

Strategie e analisi dell'errore per problemi di ottimizzazione vincolati regolati da equazioni di evoluzione

Claudia Bonomi Edoardo Arbib

POLITECNICO DI MILANO

Progetto per il corso di Analisi Numerica per le Equazioni a Derivate Parziali II

Contenuti

- ➊ Problema continuo
- ➋ Problema discreto
 - Equazioni di stato e aggiunta
 - Discretizzazione variazionale
- ➌ Algoritmi risolutivi
- ➍ Implementazione
- ➎ Risultati
- ➏ Conclusioni e Lavori Futuri

Setting

dominio $\Omega \times I$, $\Omega \subset \mathbb{R}^n$, $I = (0, T)$, $U_{ad} \subset U = L^2(I, \mathbb{R}^D)$

$$\min_{y \in Y, u \in U_{ad}} J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2(I, L^2(\Omega))}^2 + \frac{\alpha}{2} \|u\|_U^2 \quad (1)$$

$$\text{s.t.} \quad y = S(Bu, y_0)$$

Equazione di stato

Condizione di ottimalità

$$\begin{aligned} \partial_t \bar{y} - \Delta \bar{y} &= f && \text{in } I \times \Omega \\ \bar{y} &= 0 && \text{in } I \times \Omega \\ \bar{y}(0) &= \kappa && \text{in } \Omega \end{aligned}$$

$$\bar{u} = P_{U_{ad}} \left(-\frac{1}{\alpha} B' \bar{p} \right)$$

\bar{p} è la soluzione del problema aggiunto

$$\begin{aligned} -\partial_t \bar{p} - \Delta \bar{p} &= \bar{y} - y_d && \text{in } I \times \Omega \\ \bar{p} &= 0 && \text{in } I \times \partial\Omega \\ \bar{p}(T) &= 0 && \text{su } \Omega \end{aligned} \quad (2)$$

Spazio e operatore di controllo

Spazio di controllo

$$U_{ad} = \{u \in U \mid a_i \leq u_i(t) \leq b_i \forall i = 1 : d\} \quad (3)$$

con $a_i, b_i \in \mathbb{R}$ t.c. $a_i < b_i \forall i = 1 : d$

Operatore di controllo

$$B : U \rightarrow L^2(I, H^{-1}(\Omega)), \quad u \mapsto \left(t \mapsto \sum_{i=1}^d u_i(t) g_i \right) \quad (4)$$

con funzionali noti $g_i \in H^{-1}(\Omega)$

Discretizzazione temporale

Partizione di $[0, T)$ in sottointervalli $I_m = [t_{m-1}, t_m)$, dove

$0 = t_0 < t_1 < \dots < t_M = T \longrightarrow$ griglia primale

Seconda partizione di $[0, T)$ in intervalli $I_m^* = [t_{m-1}^*, t_m^*)$, con

$0 = t_0^* < t_1^* < \dots < t_M^* = T$ e

$t_m^* = \frac{t_{m-1} + t_m}{2}$ per $m = 1, \dots, M \longrightarrow$ griglia duale

Ambientazione funzionale

$$P_k := \left\{ v \in C([0, T], H_0^1(\Omega)) \mid v|_{I_m} \in \mathcal{P}_1(I_m, H_0^1(\Omega)) \right\},$$

$$P_k^* := \left\{ v \in C([0, T], H_0^1(\Omega)) \mid v|_{I_m^*} \in \mathcal{P}_1(I_m^*, H_0^1(\Omega)) \right\},$$

$$Y_k := \left\{ v : [0, T] \rightarrow H_0^1(\Omega) \mid v|_{I_m} \in \mathcal{P}_0(I_m, H_0^1(\Omega)) \right\}.$$

Operatori di interpolazione

$$\textcircled{1} \mathcal{P}_{Y_k} : L^2(I, H_0^1(\Omega)) \rightarrow Y_k$$

$$\mathcal{P}_{Y_k} v|_{I_m} := \frac{1}{k_m} \int_{t_{m-1}}^{t_m} v dt \text{ for } m = 1, \dots, M, \text{ e } \mathcal{P}_{Y_k} v(T) := 0$$

$$\textcircled{2} \Pi_{Y_k} : C([0, T], H_0^1(\Omega)) \rightarrow Y_k$$

$$\Pi_{Y_k} v|_{I_m} := v(t_m^*) \text{ per } m = 1, \dots, M, \quad \Pi_{Y_k} v(T) := v(T)$$

$$\textcircled{3} \pi_{P_k^*} : C([0, T], H_0^1(\Omega)) \cup Y_k \rightarrow P_k^*$$

$$\pi_{P_k^*} v|_{I_1^* \cup I_2^*} := v(t_1^*) + \frac{t - t_1^*}{t_2^* - t_1^*} (v(t_2^*) - v(t_1^*)),$$

$$\pi_{P_k^*} v|_{I_m^*} := v(t_{m-1}^*) + \frac{t - t_{m-1}^*}{t_m^* - t_{m-1}^*} (v(t_m^*) - v(t_{m-1}^*)),$$

$$\pi_{P_k^*} v|_{I_M^* \cup I_{M+1}^*} := v(t_{M-1}^*) + \frac{t - t_{M-1}^*}{t_M^* - t_{M-1}^*} (v(t_M^*) - v(t_{M-1}^*)).$$

Operatori di interpolazione

$$\textcircled{1} \mathcal{P}_{Y_k} : L^2(I, H_0^1(\Omega)) \rightarrow Y_k$$

$$\mathcal{P}_{Y_k} v|_{I_m} := \frac{1}{k_m} \int_{t_{m-1}}^{t_m} v dt \text{ for } m = 1, \dots, M, \text{ e } \mathcal{P}_{Y_k} v(T) := 0$$

$$\textcircled{2} \Pi_{Y_k} : C([0, T], H_0^1(\Omega)) \rightarrow Y_k$$

$$\Pi_{Y_k} v|_{I_m} := v(t_m^*) \text{ per } m = 1, \dots, M, \quad \Pi_{Y_k} v(T) := v(T)$$

$$\textcircled{3} \pi_{P_k^*} : C([0, T], H_0^1(\Omega)) \cup Y_k \rightarrow P_k^*$$

$$\pi_{P_k^*} v|_{I_1^* \cup I_2^*} := v(t_1^*) + \frac{t - t_1^*}{t_2^* - t_1^*} (v(t_2^*) - v(t_1^*)),$$

$$\pi_{P_k^*} v|_{I_m^*} := v(t_{m-1}^*) + \frac{t - t_{m-1}^*}{t_m^* - t_{m-1}^*} (v(t_m^*) - v(t_{m-1}^*)),$$

$$\pi_{P_k^*} v|_{I_M^* \cup I_{M+1}^*} := v(t_{M-1}^*) + \frac{t - t_{M-1}^*}{t_M^* - t_{M-1}^*} (v(t_M^*) - v(t_{M-1}^*)).$$

Operatori di interpolazione

$$\textcircled{1} \mathcal{P}_{Y_k} : L^2(I, H_0^1(\Omega)) \rightarrow Y_k$$

$$\mathcal{P}_{Y_k} v|_{I_m} := \frac{1}{k_m} \int_{t_{m-1}}^{t_m} v dt \text{ for } m = 1, \dots, M, \text{ e } \mathcal{P}_{Y_k} v(T) := 0$$

$$\textcircled{2} \Pi_{Y_k} : C([0, T], H_0^1(\Omega)) \rightarrow Y_k$$

$$\Pi_{Y_k} v|_{I_m} := v(t_m^*) \text{ per } m = 1, \dots, M, \quad \Pi_{Y_k} v(T) := v(T)$$

$$\textcircled{3} \pi_{P_k^*} : C([0, T], H_0^1(\Omega)) \cup Y_k \rightarrow P_k^*$$

$$\pi_{P_k^*} v|_{I_1^* \cup I_2^*} := v(t_1^*) + \frac{t - t_1^*}{t_2^* - t_1^*} (v(t_2^*) - v(t_1^*)),$$

$$\pi_{P_k^*} v|_{I_m^*} := v(t_{m-1}^*) + \frac{t - t_{m-1}^*}{t_m^* - t_{m-1}^*} (v(t_m^*) - v(t_{m-1}^*)),$$

$$\pi_{P_k^*} v|_{I_M^* \cup I_{M+1}^*} := v(t_{M-1}^*) + \frac{t - t_{M-1}^*}{t_M^* - t_{M-1}^*} (v(t_M^*) - v(t_{M-1}^*)).$$

Equazione di stato

Formulazione debole trovare $y_k \in Y_k$ tale che

$$\begin{aligned} \int_0^T -\langle \partial_t v(t), y(t) \rangle_{H^{-1}H_0^1} dt + \int_0^T a(y(t), v(t)) dt + (y(T), v(T))_{L^2} \\ = \int_0^T \langle f(t), v_k(t) \rangle_{H^{-1}H_0^1} dt + (\kappa, v_k(0))_{L^2} \quad \forall v_k \in P_k. \quad (5) \end{aligned}$$

\Rightarrow variante di CN con passo di Rannacher. Lo schema è consistente, stabile, convergente.

Analisi errore

$y_k \in Y_k \Rightarrow$ ordine $\mathcal{O}(k)$, ma $\pi_{P_k^*} y_k$ converge con ordine due

Lo studio dell'equazione aggiunta è analogo, solo con spazi di soluzione e test scambiati. Lo schema risultante è una variante di CN.

Discretizzazione variazionale

Problema di controllo ottimo discretizzato

$$\begin{aligned} \min_{y_k \in Y_k, u \in U_{ad}} \quad & J(y_k, u) = \frac{1}{2} \|y_k - y_d\|_{L^2(I, L^2(\Omega))}^2 + \frac{\alpha}{2} \|u\|_U^2 \\ \text{s.t.} \quad & y_k = S_k(Bu, y_0) \end{aligned} \quad (\mathbb{P}_k)$$

dove S_k è la discretizzazione di S tramite lo schema di PG.

Osservazioni

- 1 Il metodo si basa sulla discretizzazione dei soli spazi di stato e aggiunto, utilizzando implicitamente le condizioni di ottimalità del primo ordine per la discretizzazione del controllo.
- 2 Il metodo permette di disaccoppiare l'approssimazione dell'*active set* dalla scelta della griglia temporale
- 3 Il metodo è ben posto e convergente con ordine 2 rispetto al controllo u .

Discretizzazione variazionale

Problema di controllo ottimo discretizzato

$$\begin{aligned} \min_{y_k \in Y_k, u \in U_{ad}} \quad & J(y_k, u) = \frac{1}{2} \|y_k - y_d\|_{L^2(I, L^2(\Omega))}^2 + \frac{\alpha}{2} \|u\|_U^2 \\ \text{s.t.} \quad & y_k = S_k(Bu, y_0) \end{aligned} \quad (\mathbb{P}_k)$$

dove S_k è la discretizzazione di S tramite lo schema di PG.

Osservazioni

- 1 Il metodo si basa sulla discretizzazione dei soli spazi di stato e aggiunto, utilizzando implicitamente le condizioni di ottimalità del primo ordine per la discretizzazione del controllo.
- 2 Il metodo permette di disaccoppiare l'approssimazione dell'*active set* dalla scelta della griglia temporale
- 3 Il metodo è ben posto e convergente con ordine 2 rispetto al controllo u .

Discretizzazione variazionale

Problema di controllo ottimo discretizzato

$$\begin{aligned} \min_{y_k \in Y_k, u \in U_{ad}} \quad & J(y_k, u) = \frac{1}{2} \|y_k - y_d\|_{L^2(I, L^2(\Omega))}^2 + \frac{\alpha}{2} \|u\|_U^2 \\ \text{s.t.} \quad & y_k = S_k(Bu, y_0) \end{aligned} \quad (\mathbb{P}_k)$$

dove S_k è la discretizzazione di S tramite lo schema di PG.

Osservazioni

- 1 Il metodo si basa sulla discretizzazione dei soli spazi di stato e aggiunto, utilizzando implicitamente le condizioni di ottimalità del primo ordine per la discretizzazione del controllo.
- 2 Il metodo permette di disaccoppiare l'approssimazione dell'*active set* dalla scelta della griglia temporale
- 3 Il metodo è ben posto e convergente con ordine 2 rispetto al controllo u .

Punto fisso

La CNES di ottimalità del problema discreto è sempre

$\bar{u}_k = P_{U_{ad}} \left(-\frac{1}{\alpha} B' \bar{p}_k \right)$. Le iterazioni di punto fisso si applicano proprio a quest'equazione \Rightarrow

Algoritmo

- ➊ Inizializzare $u_h^0 \in U_{ad}$, $n := 0$.
- ➋ Ripetere fino a convergenza
 - ➊ calcolare Bu_h^n ,
 - ➋ calcolare $y_h^n = S_h(y_0, Bu_h^n)$,
 - ➌ calcolare $p_h^n = S_h^*(y_h^n - y_d)$,
 - ➍ calcolare $u_h^{n+1} = P_{U_{ad}}(-\frac{1}{\alpha} B' p_h^n)$,
 - ➎ porre $n=n+1$.

Criterio di arresto: $\|B'(p_h^{n+1} - p_h^n)\|_{L^\infty(\Omega \times I)} < \epsilon$

Non converge per α piccoli

Semi-Newton

Metodo di Newton con minimizzazione monodimensionale (Armijo)
 \Rightarrow formulazione del problema tramite Lagrangiana primale e duale

$$\phi(w) = - \inf_{u, y \in L^2(I, L^2)} \underbrace{\left(\frac{1}{2} \|y - y_d\|^2 + \frac{\alpha}{2} \|u\|^2 + \chi_{U_{ad}}(u) - (w, y - S_h u) \right)}_{\mathcal{L}(u, y, w)} \quad (6)$$

che diventa quindi un caso di minimizzazione non vincolata
 $\min_{w \in L^2(I, L^2(\Omega))} \phi(w)$. Qual è la CNES di ottimalità? Si può
 applicare un metodo di Newton? Risponde il seguente

Semi-Newton II

Lemma

La funzione $\phi : L^2(I, L^2(\Omega)) \rightarrow \mathbb{R}$ è fortemente convessa e Frechet-differenziabile con gradiente lipschitziano

$$\nabla \phi(w) = y(w) - S_h u(w), \quad (7)$$

dove $y_h(w) = w + y_d$ e $u(w) = P_{U_{ad}}(-\frac{1}{\alpha} S_h^ w)$ sono gli unici punti di minimo della lagrangiana $\mathcal{L}(u, y, w)$ per ogni $w \in L^2(I, L^2(\Omega))$ data.*

Ma allora calcolato l'Hessiano generalizzato di ϕ , ogni iterazione di Newton risolve

$$\left(I + \frac{1}{\alpha} S_h \mathbb{1}_{S_h^* w} S_h^*\right) \delta w = -(w + y_d) + S_h P_{U_{ad}}\left(-\frac{1}{\alpha} S_h^* w\right). \quad (8)$$

Ecco quindi l'algoritmo promesso

Semi-Newton III

Algoritmo

- ➊ Inizializzare $w^0 \in L^2(I, L^2(\Omega))$, $\beta \in (0, 1)$, $k = 0$,
- ➋ Ripetere fino a convergenza
 - ➊ Risolvere l'equazione (8) per δw^k tramite CG,
 - ➋ Porre $\lambda := 1$,
 - ➌ Finché risulta vera la condizione
 $\phi(w^k + \lambda \delta w^k) > \phi(w) + \frac{1}{3} \lambda (\nabla \phi(w^k), \delta w^k)_{L^2(I, L^2(\Omega))}$, porre
 $\lambda := \beta \lambda$,
 - ➍ Porre $w^{k+1} = w^k + \lambda \delta w^k$
 - ➎ Porre $k := k + 1$.

criterio di arresto: $\|\nabla \phi(w^k)\| \leq t_0$

L'algoritmo converge e il criterio di arresto è plausibile.

Significato della simbologia

$\chi_{U_{ad}}$

Indica la funzione caratteristica dell'insieme U_{ad} nel senso dell'analisi convessa, ovvero

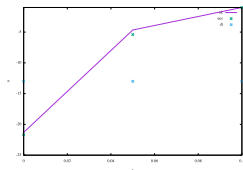
$$\chi_{U_{ad}} = \begin{cases} 0, & \text{su } U_{ad}, \\ \infty & \text{su } L^2(I, L^2(\Omega)) \setminus U_{ad}. \end{cases} \quad (9)$$

$\mathbb{1}_{p_h(v)}$

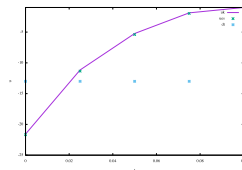
Introdotta l'*inactive set* della funzione p_h come l'insieme $\mathcal{I}(p_h) = \{ \omega \in \Omega \times [0, T] \mid (-\frac{1}{\alpha} p_h(v))(\omega) \in (a(\omega), b(\omega)) \}$ e $\mathbb{1}_{\mathcal{I}(p_h)}$ come la funzione indicatrice di tale insieme con $\mathbb{1}_{p_h(v)}$ si denota l'endomorfismo auto-aggiunto in $L^2(I, L^2(\Omega))$ dato dalla moltiplicazione puntuale con $\mathbb{1}_{\mathcal{I}(p_h)}$.

Gli strumenti di sviluppo utilizzati sono **Freefem++** e **GitHub**

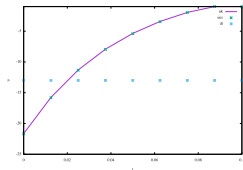
Test Case 01 Punto fisso \bar{u} e u_k



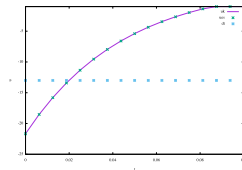
(a) $l=1$



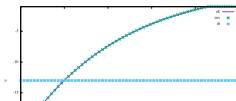
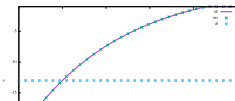
(b) $l=2$



(c) $l=3$



(d) $l=4$



Test Case 01 Punto fisso

Tabella: Punto fisso per Test case 01: errori e EOC

l	$\ \bar{u} - u_{kh}\ _{L^2(L^2)}$	$\ \bar{y} - y_{kh}\ _{L^2(L^2)}$	EOC_u	EOC_y
1	0.31667	0.981285	—	—
2	0.0835064	0.496296	2.60937	1.33449
3	0.0209608	0.248822	2.35165	1.17464
4	0.00500916	0.124494	2.25065	1.08882
5	0.00109219	0.0622586	2.29624	1.04473
6	0.000497644	0.0311327	1.15957	1.02236

Test Case 01 Punto fisso

Tabella: Punto fisso per Test case 01: errori e EOC

l	$\ \bar{y} - \pi_{P_k^*} y_{kh}\ _{L^2(L^2)}$	$\ \bar{p} - p_{kh}\ _{L^2(L^2)}$	EOC_{π_y}	EOC_p
1	0.520894	0.00660747	—	—
2	0.15134	0.00173155	2.41965	2.6216
3	0.0393476	0.00043334	2.29181	2.35673
4	0.00970087	0.000103613	2.20164	2.24982
5	0.00221619	0.000022824	2.2259	2.28078
6	0.000432024	0.000010724	2.41203	1.11429

Test Case 01 Semi-Newton

