi, in particolare quando γ è una funzione della variabile spaziale. Considerazione che ci porta ad impostare il dato di bordo di Neumann sul flusso totale e non solamente sulla derivata di u .

Deriviamo innanzitutto la forma debole del problema, avendo un dato di Dirichlet in a scelgo lo spazio (nel continuo) come

$$V = \{v \in H^1(\Omega) : v(a) = 0\}$$

La forma debole è quindi: sia v funzione test, moltiplichiamo il precedente problema per v e successivamente integriamo sul dominio Ω , abbiamo

$$-\int_a^b (\gamma u')' v dx = \int_a^b f v dx$$

integrando per parti risulta

$$\int_a^b \gamma u' v' dx - [\gamma u' v]_a^b = \int_a^b f v dx.$$

Dato che abbiamo scelto $v \in V$, quindi nullo al bordo “sinistro”, otteniamo che i termini di bordo diventano

$$-[\gamma u' v]_a^b = \gamma u'(a)v(a) - \gamma u'(b)v(b) = -\gamma u'(b)v(b) = gv(b),$$

la formulazione debole del problema differenziale con condizione di Neumann risulta: trovare $u \in V$ tale che il seguente problema è soddisfatto

$$\int_a^b \gamma u' v' dx = \int_a^b f v dx - gv(b) \quad \forall v \in V.$$

Per riscrivere il problema in forma astratta, possiamo introdurre, come nei casi precedenti, la forma bilineare a e il funzionale F come

$$a(u, v) = \int_a^b \gamma u' v' dx \quad \text{e} \quad F(v) = \int_a^b f v dx - gv(b),$$

il problema viene quindi riscritto: trovare $u \in V$ tale che

$$a(u, v) = F(v) \quad \forall v \in V.$$

Sfruttando ancora una volta il lemma di Lax-Millgram possiamo concludere che il problema precedente è ben posto.

L'approssimazione numerica del problema segue quanto fatto in precedenza, dobbiamo tuttavia prestare particolare attenzione a come poter gestire il dato di Neumann. Lo spazio discreto V_h risulta ora un sottospazio di V , una sua funzione ha quindi valore nullo in a ma nulla è prescritto a priori in b . Come conseguenza, la dimensione dello spazio è ora $N + 1$ in quanto abbiamo un grado di libertà anche in corrispondenza dell'ultimo nodo. Il termine $-gv(b)$, associato al nodo $N + 1$, viene approssimato sfruttando l'espansione sulla base e risulta quindi

$$-gv(b) \approx -g\phi_{N+1}(b) = -g$$

avendo sfruttato il fatto che il nodo $(N + 1)$ -esimo è appunto b e quindi $\phi_{N+1}(b) = 1$. Come anticipato, la soluzione in b non è nota e quindi dobbiamo includere fra le incognite anche u_{N+1} .

Supponendo per semplicità che γ sia costante, le componenti dell'ultima riga della matrice risultano

$$\begin{aligned} a_{N+1,N+1} &= \gamma \int_{x_N}^{x_{N+1}} \phi'_{N+1} \phi'_{N+1} dx = \gamma \int_{x_N}^{x_{N+1}} \frac{1}{h^2} dx = \frac{\gamma}{h} \\ a_{N+1,N} &= \gamma \int_{x_N}^{x_{N+1}} \phi'_{N+1} \phi'_N dx = \gamma \int_{x_N}^{x_{N+1}} \frac{1}{h} \left(-\frac{1}{h} \right) dx = -\frac{\gamma}{h} \end{aligned}$$

L'ultima equazione che lega u_{N+1} con u_N è quindi della forma

$$\gamma \frac{-u_N + u_{N+1}}{h} = \int_{x_N}^{x_{N+1}} f dx - g.$$

Quindi, supponendo per semplicità che γ sia costante, la matrice A ha la seguente struttura: ha dimensione $N+1 \times N+1$, è tridiagonale e data da

$$A = \frac{\gamma}{h} \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & \dots & 0 & 0 & -1 & 2 & -1 \\ 0 & \dots & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

Notiamo l'ultima riga che è differente da quello che abbiamo visto prima e risulta dalla condizione al bordo di tipo Neumann.

Problemi più generali

Consideriamo ora il caso in cui abbiamo la seguente equazione

$$\begin{cases} (-\gamma u' + \delta u)' + \nu u = f & \text{in } \Omega = (a, b) \\ u(a) = 0 \\ -\gamma u'(b) + \delta u(b) = g. \end{cases}$$

Si tratta ancora di un problema differenziale di tipo ellittico, ma più generale di quello che abbiamo trattato finora. In questo caso il flusso totale è definito da $-\gamma u' + \delta u$ che coinvolge un termine diffusivo, il primo, e di trasporto, il secondo. δ è una velocità di trasporto che può essere data o calcolata da un altro modello. Infine il termine di ordine zero che dipende dal coefficiente ν modella un termine di reazione. Al bordo b imponiamo una condizione sul flusso totale, ottenendo una condizione più complessa, che lega, al bordo, la derivata $u'(b)$, il valore della funzione $u(b)$ e un valore dato. Questa condizione può essere, in molti casi, la più realistica. Nel caso di coefficienti costanti l'equazione precedente può essere scritta anche come

$$(-\gamma u' + \delta u)' + \nu u = -\gamma u'' + \delta u' + \nu u = f,$$

da cui risultano più evidenti i termini di differente ordine. Per determinare la formulazione agli elementi finiti si procede in modo analogo a quanto visto precedentemente, prima nel continuo considerando lo spazio $V = \{v \in H^1(\Omega) : v(a) = 0\}$ e una generica funzione test $v \in V$ ottenendo

$$\begin{aligned} \int_a^b (-\gamma u' + \delta u)' v dx + \int_a^b \nu u v dx &= \int_a^b \gamma u' v' dx - [\gamma u' v]_a^b - \int_a^b \delta u v' dx + [\delta u v]_a^b + \int_a^b \nu u v dx = \\ &= \int_a^b \gamma u' v' dx - \int_a^b \delta u v' dx + [-\gamma u'(b) + \delta u(b)] v(b) + \int_a^b \nu u v dx = \\ &= \int_a^b \gamma u' v' dx - \int_a^b \delta u v' dx + g v(b) + \int_a^b \nu u v dx. \end{aligned}$$

Possiamo introdurre una forma bilineare per ogni integrale presente, ovvero le seguenti

$$a(u, v) = \int_a^b \gamma u' v' dx \quad b(u, v) = - \int_a^b \delta u v' dx \quad c(u, v) = \int_a^b \nu u v dx$$

e un funzionale lineare dato da

$$F(v) = \int_a^b f v dx - g v(b).$$

Il problema in forma debole risulta quindi trovare $u \in V$ tale che

$$a(u, v) + b(u, v) + c(u, v) = F(v) \quad \forall v \in V$$

la cui buona posizione può essere dimostrata verificando le ipotesi del teorema di Lax-Millgram ora sulla forma bilineare $a + b + c$.

Possiamo ora introdurre la fomulazione discreta del problema usando gli elementi finiti lineari, per semplicità, su una griglia uniforme e a coefficienti costanti. Associamo le matrici A , B e C alle tre forme bilineari viste in precedenza a , b e c , rispettivamente. La matrice A è già stata introdotta precedentemente e risulta pari a

$$A = \frac{\gamma}{h} \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & \dots & 0 & 0 & -1 & 2 & -1 \\ 0 & \dots & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

Per la matrice B possiamo calcolare i coefficienti $b_{ij} = -\delta \int \phi_i' \phi_j$ che risultano essere dati da

$$\begin{aligned} b_{i,i} &= -\delta \int_{x_{i-1}}^{x_{i+1}} \phi_i \phi_i' dx = -\delta \int_{x_{i-1}}^{x_i} \phi_i \phi_i' dx - \delta \int_{x_i}^{x_{i+1}} \phi_i \phi_i' dx = \\ &= -\frac{\delta}{h} \int_{x_{i-1}}^{x_i} \phi_i dx + \frac{\delta}{h} \int_{x_i}^{x_{i+1}} \phi_i dx = -\frac{\delta}{h} \frac{h}{2} + \frac{\delta}{h} \frac{h}{2} = 0 \\ b_{i,i+1} &= -\delta \int_{x_{i-1}}^{x_{i+1}} \phi_{i+1} \phi_i' dx = -\delta \int_{x_i}^{x_{i+1}} \phi_{i+1} \phi_i' dx = \frac{\delta}{h} \int_{x_i}^{x_{i+1}} \phi_{i+1} dx = \frac{\delta}{h} \frac{h}{2} = \frac{\delta}{2} \\ b_{i,i-1} &= -\delta \int_{x_{i-1}}^{x_{i+1}} \phi_{i-1} \phi_i' dx = -\delta \int_{x_{i-1}}^{x_i} \phi_{i-1} \phi_i' dx = -\frac{\delta}{h} \int_{x_{i-1}}^{x_i} \phi_i dx = -\frac{\delta}{h} \frac{h}{2} = -\frac{\delta}{2} \end{aligned}$$

Dato che $v(b) \neq 0$ l'ultima riga, corrispondente al nodo $N+1$, risulta diversa,

$$\begin{aligned} b_{N+1,N+1} &= -\delta \int_{x_N}^{x_{N+1}} \phi_{N+1} \phi_{N+1}' dx = -\frac{\delta}{h} \int_{x_N}^{x_{N+1}} \phi_{N+1} dx = -\frac{\delta}{h} \frac{h}{2} = -\frac{\delta}{2} \\ b_{N+1,N} &= -\delta \int_{x_N}^{x_{N+1}} \phi_N \phi_{N+1}' dx = -\frac{\delta}{h} \int_{x_N}^{x_{N+1}} \phi_N dx = -\frac{\delta}{h} \frac{h}{2} = -\frac{\delta}{2} \end{aligned}$$

La matrice associata risulta quindi la seguente

$$B = \frac{\delta}{2} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & 1 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & \dots & 0 & 0 & -1 & 0 & 1 \\ 0 & \dots & 0 & 0 & 0 & -1 & -1 \end{bmatrix}$$

Infine per il termine di reazione i coefficienti $c_{ij} = \nu \int \phi_i \phi_j$ della matrice C risultano dati da

$$\begin{aligned}
c_{i,i} &= \nu \int_{x_{i-1}}^{x_{i+1}} \phi_i \phi_i dx = \nu \int_{x_{i-1}}^{x_i} \phi_i^2 dx + \nu \int_{x_i}^{x_{i+1}} \phi_i^2 dx = \\
&= \frac{\nu}{h^2} \left[\int_{x_{i-1}}^{x_i} (x - x_{i-1})^2 dx + \int_{x_i}^{x_{i+1}} (x_{i+1} - x)^2 dx \right] = \frac{\nu}{3h^2} [(x_i - x_{i-1})^3 + (x_{i+1} - x_i)^3] = \frac{2h\nu}{3} \\
c_{i,i+1} &= \nu \int_{x_{i-1}}^{x_{i+1}} \phi_{i+1} \phi_i dx = \frac{\nu}{h^2} \int_{x_i}^{x_{i+1}} (x_{i+1} - x)(x - x_i) dx = \\
&= \frac{\nu}{h^2} \int_{x_i}^{x_{i+1}} (x_{i+1}x - x_{i+1}x_i - x^2 + xx_i) dx = \frac{\nu}{h^2} \left[(x_{i+1} + x_i) \frac{x^2}{2} - x_{i+1}x_i x - \frac{x^3}{3} \right]_{x_i}^{x_{i+1}} = \\
&= \frac{\nu}{h^2} \left((x_{i+1} + x_i) \frac{x_{i+1}^2}{2} - x_{i+1}x_i x_{i+1} - \frac{x_{i+1}^3}{3} - (x_{i+1} + x_i) \frac{x_i^2}{2} + x_{i+1}x_i x_i + \frac{x_i^3}{3} \right) = \\
&= \frac{\nu}{h^2} \left((x_{i+1} + x_i) \frac{x_{i+1}^2}{2} - x_{i+1}x_i x_{i+1} - \frac{x_{i+1}^3}{3} - (x_{i+1} + x_i) \frac{x_i^2}{2} + x_{i+1}x_i x_i + \frac{x_i^3}{3} \right) = \\
&= \frac{\nu}{6h^2} (x_{i+1}^3 - x_i^3 - 3x_{i+1}^2 x_i + 3x_{i+1} x_i^2) = \frac{\nu}{6h^2} (x_{i+1} - x_i)^3 = \frac{h\nu}{6} \\
c_{i,i-1} &= \nu \int_{x_{i-1}}^{x_{i+1}} \phi_{i-1} \phi_i dx = \frac{\nu}{h^2} \int_{x_{i-1}}^{x_i} (x_{i-1} - x)(x - x_i) dx = \\
&= \frac{\nu}{h^2} \int_{x_{i-1}}^{x_i} (x_{i-1}x - x^2 - x_i x_{i-1} + xx_i) dx = \frac{\nu}{h^2} \left[(x_i + x_{i-1}) \frac{x^2}{2} - x_i x_{i-1} x - \frac{x^3}{3} \right]_{x_{i-1}}^{x_i} = \\
&= \frac{\nu}{h^2} \left((x_i + x_{i-1}) \frac{x_i^2}{2} - x_i x_{i-1} x_i - \frac{x_i^3}{3} - (x_i + x_{i-1}) \frac{x_{i-1}^2}{2} + x_i x_{i-1} x_{i-1} + \frac{x_{i-1}^3}{3} \right) = \\
&= \frac{\nu}{6h^2} (x_i^3 - x_{i-1}^3 - 3x_i^2 x_{i-1} + 3x_i x_{i-1}^2) = \frac{\nu}{6h^2} (x_i - x_{i-1})^3 = \frac{h\nu}{6}
\end{aligned}$$

Data la condizione al bordo sul flusso totale alla fine dell'intervallo gli ultimi due termini sono dati da

$$\begin{aligned}
c_{N+1,N+1} &= \nu \int_{x_N}^{x_{N+1}} \phi_{N+1} \phi_{N+1} dx = \nu \int_{x_N}^{x_{N+1}} \phi_{N+1}^2 dx = \frac{\nu}{h^2} \int_{x_N}^{x_{N+1}} (x - x_N)^2 dx = \\
&= \frac{\nu}{3h^2} (x_{N+1} - x_N)^3 = \frac{h\nu}{3} \\
c_{N+1,N} &= \nu \int_{x_N}^{x_{N+1}} \phi_N \phi_{N+1} dx = \nu \frac{\nu}{h^2} \int_{x_N}^{x_{N+1}} (x_N - x)(x - x_{N+1}) dx = \\
&= \frac{\nu}{h^2} \int_{x_N}^{x_{N+1}} (x_N x - x^2 - x_{N+1} x_N + x x_{N+1}) dx = \frac{\nu}{h^2} \left[(x_{N+1} + x_N) \frac{x^2}{2} - x_{N+1} x_N x - \frac{x^3}{3} \right]_{x_N}^{x_{N+1}} = \\
&= \frac{\nu}{h^2} \left((x_{N+1} + x_N) \frac{x_{N+1}^2}{2} - x_{N+1} x_N x_{N+1} - \frac{x_{N+1}^3}{3} - (x_{N+1} + x_N) \frac{x_N^2}{2} + x_{N+1} x_N x_N + \frac{x_N^3}{3} \right) = \\
&= \frac{\nu}{6h^2} (x_{N+1}^3 - x_N^3 - 3x_{N+1}^2 x_N + 3x_{N+1} x_N^2) = \frac{\nu}{6h^2} (x_{N+1} - x_N)^3 = \frac{h\nu}{6}
\end{aligned}$$

Otteniamo che quindi la matrice è data dalla seguente struttura

$$C = h\nu \begin{bmatrix} \frac{2}{3} & \frac{1}{6} & 0 & 0 & 0 & \dots & 0 \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & \dots & 0 & 0 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ 0 & \dots & 0 & 0 & 0 & \frac{1}{6} & \frac{2}{3} \end{bmatrix}$$

Notiamo che le tre matrici scalano diversamente rispetto ad h , infatti corrispondono a termini di diverso ordine. Il problema algebrico associato risulta quindi scritto nel seguente modo

$$(A + B + C)u = f.$$

Il caso multi-dimensionale

Nel caso multi-dimensionale l'approccio risulta simile, infatti dobbiamo ricavare la formulazione debole del problema e costruire le basi ϕ_i ora su una griglia computazionale multi-dimensionale.

Analizziamo in particolare gli aspetti distintivi del caso multi-dimensionale rispetto a quello monodimensionale.

Formulazione debole multidimensionale

Sia dato un dominio regolare $\Omega \subset \mathbb{R}^d$ con bordo $\partial\Omega$, supponiamo che possiamo dividere $\partial\Omega$ in due porzioni disgiunte $\partial_D\Omega$ e $\partial_N\Omega$. Partendo dal problema

$$\begin{aligned} -\Delta u &= f & \text{in } \Omega \\ u &= 0 & \text{su } \partial_D\Omega \\ \nabla u \cdot \mathbf{n} &= g & \text{su } \partial_N\Omega \end{aligned}$$

Notiamo innanzitutto che la condizione di Neumann imposta sulla porzione di bordo $\partial_N\Omega$ è una condizione che prescrive il valore della derivata *in direzione normale al bordo*, infatti \mathbf{n} indica la normale unitaria uscente.

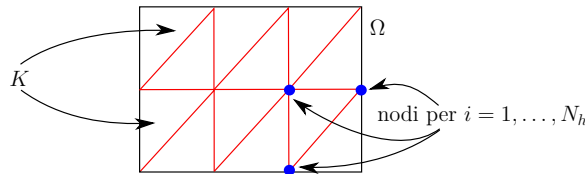
Moltiplichiamo l'equazione per una funzione test v e integriamo sul dominio:

$$\begin{aligned} \int_{\Omega} -\Delta u v &= \int_{\Omega} f v \\ \int_{\Omega} -\nabla \cdot (\nabla u v) + \int_{\Omega} \nabla u \cdot \nabla v &= \int_{\Omega} f v \\ \int_{\Omega} \nabla u \cdot \nabla v &= \int_{\Omega} f v + \int_{\partial\Omega} \nabla u \cdot \mathbf{n} v \\ \int_{\Omega} \nabla u \cdot \nabla v &= \int_{\Omega} f v + \int_{\partial_N\Omega} g v \end{aligned}$$

Notiamo che *i)* nel caso multidimensionale la derivata spaziale è sostituita dal gradiente, per cui dobbiamo calcolare il prodotto scalare fra i gradienti di u e v e *ii)* sulla porzione di bordo di Neumann sostituiamo la condizione corrispondente, mentre sul bordo di Dirichlet l'integrale di bordo si annulla scegliendo funzioni test v che siano nulle su tale bordo.

Costruzione e gestione della griglia

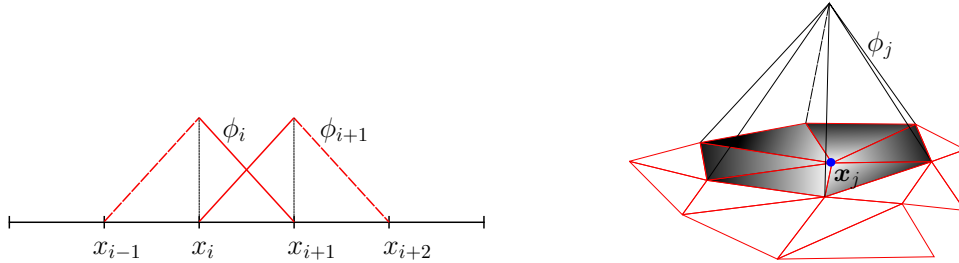
Nel caso 2D la griglia può suddividere il dominio in triangoli, come mostrato in figura



Nel caso tridimensionale, analogamente, un dominio poliedrico può essere suddiviso in tetraedri. Il metodo degli elementi finiti può però essere applicato anche a griglie formate da quadrati/rettangoli o esaedri in 3D.

Le basi ϕ_i vengono scelte in modo tale da essere diverse da zero solo in una piccola parte del dominio, su pochi elementi, così da ottenere una matrice A risultante dalla discretizzazione sparsa

e quindi che richiede poche risorse per poter essere memorizzata. Le funzioni di base lineari, in 2D, sono analoghe a quelle 1D come mostrato in figura



La principale differenza e difficoltà rispetto al caso risiede nella numerazione non più ovvia dei nodi e degli elementi. Se infatti le funzioni di base sono costruite in modo da avere valore 1 in uno specifico nodo il processo di *assemblaggio* della matrice si basa sul principio di additività dell'integrale. In altre parole, gli integrali alla base della forma bilineare e del funzionale F vengono calcolati come somma degli integrali sui singoli triangoli/tetraedri. Questi integrali saranno non nulli se la funzione di base coinvolta corrisponde a un nodo di tale triangolo. È quindi necessario conoscere questa mappa detta *connettività*: ad esempio, il triangolo numero 10 della griglia ha come vertici i nodi 12, 46, 47.

Una conseguenza di questa numerazione non "ovvia" è la struttura della matrice, che risulterà ancora sparsa ma non tridiagonale come nel caso 1D.

Un'ulteriore difficoltà è dovuta al fatto che, a differenza dei segmenti in 1D, tipicamente i triangoli sono tutti diversi fra loro. Per quanto sia semplice l'espressione delle funzioni di base, dipende quindi dalla posizione dei vertici. Tipicamente, le funzioni di base vengono definite sul cosiddetto *elemento di riferimento*, un triangolo di vertici $(0,0)$, $(1,0)$, $(0,1)$ e opportunamente ri-mappate sul singolo triangolo della griglia.

Infine, gli integrali possono essere calcolati con formule di quadratura abbastanza accurate (tipicamente Gaussiane).

Equazione del calore

Consideriamo ora un problema tempo dipendente, l'equazione del calore. Nel caso mono-dimensionale, dato $\Omega_T = \Omega \times (0, T)$ con $\Omega = (a, b)$, risulta

$$\begin{cases} \partial_t u - \partial_x(\gamma \partial_x u) = f & \text{in } \Omega_T \\ u(a, t) = \alpha(t) & \text{per } t \in [0, T) \\ u(b, t) = \beta(t) & \text{per } t \in [0, T) \\ u(x, t = 0) = u_0 & \text{in } \Omega \times \{0\} \end{cases}$$

La soluzione ora è una variabile che dipende sia dallo spazio che dal tempo $u(x, t) : \Omega_T \rightarrow \mathbb{R}$. Assumiamo per semplicità dati di Dirichlet nulli, $\alpha(t) = 0$, $\beta(t) = 0$, consideriamo ora lo spazio funzionale $V = H_0^1(\Omega)$ e prendiamo delle funzioni test $v \in V$ che sono indipendenti da t ; la forma debole del problema definita dall'equazione del calore è data da: trovare, per ogni $t \in (0, T)$, $u(x, t) \in V$ tale che

$$\int_a^b \partial_t u v dx + \int_a^b \gamma \partial_x u \partial_x v dx = \int_a^b f v dx \quad \forall v \in V.$$

La formulazione discreta del problema è derivata fissando il tempo ed operando prima in spazio. Scegliendo $V_h \subset V$ sotto-spazio finito dimensionale otteniamo che la semi-discretizzazione in spazio è: trovare, per ogni $t \in (0, T)$, $u_h(t) \in V_h$ tale che

$$\int_a^b \partial_t u_h v_h dx + \int_a^b \gamma \partial_x u_h \partial_x v_h dx = \int_a^b f v_h dx \quad \forall v_h \in V_h.$$

Introduciamo ora le forme a e m e il funzionale F definiti come

$$a(u_h, v_h) = \int_a^b \gamma \partial_x u_h \partial_x v_h \quad \text{e} \quad m(u_h, v_h) = \int_a^b u_h v_h \quad \text{e} \quad F(v_h) = \int_a^b f v_h dx$$

il precedente problema può quindi essere scritto come:

$$m(\partial_t u_h(t), v_h) + a(u_h, v_h) = F(v_h) \quad \forall v_h \in V_h.$$

Anche in questo caso consideriamo una base $\{\phi_j\}$ per V_h di funzioni linearmente indipendenti, consideriamo lo sviluppo di v_h e u_h sulla base $\{\phi_j\}$ di V_h otteniamo

$$v_h(x) = \sum_{i=1}^{N_h} \alpha_i \phi_i(x) \quad \text{e} \quad u_h(x, t) = \sum_{j=1}^{N_h} u_j(t) \phi_j(x)$$

Sfruttando la bi-linearità della forma a , il fatto che il dominio non dipende dal tempo e la linearità del funzionale F otteniamo la seguente espressione

$$\begin{aligned} \partial_t m \left(\sum_{j=1}^{N_h} u_j(t) \phi_j, \phi_i \right) + a \left(\sum_{j=1}^{N_h} u_j(t) \phi_j, \phi_i \right) &= F(\phi_i) \quad \forall i = 1, \dots, N_h \\ \sum_{j=1}^{N_h} \partial_t u_j(t) m(\phi_j, \phi_i) + \sum_{j=1}^{N_h} u_j(t) a(\phi_j, \phi_i) &= F(\phi_i) \quad \forall i = 1, \dots, N_h, \end{aligned}$$

introduciamo la matrice di rigidità A e di massa M date rispettivamente da

$$\begin{aligned} A \in \mathbb{R}^{N_h \times N_h} : \quad a_{ij} &= a(\phi_j, \phi_i) \\ M \in \mathbb{R}^{N_h \times N_h} : \quad m_{ij} &= m(\phi_j, \phi_i) \end{aligned}$$

Notiamo che la matrice di massa ha la stessa struttura della matrice introdotta per descrivere il termine di reazione in precedenza. Il problema semi-discreto risulta quindi di determinare il vettore tempo dipendente $\mathbf{u}(t)$ tale che

$$M d_t \mathbf{u}(t) + A \mathbf{u}(t) = \mathbf{f}.$$

Notiamo che abbiamo un sistema di equazioni differenziali ordinarie in tempo di dimensione N_h , che può essere scritto formalmente nel seguente modo

$$\begin{cases} d_t \mathbf{u}(t) = \tilde{\mathbf{f}}(t, \mathbf{u}) \\ \mathbf{u}(0) = \mathbf{u}_0 \end{cases}$$

dove appunto $\tilde{\mathbf{f}}(t, \mathbf{u}) = -M^{-1} A \mathbf{u}(t) + M^{-1} \mathbf{f}$. La soluzione iniziale \mathbf{u}_0 dipende solamente dallo spazio e può essere approssimata come

$$u_{0,h}(x) = \sum_{j=1}^{N_h} u_{0,j} \phi_j(x).$$

Introduciamo quindi ora una discretizzazione in tempo, ovvero dividiamo l'intervallo $(0, T]$ in N_t sotto-intervalli temporali tale che

$$t_0 = 0 \quad t_{N_t} = T \quad t^n = n \Delta t,$$

dove Δt è il passo di discretizzazione temporale definito come $\Delta t = T/N_t$. Possiamo quindi utilizzare gli schemi di avanzamento temporale già studiati, in particolare la derivata in tempo viene discretizzata come

$$d_t \mathbf{u} \approx \frac{\mathbf{u}(t^{n+1}) - \mathbf{u}(t^n)}{\Delta t} = \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t}.$$

Il θ -metodo può essere ora scritto come

$$M \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} + \theta A \mathbf{u}^{n+1} + (1 - \theta) A \mathbf{u}^n = \theta \mathbf{f}^{n+1} + (1 - \theta) \mathbf{f}^n$$

che è un sistema algebrico che dobbiamo risolvere ad ogni passo di discretizzazione temporale. Ad ogni istante temporale dobbiamo risolvere un problema del tipo

$$\left(\frac{M}{\Delta t} + \theta A \right) \mathbf{u}^{n+1} = \left(\frac{M}{\Delta t} - (1 - \theta) A \right) \mathbf{u}^n + \theta \mathbf{f}^{n+1} + (1 - \theta) \mathbf{f}^n.$$

Possiamo quindi fattorizzare una sola volta la matrice $M/\Delta t + \theta A$ con un costo computazionale pari ad $O(N_h^3)$, successivamente possiamo utilizzare la sostituzione in avanti e all'indietro per determinare il valore di \mathbf{u}^{n+1} , ad un costo $O(N_h^2)$. Notiamo che nel caso utilizzassimo elementi finiti lineari, le matrici A e M sono tridiagonali, possiamo quindi utilizzare l'algoritmo di Thomas per determinarne la soluzione ad un costo computazionale pari a $O(N_h)$.

Il θ -metodo, per una equazione differenziale ordinaria, è condizionatamente assolutamente stabile per $0 \leq \theta < 0.5$ e richiede di soddisfare una condizione per il passo temporale massimo, che nel caso di Eulero Esplicito è

$$\Delta t < \frac{2}{|\lambda|}$$

dove $\lambda < 0$ è il parametro associato al problema modello

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

Al contrario il θ -metodo risulta incondizionatamente assolutamente stabile per valori di $\theta \in [0.5, 1]$. Applicando questi concetti all'equazione del calore discretizzata sia in spazio che in tempo, la condizione di assoluta stabilità su Δt è ora scritta come

$$\Delta t \leq \frac{2}{\max |\lambda(M^{-1}A)|}$$

ossia dipende dagli autovalori di $M^{-1}A$. Dato che $\lambda \sim O(h^{-2})$, infatti gli elementi della matrice A scalano come $O(h^{-1})$ dato che la forma bilineare a contiene due derivate in spazio e un integrale mentre quelli della matrice M come $O(h)$ dato che m contiene solamente l'integrale, allora la condizione di assoluta stabilità, nel caso in cui $0 \leq \theta < 0.5$ diviene un legame tra Δt e h

$$\Delta t \leq ch^2.$$

Per questo motivo un metodo esplicito sarà "costretto" ad utilizzare un passo temporale molto fine, e questo criterio diventa *molto* più stringente per griglie raffinate.

Approfondimento - Approssimazione alle differenze finite

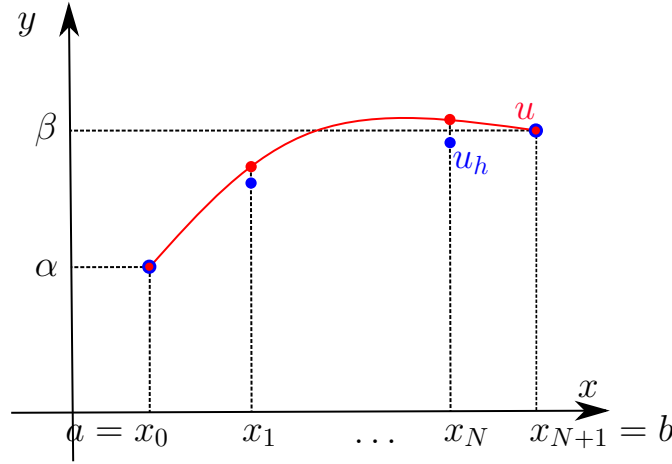
In questo corso approfondiremo un metodo basato sulla formulazione debole del problema, il metodo degli elementi finiti. Tuttavia, esistono altre possibili strategie tra cui il metodo delle differenze finite, di cui forniamo una breve descrizione come approfondimento.

Consideriamo il problema di Poisson monodimensionale, definito sul dominio $\Omega = (a, b)$

$$\begin{aligned} -u'' &= f \quad \text{in } \Omega \\ u(a) &= \alpha \\ u(b) &= \beta. \end{aligned}$$

Le condizioni imposte al bordo di Ω sono dette *di Dirichlet* e prescrivono il valore della soluzione u in a con valore $\alpha \in \mathbb{R}$ e in b con valore $\beta \in \mathbb{R}$. Una volta definita una griglia di passo h come

descritto nella sezione precedente, approssimiamo la soluzione esatta u nei nodi $\{x_j\}$ introducendo quindi $u_j \approx u(x_j) \in \mathbb{R}$.



Per approssimare il problema di Poisson, utilizziamo una differenza finita centrata per la derivata seconda, in particolare (si veda la lezione 7)

$$u''(x_j) \approx \frac{u_{j-1} - 2u_j + u_{j+1}}{h^2}$$

Abbiamo quindi le seguenti equazioni da risolvere per ogni nodo x_j

$$\text{per } x_1 \rightarrow -\frac{u_0 - 2u_1 + u_2}{h^2} = f(x_1)$$

$$\text{per } x_2 \rightarrow -\frac{u_1 - 2u_2 + u_3}{h^2} = f(x_2)$$

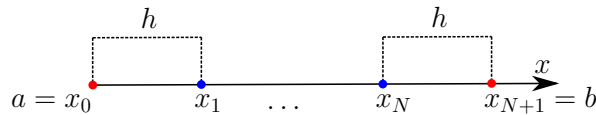
...

$$\text{per } x_j \rightarrow -\frac{u_{j-1} - 2u_j + u_{j+1}}{h^2} = f(x_j)$$

...

$$\text{per } x_N \rightarrow -\frac{u_{N-1} - 2u_N + u_{N+1}}{h^2} = f(x_N)$$

Indicando con rosso i nodi in cui la soluzione è nota e in blu in cui è incognita, osserviamo che la soluzione è incognita solo nei nodi interni, che sono N sui $N + 2$ nodi totali



Date le condizioni al bordo in $a = x_0$ e $b = x_{N+1}$, ovvero $u_0 = \alpha$ e $u_{N+1} = \beta$, allora la prima e l'ultima equazione possono essere riscritte nel seguente modo

$$\text{per } x_1 \rightarrow -\frac{u_0 - 2u_1 + u_2}{h^2} = f(x_1) \rightarrow \frac{2u_1 - u_2}{h^2} = f(x_1) + \frac{\alpha}{h^2}$$

$$\text{per } x_N \rightarrow -\frac{u_{N-1} - 2u_N + u_{N+1}}{h^2} = f(x_N) \rightarrow \frac{-u_{N-1} + 2u_N}{h^2} = f(x_N) + \frac{\beta}{h^2}.$$

Otteniamo quindi un sistema di N equazioni algebriche nelle N incognite $\{u_j\}$ dove abbiamo imposto $u_0 = \alpha$ e $u_{N+1} = \beta$ usando le condizioni al bordo. Possiamo scrivere il precedente

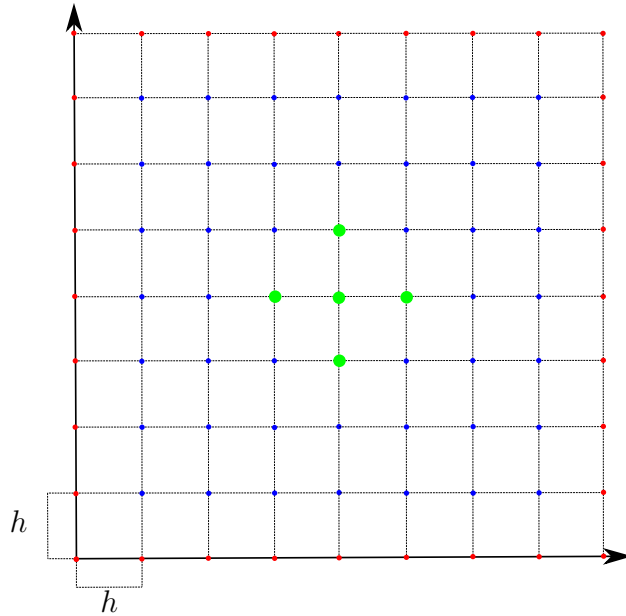
insieme di equazioni come un sistema lineare $A\mathbf{u} = \mathbf{b}$ dove

$$A = \frac{1}{h^2} \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & \dots & 0 & 0 & 0 & -1 & 2 \end{bmatrix} \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \dots \\ u_N \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} f(x_1) + \frac{\alpha}{h^2} \\ f(x_2) \\ f(x_3) \\ \dots \\ f(x_N) + \frac{\beta}{h^2} \end{bmatrix}$$

Abbiamo il seguente risultato di convergenza per il metodo appena descritto, per ogni $f \in C^2(a, b)$ risulta

$$\max_{i=1, \dots, N} |u_j - u(x_j)| \leq ch^2,$$

l'errore quindi decade quadraticamente rispetto al passo di griglia h . Se volessimo estendere le differenze finite al caso bidimensionale o tridimensionale, dovremmo innanzitutto introdurre una griglia, detta Cartesiana, che è una griglia bi o tridimensionale i cui nodi \mathbf{x}_i sono il prodotto tensore delle discretizzazioni lungo ogni asse coordinato. La griglia viene rappresentata come



Quindi, per ogni direzione si applica la differenza finita centrata vista precedentemente, ricordando che il laplaciano viene scritto come

$$\Delta u = \partial_{x^2} u + \partial_{y^2} u + \partial_{z^2} u.$$

Questo metodo è tuttavia limitato a geometrie di Ω molto semplici e quindi la sua applicazione a problemi ingegneristici è limitata.