

Appunti per esercitazioni di istituzioni di analisi numerica

Federico Poloni
f.poloni@sns.it

10 febbraio 2010

1 Esercizi

1.1 Esercizi 6.10–6.14 del libro di Bevilacqua et al. [1]

Alcuni esercizi tratti dal libro. Le dimostrazioni sono in parte quelle degli “hint” del libro e in parte mie.

Esercizio 1. *Come mai la base dei monomi $p_k(x) = x^k$ non è ortogonale rispetto a nessun prodotto scalare $\langle f, g \rangle := \int \omega(x)f(x)g(x)dx$?*

Ipotesi sottintesa: $\langle \cdot, \cdot \rangle$ dev'essere definito positivo sulle funzioni continue. Quindi se $\omega(x) > 0$ tutto bene; se ω ha qualche zero isolato tutto bene di nuovo, se si annulla su un intero intervallo non più.

Quick proof

Dimostrazione. Per assurdo:

$$0 = \langle 1, x^2 \rangle = \int \omega(x)x^2 dx = \langle x, x \rangle > 0$$

perché è definito positivo. □

Standard proof Dal conto per ricavare la relazione a tre termini, si aveva

$$C = \frac{A_{n+1}}{A_n} \frac{\langle p_n, p_n \rangle}{\langle p_{n-1}, p_{n-1} \rangle},$$

quindi C non può essere nullo (se il prodotto scalare è definito positivo).

Una terza dimostrazione segue dal prossimo esercizio.

Esercizio 2. *Siano p_i polinomi ortogonali su $[a, b]$. Tutte le funzioni continue f ortogonali a p_0, p_1, \dots, p_{n-1} si annullano (con cambio di segno) almeno n volte nell'intervallo $[a, b]$*

Dimostrazione. Supponiamo per assurdo che f cambi segno meno di n volte. Costruiamo il polinomio q di grado al più $n - 1$ che ha lo stesso segno di f in tutti i punti di $[a, b]$ (basta prendere il polinomio che si annulla in...). Ora $\langle f, q \rangle = \int_a^b \omega(x) f(x) q(x) dx > 0$ perché $f q \geq 0$. Ma q si scrive come combinazione lineare di p_0, p_1, \dots, p_{n-1} , quindi dovrebbe essere ortogonale a f , assurdo. \square

Esercizio 3. Le radici di due polinomi ortogonali successivi p_n e p_{n+1} sono “interleaved”: cioè, se $x_1 < x_2 < \dots < x_n$ sono le radici di p_n e $y_1 < y_2 < \dots < y_{n+1}$ sono le radici di p_{n+1} , si ha

$$y_1 < x_1 < y_2 < x_2 < y_3 < \dots < x_n < y_{n+1}.$$

(come si dice interleaved in italiano?)

Dimostrazione. Notiamo innanzitutto che per l'esercizio 2 p_{n+1} ha $n + 1$ zeri nell'intervallo, che chiamiamo $y_1 < y_2 < \dots < y_{n+1}$. Derivando Christoffel-Darboux rispetto a x e ponendo $y = x$, si ha

$$0 < \sum_{i=0}^n \frac{[p_i(x)]^2}{h_i} = \gamma_n [p'_{n+1}(x)p_n(x) - p_{n+1}(x)p'_n(x)].$$

Quindi in particolare

$$p'_{n+1}(x)p_n(x) - p_{n+1}(x)p'_n(x) \tag{1}$$

ha lo stesso segno per ogni x (positivo o negativo, a seconda di come si normalizzano p_n e p_{n+1}). Prendiamo ora $x = y_i$ e $x = y_{i+1}$. Si ha $p_{n+1}(y_i) = p_{n+1}(y_{i+1}) = 0$, mentre $p'_{n+1}(y_i)$ e $p'_{n+1}(y_{i+1})$ hanno segni opposti (visto che tutte le radici sono semplici e il polinomio deve “cambiare segno” a ogni radice). Quindi controllando i segni nella (1) si ha che anche $p_n(y_i)$ e $p_n(y_{i+1})$ hanno segni opposti, e perciò p_n ha (almeno) uno zero compreso strettamente tra y_i e y_{i+1} . Abbiamo trovato in questo modo n zeri distinti di p_n , uno in ogni intervallo (y_i, y_{i+1}) , quindi devono essere tutti.

Sul libro lo fa invertendo il ruolo degli x_i e degli y_i : parte dal fatto che p_n ha n radici nell'intervallo, e dimostra che p_{n+1} ha una radice in ognuno degli intervalli (x_i, x_{i+1}) e che ce ne sono due “più in fuori”. \square

1.2 Equazioni differenziali dei polinomi ortogonali classici

I polinomi ortogonali classici possono essere definiti anche come soluzioni di una famiglia di equazioni differenziali. Cercheremo ora di ricavare questi risultati senza troppi conti¹.

Ricorda la formula di Leibniz

$$\frac{d^n}{dx^n} fg = \sum_{k=0}^n \binom{n}{k} \frac{d^k}{dx^k} f \frac{d^{n-k}}{dx^{n-k}} g.$$

¹Se studiando trovate una dimostrazione più veloce, ditemela che mi interessa.

Legendre

Esercizio 4. *L' n -esimo polinomio di Legendre soddisfa l'equazione differenziale*

$$(1 - x^2)P_n'' - 2xP_n' + n(n + 1)P_n = 0. \quad (2)$$

Breve approfondimento: questa equazione differenziale si incontra spesso in fisica e in chimica quantistica: i polinomi di Legendre sono associati alle *armoniche sferiche*, che si usano per risolvere l'equazione di Maxwell (cioè determinare come si comporta il campo elettrico nel vuoto intorno a una sorgente puntiforme) e l'equazione di Schrödinger per l'atomo di idrogeno (cioè per determinare la forma degli orbitali — quelli che compaiono con dei disegni di forme buffe nei libri di chimica) in coordinate sferiche.

Dimostrazione. Si utilizza la formula di Rodrigues

$$P_n = \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} (1 - x^2)^n.$$

Si parte da

$$X = \frac{d^{n+2}}{dx^{n+2}} (1 - x^2)^{n+1} \quad (3)$$

e si espande usando la formula di Leibniz con $f = (1 - x^2)^n$, $g = (1 - x^2)$. I sommandi dove g è derivata più di due volte muoiono. Quindi resta

$$X = C_n \left((1 - x^2)P_n'' - (n + 2)2xP_n' - 2 \binom{n + 2}{2} P_n \right), \quad (4)$$

dove $C_n = \frac{(-1)^n}{2^n n!}$ è una costante che tanto si semplificherà.

Ora si riparte da (3), si calcola esplicitamente una derivata

$$X = \frac{d^{n+1}}{dx^{n+1}} ((n + 1)(-2x)(1 - x^2)^n),$$

e si riapplica Leibniz a quest'ultima espressione con $f = (1 - x^2)^n$, $g = x$:

$$X = C_n (-2(n + 1)xP_n' - 2(n + 1)^2 P_n). \quad (5)$$

Uguagliare (4) e (5), e semplificare. \square

Chebyshev

Esercizio 5. *Gli n -esimi polinomi di Chebyshev di primo e secondo tipo soddisfano rispettivamente le equazioni differenziali*

$$\begin{aligned} (1 - x^2)T_n'' - xT_n' + n^2 T_n &= 0 \\ (1 - x^2)U_n'' - 3xU_n' + n(n + 2)U_n &= 0 \end{aligned}$$

Dimostrazione. Si fa esplicitamente il cambio di variabile $x = \cos \theta$. Così $\frac{dx}{d\theta} = -\sin \theta$, e $\frac{dT_n}{d\theta} = \frac{d}{d\theta} \cos n\theta = -n \sin n\theta$. Allora

$$\frac{dT_n}{dx} = \frac{dT_n}{d\theta} \frac{d\theta}{dx} = \frac{n \sin n\theta}{\sin \theta}. \quad (6)$$

Chiamiamo $W_n := \frac{n \sin n\theta}{\sin \theta}$ e deriviamo ancora rispetto a x con lo stesso trucco

$$\frac{dW_n}{d\theta} = \frac{n^2 \cos n\theta \sin \theta - n \cos \theta \sin n\theta}{\sin^2 \theta}.$$

A questo punto si sostituiscono $(1 - x^2) = \sin^2 \theta$, $x = \cos \theta$ e tutti i valori calcolati nell'espressione

$$(1 - x^2)T_n'' - xT_n' + n^2T_n \quad (7)$$

e si verifica a mano che fa 0.

Per i Chebyshev di secondo tipo, si nota che essendo $U_n = \frac{\sin(n+1)\theta}{\sin \theta}$, la (6) implica $T_n' = nU_{n-1}$; quindi basta derivare una volta la (7) e viene un'equazione del second'ordine nel solo U_{n-1} . \square

Laguerre L'equazione è

$$xL_n'' + (1 - x)L_n' + nL_n = 0.$$

Dimostrazione. (questa secondo me è migliorabile) Si ha la formula di Rodrigues

$$L_n = \frac{1}{n!} e^x \frac{d^n}{dx^n} (e^{-x} x^n).$$

Notare che usare lo stesso trucco usato per i Legendre genera qualche complicazione perché $L_n' \neq \frac{d^{n+1}}{dx^{n+1}} (e^{-x} x^n)$. Invece definiamo $W_n := \frac{d^n}{dx^n} (e^{-x} x^n)$; si ha

$$\begin{aligned} n!L_n &= e^x W_n \\ n!L_n' &= e^x (W_n + W_n') \\ n!L_n'' &= e^x (W_n'' + 2W_n' + W_n). \end{aligned} \quad (8)$$

Possiamo applicare il trucco di prima ai W_n :

$$X = \frac{d^{n+2}}{dx^{n+2}} e^{-x} x^{n+1} = W_n'' x + (n+2)W_n'; \quad (9)$$

$$X = \frac{d^{n+1}}{dx^{n+1}} (-e^{-x} x^{n+1} + (n+1)e^{-x} x^n) = -W_n' x - (n+2)W_n + (n+1)W_n' \quad (10)$$

(i passaggi non ovvi si ottengono da Leibniz con $f = e^{-x} x^n$, $g = x$). Uguagliare (9) e (10) e scrivere tutto in funzione dei L_n utilizzando (8). \square

Hermite L'equazione è

$$H_n'' - 2xH_n' + 2nH_n = 0.$$

Dimostrazione. Stavolta usiamo non solo la formula di Rodrigues

$$H_n = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$$

ma anche la relazione a tre termini

$$H_{n+1} = 2xH_n - 2nH_{n-1}. \quad (11)$$

(notare che con solo la relazione a tre termini non si arriva da nessuna parte perché non definisce univocamente i polinomi — difatti servono anche i dati iniziali).

Derivando H_n si ha

$$H_n' = (-1)^n 2xe^{x^2} \frac{d^n}{dx^n} e^{-x^2} + (-1)^n e^{x^2} \frac{d^{n+1}}{dx^{n+1}} e^{-x^2} = 2xH_n - H_{n+1} = 2nH_{n-1} \quad (12)$$

(l'ultimo passaggio è la (11)).

Analogamente

$$H_n'' = 2n(H_{n-1})' = 2n2(n-1)H_{n-2} = 2n(H_n - 2xH_{n-1}), \quad (13)$$

usando ripetutamente la (11).

Ora combiniamo la (12) e la (13) in modo da eliminare H_{n-1} e abbiamo vinto. \square

1.3 Approssimazione dei coseni vs. Chebyshev

Data una funzione $f : [0, \pi] \rightarrow \mathbb{R}$, possiamo costruire la funzione $\tilde{f} : [-\pi, \pi] \rightarrow \mathbb{R}$ data da $\tilde{f}(x) := f(|x|)$ (cioè f “simmetrizzata” rispetto all'asse y). Possiamo anche fare la “simmetrizzata dispari” $\hat{f}(x) := \operatorname{sgn}(x)f(|x|)$. Notare che \tilde{f} è pari e \hat{f} è dispari.

Posso approssimare \tilde{f} con una serie di Fourier, che verrà composta di soli coseni perché \tilde{f} è pari. Posso approssimare \hat{f} con una serie di Fourier, che verrà composta di soli seni.

Esercizio 6. *Su una funzione “normale”, quale delle due approssimazioni “funziona meglio”? (esercizio fumoso e mal posto)*

Dimostrazione. La \tilde{f} , perché i risultati classici assicurano la “buona convergenza” delle serie di Fourier per funzioni di classe C^1 , o almeno Lipschitziane: e mentre la \tilde{f} lo è (per funzioni “normali”), la \hat{f} è discontinua in zero (a meno che $f(0) = 0$). \square

Esercizio 7. *Chi sono i coefficienti della serie di \tilde{f} ?*

Dimostrazione.

$$a_n = \frac{2}{\pi} \int_0^\pi f(y) \cos(ny) dy = \frac{2}{\pi} \int_{-1}^1 f(\arccos x) T_n(x) \frac{-1}{\sqrt{1-x^2}} dx,$$

(cambio di variabile $\cos y = x$) quindi è come se facessi un'approssimazione con polinomi di Chebyshev di $f(\arccos x)$ (coefficienti da aggiustare probabilmente). Esercizio extra: cosa succede con la \hat{f} ? (se non sbaglio i conti, dovrebbe venire un'approssimazione di $f(\arccos x)/\sin(x)$ con polinomi di Chebyshev di secondo tipo U_n). \square

2 Metodi variazionali

Ho seguito il libro di Stoer [3], ma aggiungendo qualche dettaglio per cui mi sono rifatto al libro di Quarteroni e Valli [4]

2.1 Precisazione

Sia B uno spazio di funzioni. Per evitare di generare confusione chiamando troppe cose con il nome “funzione”, solitamente in analisi si preferisce chiamare *operatore* una mappa da B in sé (o comunque da uno spazio di funzioni a un altro spazio di funzioni), e *funzionale* una mappa da B a \mathbb{R} (o \mathbb{C}). Notate anche che nella frase qui sopra, per evitare confusione, ho usato il termine *mappa*, che in fondo è un altro sinonimo di funzione.

Per evitare di avere troppe parentesi tonde, qui useremo le parentesi quadre per indicare l'applicazione di un operatore: per esempio, $L[f]$ indicherà l'applicazione dell'operatore L alla funzione f . La notazione $L[f](x)$, invece, indica che dobbiamo applicare l'operatore L alla funzione f , e valutare la funzione risultante nel punto x .

2.2 Introduzione

I metodi alle differenze finite hanno due difetti:

1. Funzionano bene su griglie quadrate regolari, ma si adattano male a domini di forma più strana, per problemi in 2 o più dimensioni, che invece spesso capitano nelle applicazioni.
2. Richiedono un'alta regolarità della soluzione: per dimostrarne il funzionamento, abbiamo usato sviluppi di Taylor fino al IV ordine, richiedendo quindi che la funzione sia C^4 (e che $\|u^{(4)}\|_\infty$ non sia troppo grossa).

I metodi chiamati variazionali, o di Ritz-Galerkin, o degli elementi finiti, invece funzionano meglio sotto questi punti di vista.

2.3 Problema modello

Consideriamo il generico *problema di Sturm-Liouville*; cerchiamo $u : [a, b] \rightarrow \mathbb{R}$ di classe C^2 tale che

$$\begin{aligned}L[u] &= f; \\ u(a) &= \alpha \\ u(b) &= \beta,\end{aligned}\tag{14}$$

dove l'operatore L è definito come

$$L[v] := - [p(x)v'(x)]' + q(x)v(x)$$

per due funzioni $p(x) > 0, p \in C^1$, e $q(x) > 0, q \in C^0$ date. Nel seguito, quando possibile ometteremo i (x) dalle funzioni: per esempio l'equazione qui sopra sarà scritta in modo più compatto

$$L[v] = -[pv']' + qv.$$

Notate che L è un operatore lineare. Notate che l'equazione (2) è un esempio di un problema di questa classe.

Innanzitutto mi tolgo dai piedi le condizioni al contorno, cercando di rimpiazzarle con $u(a) = u(b) = 0$: per questo, scelgo una qualunque funzione $l(x) \in C^2$ che soddisfi $l(a) = \alpha, l(b) = \beta$ (per esempio c'è un polinomio di grado 1 che va bene), e noto che $\tilde{u} = u - l$ soddisfa

$$\begin{aligned}L[\tilde{u}] &= f - L[l] \\ \tilde{u}(a) &= 0 \\ \tilde{u}(b) &= 0\end{aligned}$$

che è un problema dello stesso tipo ma con condizioni al bordo omogenee. Con qualche altro trucco (che non vediamo) è possibile trattare più o meno tutte le condizioni al contorno, incluse quelle di Neumann.

2.4 Forma debole

Consideriamo il prodotto scalare L^2 classico

$$\langle u, v \rangle = \int_a^b uv dx.$$

Nota che se u soddisfa (14), allora (integrando) soddisfa anche

$$\langle L[u], v \rangle = \langle f, v \rangle \quad \forall v \in C_a^1 \text{ tratti, } v(a) = v(b) = 0.\tag{15}$$

Nota che con quelle ipotesi su v possiamo scrivere

$$\langle L[u], v \rangle = - \int_a^b (pu')' v dx + \int_a^b quv dx = \int_a^b pu' v' dx + \int_a^b quv dx,\tag{16}$$

dove abbiamo integrato per parti il primo pezzo.

Con questa espressione per $\langle L[u], v \rangle$, nota che il problema (15) ha senso anche se u e v sono due funzioni soltanto $C^1_{\text{a tratti}}$ e nulle ai bordi. Anzi, possiamo chiedere solo che u' e v' esistano come derivate distribuzionali. Lo spazio delle funzioni

$$\{v : v \in L^2[a, b], v' \in L^2[a, b], v(a) = v(b) = 0\} =: H_0^1, \quad (17)$$

dove la derivata è fatta in senso distribuzionale, è detto uno *spazio di Sobolev*. Se non sapete cos'è una derivata distribuzionale, ignorate pure questa parte e prendete la (17) come definizione di H_0^1 solo con le funzioni C^1 a tratti, e tutto funzionerà lo stesso.

2.5 Proprietà di L

Su H_0^1 , l'operatore L è simmetrico rispetto al prodotto scalare $\langle \cdot, \cdot \rangle$ (ovvio dalla (16)); possiamo anche provare che è definito positivo (in un certo senso).

Lemma 8 (versione 1D del lemma di Poincaré). *Esiste una costante C (che dipende solo dal dominio $[a, b]$) tale che per ogni $u \in H_0^1$ si ha*

$$\|u\|_N \leq C \|u'\|_{L^2}$$

per le due norme $N = \infty$ e $N = L^2$.

Dimostrazione. Usando Cauchy-Schwarz in modo furbo, si ha $\forall z \in [a, b]$

$$u(z)^2 = \left(\int_a^z u' dx \right)^2 \leq \left(\int_a^z 1^2 dx \right) \left(\int_a^z u'^2 dx \right) \leq (z - a) \|u'\|_{L^2}^2.$$

Visto che questa disuguaglianza vale per ogni $z \in [a, b]$, si ha

$$\|u\|_{\infty}^2 \leq (b - a) \|u'\|_{L^2}^2.$$

Similmente, integrando entrambi i lati della disuguaglianza si ottiene

$$\|u\|_{L^2}^2 = \int_a^b u(z)^2 dz \leq \left(\int_a^b (z - a) dz \right) \|u'\|_{L^2}^2$$

□

Teorema 9. *Esistono costanti m, M che dipendono solo dal dominio tali che per ogni $u \in H_0^1$ si ha*

$$m \|u\|_N^2 \leq \langle L[u], u \rangle \leq M \|u\|_N^2$$

Dimostrazione.

$$\begin{aligned}\langle L[u], u \rangle &= \int_a^b p u'^2 dx + \int_a^b q u^2 dx \geq (\min p) \int_a^b u'^2 dx \\ &\geq (\min p) \|u'\|_{L^2}^2 \geq (\min p) C^2 \|u\|_N^2\end{aligned}$$

$$\begin{aligned}\langle L[u], u \rangle &= \int_a^b p u'^2 dx + \int_a^b q u^2 dx \leq (\max p) \|u'\|_{L^2}^2 + (\max q) \|u\|_{L^2}^2 \\ &\leq (\max p) \|u'\|_{L^2}^2 + (\max q) C^2 \|u'\|_{L^2}^2\end{aligned}$$

□

2.6 Formulazione variazionale

Teorema 10. *Se u soddisfa (15), allora u è un punto di minimo (stretto) in H_0^1 del funzionale $F(u) := \langle L[v], v \rangle - 2 \langle f, v \rangle$. In particolare, (15) ha al più una soluzione.*

Dimostrazione. Usando la simmetria di L ,

$$F(v) + \langle L[u], u \rangle = \langle L[v - u], v - u \rangle \geq \|v - u\|_N \geq 0,$$

con l'ultimo \geq sempre stretto a meno che $u = v$.

□

2.7 Discretizzazione

Finora abbiamo dimostrato tante belle proprietà ma non abbiamo enunciato algoritmi per risolvere la (14). Lo facciamo ora: nella (15), rimpiazziamo H_0^1 con un suo sottospazio S con $\dim S = n < \infty$: cioè, cerchiamo $u_S \in S$ tale che

$$\langle L[u], v \rangle = \langle f, v \rangle \quad \forall v \in S. \quad (18)$$

Fissiamo una base di funzioni $(\phi_i(x))_{i=1}^n$ per S . Allora, u_S si scriverà in funzione della base come $u_S = \sum_{j=1}^n x_j \phi_j$. Scrivendo l'equazione (18) per $v = \phi_i$, $i = 1, 2, \dots, n$ (perché bastano?) otteniamo il sistema

$$\sum_{j=1}^n \langle \phi_i, L[\phi_j] \rangle x_j = \langle f, \phi_i \rangle \quad i = 1, 2, \dots, n. \quad (19)$$

È un sistema con matrice A_S simmetrica e positiva definita (perché $\langle f, L[f] \rangle \geq \|f\|_{L^2}^2 > 0$ per ogni $0 \neq f \in S$). Quindi possiamo risolverlo e ricavare gli x_i , e quindi u_S .

Nota che il teorema 10 funziona (con la stessa dimostrazione) anche se rimpiazziamo H_0^1 con S , per cui u_S è il punto di minimo (stretto) del funzionale F su S .

2.8 Errore di discretizzazione

Teorema 11 (Lemma di Céa). *Esiste una costante (che dipende solo dal dominio) D per cui per ogni $v \in S$ si ha*

$$\|u - u_S\|_N \leq D \|v' - u'\|_N$$

per le norme $N = 2$, $N = \infty$.

Dimostrazione.

$$\begin{aligned} m \|u - u_S\|_N &\leq \langle L[u - u_S], u - u_S \rangle = \langle u, L[u] \rangle + F(u_S) \\ &\leq \langle u, L[u] \rangle + F(v) = \langle L[u - v], u - v \rangle \leq M \|v' - u'\|_N \end{aligned}$$

dove nel \leq a cavallo dell'andata a capo abbiamo usato l'osservazione appena fatta che u_S è il minimo di F su S , e quindi $F(u_S) \leq F(v)$. \square

Quindi l'errore globale sulla soluzione dipende solo dall'*errore di approssimazione* della derivata prima v nello spazio S : cioè, se scelgo un S dove la derivata prima di u può venire approssimata "bene" (cioè c'è una $v \in S$ tale che $v' - u'$ è piccolo), allora ottengo una buona stima della soluzione.

2.9 Esempio: Spline cubiche

Prendiamo come S lo spazio delle spline cubiche sui punti equispaziati x_0, x_1, \dots, x_N . Ha dimensione $N + 1$ (infatti basta fissare i valori della u su x_0, x_1, \dots, x_N e possiamo costruire la sua spline). È possibile (non lo vediamo qui) costruire una sua base tale che per ogni i la funzione di base ϕ_i è diversa da zero solo nell'intervallo $[x_{i-2}, x_{i+2}]$. Ciò è bello perché implica che gli elementi della matrice del sistema $A_S = \langle \phi_i, L[\phi_j] \rangle$ sono nulli tutte le volte che $|j - i| > 2$ (è evidente dalla (16)). Quindi la matrice del sistema è pentadiagonale, e possiamo applicarci velocemente la maggior parte degli algoritmi di algebra lineare (eliminazione di Gauss in $O(n)$, un passo dei principali metodi iterativi in $O(n)$).

Per quanto riguarda l'errore globale, si può dimostrare (noi lo omettiamo) il seguente risultato:

Lemma 12. *Sia $u \in C^4$. La spline cubica che approssima u nei punti x_0, x_1, \dots, x_N soddisfa*

$$\|v' - u'\|_\infty \leq C \|u^{(4)}\|_\infty h^3$$

per una costante moderata C .

Questo + il lemma di Céa ci permettono di dire che l'errore globale per questo metodo agli elementi finiti è al più $C \|u^{(4)}\|_\infty h^3$. Notare che per le differenze finite c'era un'espressione simile ma h^2 al posto di h^3 , quindi gli elementi finiti si avvicinano meglio alla soluzione.

2.10 Esempio: funzioni lineari a tratti

Prendiamo come S lo spazio delle funzioni di approssimazione lineare nei punti x_0, x_1, \dots, x_N (cioè le funzioni lineari a tratti con i punti di non derivabilità nelle x_i). Una sua base è data dalle funzioni ϕ_i tali che $\phi_i(x_j) = \delta_{ij}$ (“hat functions”); con questa base la A_S risulta tridiagonale.

Notare che hanno la regolarità minima che serve per far funzionare gli elementi finiti (C^1 a tratti). Si può dimostrare (non difficile, ma lo omettiamo) che per la funzione v di approssimazione lineare per una funzione $u \in C^2$ soddisfa

$$\|v' - u'\|_\infty \leq C \|u^{(2)}\|_\infty h.$$

(notare che C^2 è la minima regolarità di u per cui ha senso porsi il problema (14)). Questo + lemma di Céa ci dicono che la soluzione converge con errore h .

2.11 Curiosità: cosa succede in più dimensioni

Tutti i risultati enunciati finora non richiedono particolari proprietà di \mathbb{R} , e infatti funzionano anche quando le funzioni sono definite su un aperto Ω di \mathbb{R}^2 , per ogni operatore L per cui si riesce a dimostrare il teorema 9. Non serve che L sia simmetrico, ma se non lo è le dimostrazioni si complicano un po'. Nel caso simmetrico, quasi tutte le dimostrazioni che abbiamo visto funzionano pari pari (eccezione: lemma di Poincaré).

Sono possibili diverse scelte dello spazio S , a seconda del problema; la più comune è prendere una *triangolazione* del dominio, cioè una suddivisione di Ω in tanti piccoli triangoli, e considerare le funzioni che sono continue globalmente e lineari su ognuno dei triangoli. Con la base opportuna (le funzioni che valgono 1 in un vertice della triangolazione e 0 altrove), la matrice del sistema non ha più una struttura particolare come nel caso 1D, ma è comunque molto sparsa.

Le operazioni da fare nella pratica sono:

- Costruire la triangolazione del dominio. Non è difficile, ci sono algoritmi dedicati, ma talvolta non serve neppure applicarli perché molte applicazioni partono da programmi di progettazione al computer in cui il risultato è già naturalmente suddiviso in triangoli (o tetraedri nel caso 3D)—se qualcuno si interessa di computer-grafica 3D o di CAD ha già capito di cosa parliamo.
- Costruire la matrice e il termine noto del sistema (19). Qui ci sono molti integrali da fare. Nella pratica, questi integrali vengono calcolati utilizzando su ogni triangolo (o su ogni segmento $[x_i, x_{i+1}]$ di suddivisione dell'intervallo $[a, b]$ nel caso 1D) una formula di quadratura. Solitamente questa è la parte che porta via più tempo.

- Risolvere il sistema. Visto che di solito la matrice è molto sparsa, si usa un metodo iterativo (Jacobi, gradiente coniugato, GMRES), spesso accoppiato con preconditionatori sofisticati. Di solito si arresta il metodo iterativo con un residuo abbastanza alto: $\epsilon = 10^{-3}$, per esempio. Difatti, spesso, a causa degli errori già commessi “a monte” (nei dati sperimentali, nel modello, nella quadratura, errore globale nella soluzione dell’equazione differenziale), la precisione del risultato è già limitata a poche cifre significative.

3 Teorema di Korovkin

Ho seguito delle dispense di Bruno Iannazzo per un corso tenuto all’università dell’Insubria [2], cambiando leggermente la notazione.

Un operatore L è detto *operatore lineare positivo* (LPO) se:

1. È lineare, cioè $L[\lambda f] = \lambda L[f]$ per ogni $\lambda \in \mathbb{R}$ (o \mathbb{C}), e $L[f + g] = L[f] + L[g]$
2. È *positivo*, cioè $L[f](x) \geq 0$ per ogni funzione f tale che $f(x) \geq 0$ per ogni x . Cioè L manda funzioni non negative su tutto il dominio in funzioni non negative su tutto il dominio.

Se gli LPO sono su uno spazio di funzioni a valori in \mathbb{C} , la proprietà 2 va interpretata nel senso che ogni volta che f è a valori *reali* e ≥ 0 , allora anche $L[f]$ lo è.

Notate che le due ipotesi implicano che se $f(x) \leq g(x)$ per ogni x nel dominio, allora

$$L[f](x) \leq L[g](x) \quad \text{per ogni } x \text{ nel dominio} \quad (20)$$

e che

$$|L[f]| \leq L[|f|]. \quad (21)$$

Sia L_n una successione di operatori. Diciamo che L_n *approssima bene* una funzione f (in una certa norma $\|\cdot\|$) se $\|L_n[f] - f\| \rightarrow 0$ per $n \rightarrow \infty$.

Teorema 13 (Teorema di Korovkin, versione 1-dimensionale). *Sia $K \subseteq \mathbb{R}$ un compatto, e sia $C(K)$ lo spazio delle funzioni continue da K a \mathbb{R} con la norma del sup (convergenza uniforme). Sia $(L_n)_{n=1}^{\infty}$ una successione di LPO su $C(K)$. Se L_n approssima bene le tre funzioni:*

- $x \mapsto 1$ (la funzione costante uguale a 1),
- $x \mapsto x$ (l’identità),
- $x \mapsto x^2$ (la funzione “elevare al quadrato”),

allora L_n approssima bene tutte le funzioni di $C(K)$.

Dimostrazione. Siano $f \in C(K)$, e $\varepsilon > 0$ fissati; dobbiamo mostrare che esiste \tilde{n} tale che per ogni $n \geq \tilde{n}$ valga

$$|L_n[f](a) - f(a)| \leq \varepsilon \quad \forall a \in K.$$

Notiamo anche che al posto di ε al membro di destra possiamo mettere anche un'espressione del tipo $C\varepsilon$, a patto che C non dipenda né da n né da a .

Chiamiamo $M := \|f\|_\infty$,

$$N := \max_{x \in K} |x|,$$

e δ la costante di uniforme continuità di f , cioè quel valore per cui $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$ (che esiste perché una funzione continua su un compatto è uniformemente continua).

Ora cerchiamo di sfruttare l'ipotesi: definiamo per ogni n, y

$$\begin{aligned} \epsilon_0^n(y) &:= L_n[x \mapsto 1](y) - 1, \\ \epsilon_1^n(y) &:= L_n[x \mapsto x](y) - y, \\ \epsilon_2^n(y) &:= L_n[x \mapsto x^2](y) - y^2. \end{aligned}$$

Per l'ipotesi, possiamo prendere \tilde{n} sufficientemente grande, in modo che per $n \geq \tilde{n}$ si abbia $\|\epsilon_i^n\|_\infty < \varepsilon$ e anche $\|\epsilon_i^n\|_\infty < \delta^2 \varepsilon$ per $i = 0, 1, 2$ (la versione con il δ^2 ci servirà per far "sparire" δ nell'ultimo passaggio della (22)).

Ora, fissiamo $a \in K$ e per ogni $n \geq \tilde{n}$ scriviamo la stima

$$\begin{aligned} |L_n[f](a) - f(a)| &\leq |L_n[f](a) - f(a)L_n[x \mapsto 1](a)| + |f(a)L_n[x \mapsto 1](a) - f(a)| \\ &= |L_n[x \mapsto f(x) - f(a)](a)| + |\epsilon_0(a)f(a)| \\ &\leq L_n[x \mapsto |f(x) - f(a)](a) + M\varepsilon \end{aligned}$$

(abbiamo usato la (21) nell'ultimo passaggio). Lavoriamo separatamente ora sul primo dei due addendi; si ha

$$\begin{aligned} |f(x) - f(a)| &\leq \begin{cases} \varepsilon & \text{per } |x - a| < \delta \\ M & \text{per } |x - a| > \delta \end{cases} \\ &\leq \begin{cases} \varepsilon & \text{per } |x - a| < \delta \\ \frac{M}{\delta^2}(x - a)^2 & \text{per } |x - a| > \delta \end{cases} \\ &\leq \varepsilon + \frac{M}{\delta^2}(x - a)^2 \end{aligned}$$

quindi per la (20) e la (21), abbiamo

$$\begin{aligned}
L_n[x \mapsto |f(x) - f(a)|](a) &\leq L_n[x \mapsto \varepsilon + \frac{M}{\delta^2}(x - a)^2] \\
&= \left(\varepsilon + \frac{a^2 M}{\delta^2} \right) L_n[x \mapsto 1](a) + \frac{M}{\delta^2} L_n[x \mapsto x^2](a) - \frac{2aM}{\delta^2} L_n[x \mapsto x](a) \\
&= \left(\varepsilon + \frac{a^2 M}{\delta^2} \right) (1 + \epsilon_0(a)) + \frac{M}{\delta^2} (a^2 + \epsilon_2(a)) - \frac{2aM}{\delta^2} (a + \epsilon_1(a)) \\
&= \varepsilon + \frac{a^2 M}{\delta^2} \epsilon_0(a) + \frac{M}{\delta^2} \epsilon_2(a) - \frac{2aM}{\delta^2} \epsilon_1(a) \\
&\leq \varepsilon + N^2 M \varepsilon + M \varepsilon + 2NM \varepsilon
\end{aligned} \tag{22}$$

La quantità a cui arriviamo, come richiesto, è della forma $C\varepsilon$, con C costante che dipende solo dalle scelte di K e f . Quindi il teorema è dimostrato. \square

Notate che nelle stime compaiono quantità che dipendono da f : quindi la successione di LPO approssima bene tutte le funzioni, ma su alcune funzioni l'approssimazione converge più lentamente che su altre.

3.1 Generalizzazioni

Teorema 14 (Teorema di Korovkin, versione m -dimensionale). *Sia $K \subseteq \mathbb{R}^m$ un compatto, e sia $C(K)$ lo spazio delle funzioni continue da K a \mathbb{R} con la norma del sup (convergenza uniforme). Sia $(L_n)_{n=1}^\infty$ una successione di LPO su $C(K)$. Se L_n approssima bene le tre funzioni:*

- $x \mapsto 1$ (la funzione costante uguale a 1),
- $x \mapsto x_i$, $i = 1, 2, \dots, m$ (la proiezione sulla i -esima componente)
- $x \mapsto x^2$ (la funzione “elevare al quadrato”),

allora L_n approssima bene tutte le funzioni di $C(K)$.

Teorema 15 (Teorema di Korovkin, versione complessa). *Sia $\mathcal{B} = \{f : [0, 2\pi] \rightarrow \mathbb{C}, f(0) = f(2\pi)\}$ (funzioni complesse continue e 2π -periodiche), con la norma del sup (convergenza uniforme). Sia $(L_n)_{n=1}^\infty$ una successione di LPO su \mathcal{B} . Se L_n approssima bene le tre funzioni:*

- $x \mapsto 1$,
- $x \mapsto e^{ix}$,
- $x \mapsto e^{-ix}$,

allora L_n approssima bene tutte le funzioni di \mathcal{B} .

Rispetto alla versione 1D, cambia solo l'insieme delle funzioni su cui richiediamo la convergenza tra le ipotesi (il cosiddetto *test di Korovkin*). Anche la dimostrazione è quasi identica — potete provare a farla per esercizio.

3.2 Polinomi di Bernstein

I *polinomi di Bernstein* sono la successione di LPO su $K = [0, 1]$ definita da

$$B_n[f](x) := \sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right)$$

Esercizio 16. *Verifica che i polinomi di Bernstein sono LPO.*

3.3 Un'applicazione: dimostrazione del teorema di approssimazione di Weierstrass

Ci servirà nel seguito questo lemmetto con i binomiali.

Lemma 17 (formule “in-and-out” per i binomiali). *Valgono le seguenti identità.*

1. $\binom{n}{k} = \frac{n}{k} \binom{n-1}{k-1}$
2. $\binom{n}{k} = \frac{n(n-1)}{k(k-1)} \binom{n-2}{k-2}$

Dimostrazione. Applica le definizioni. . . □

Teorema 18 (di approssimazione di Weierstrass, caso 1D). *Sia $K \subseteq \mathbb{R}$ un compatto, e $f \in C(K)$. Allora, per ogni $\varepsilon > 0$ esiste un polinomio $p = p_{f,\varepsilon}$ tale che*

$$\|f - p\|_\infty \leq \varepsilon.$$

Dimostrazione. Ci basta provare il teorema per $K = [0, 1]$, poi con qualche semplice trasformazione del dominio possiamo estenderlo a tutti gli altri compatti (come?).

Mostreremo che

$$\|f - B_n[f]\|_\infty \rightarrow 0 \quad \text{quando } n \rightarrow \infty. \quad (23)$$

Quindi basta prendere $p = B_n[f]$ per un n sufficientemente grande. La (23) è la tesi del teorema di Korovkin per i B_n , quindi ci basta dimostrare che sono soddisfatte le ipotesi, cioè che $\|f - B_n[f]\|_\infty \rightarrow 0$ per le tre funzioni di test $x \mapsto 1$, $x \mapsto x$, $x \mapsto x^2$

- Per $x \mapsto 1$:

$$B_n[x \mapsto 1](y) = \sum_{k=0}^n \binom{n}{k} y^k (1-y)^{n-k} = 1,$$

quindi l'errore è costantemente uguale a 0.

- Per $x \mapsto x$:

$$\begin{aligned}
B_n[x \mapsto x](y) &= \sum_{k=0}^n \binom{n}{k} y^k (1-y)^{n-k} \frac{k}{n} \\
&= \sum_{k=1}^n \binom{n}{k} y^k (1-y)^{n-k} \frac{k}{n} \\
&= \sum_{k=1}^n \binom{n-1}{k-1} y^k (1-y)^{n-k} \\
&= \sum_{h=0}^{n-1} \binom{n-1}{h} y^{h+1} (1-y)^{n-1-h} = y
\end{aligned}$$

(abbiamo usato la in-and-out formula e cambiato indice $h := k - 1$)
quindi di nuovo l'errore è sempre nullo.

- Per $x \mapsto x^2$: usiamo l'identità

$$\frac{k^2}{n} = \frac{1}{n} \frac{k}{n} + \frac{n-1}{n} \frac{k(k-1)}{n(n-1)},$$

le in-and-out formulas, e il risultato del punto precedente:

$$\begin{aligned}
B_n[x \mapsto x^2](y) &= \sum_{k=0}^n \binom{n}{k} y^k (1-y)^{n-k} \frac{k^2}{n} \\
&= \frac{1}{n} \left(\sum_{k=0}^n \binom{n}{k} y^k (1-y)^{n-k} \frac{k}{n} \right) \\
&\quad + \frac{n-1}{n} \left(\sum_{k=0}^n \binom{n}{k} y^k (1-y)^{n-k} \frac{k(k-1)}{n(n-1)} \right) \\
&= \frac{1}{n} y + \frac{n-1}{n} \left(\sum_{h=0}^{n-2} \binom{n-2}{h} y^{h+2} (1-y)^{n-2-h} \right) = \frac{1}{n} y + \frac{n-1}{n} y^2
\end{aligned}$$

(stavolta abbiamo fatto il cambio di variabile $h := k - 2$). Quindi $B_n[x \mapsto x^2]$ converge uniformemente a $x \mapsto x^2$ (come ci serviva per provare che valgono le ipotesi del teorema di Korovkin), anche se stavolta c'è un errore di $O(\frac{1}{n})$. Del resto non potevamo aspettarci che l'errore fosse zero anche stavolta: difatti, se nella dimostrazione del teorema di Korovkin si ha $\epsilon_i^n = 0$ per ogni $n \dots$

Quindi i polinomi di Bernstein soddisfano il teorema di Korovkin, e in particolare ci forniscono l'approssimazione che cercavamo sopra. \square

3.4 Caso multidimensionale

Il teorema di Weierstrass funziona pari pari in dimensione maggiore:

Teorema 19 (di approssimazione di Weierstrass). *Sia $K \subseteq \mathbb{R}^m$ un compatto, e $f \in C(K)$. Allora, per ogni $\varepsilon > 0$ esiste un polinomio $p = p_{f,\varepsilon}$ tale che*

$$\|f - p\|_\infty \leq \varepsilon.$$

Dimostrazione. Se $K = [0, 1]^m$, funziona lo stesso trucco di sopra. Definiamo i polinomi di Bernstein m -dimensionali come

$$B_n^{(m)}[f](y_1, \dots, y_n) := \sum_{\substack{k_1 = 1, \dots, n \\ k_2 = 1, \dots, n \\ \vdots \\ k_m = 1, \dots, n}} \binom{n}{k_1} y_1^{k_1} (1-y_1)^{n-k_1} \binom{n}{k_2} y_2^{k_2} (1-y_2)^{n-k_2} \dots \binom{n}{k_m} y_m^{k_m} (1-y_m)^{n-k_m} f\left(\frac{k_1}{n}, \frac{k_2}{n}, \dots, \frac{k_m}{n}\right)$$

Con un po' di lavoro si riesce a provare che sono un LPO (facile) e che soddisfano le ipotesi del teorema di Korovkin (non è difficile come sembra: si riutilizzano diverse volte i calcoli fatti per il caso 1D ...), quindi concludiamo come sopra.

Stavolta però non è banale estendere il risultato da $K = [0, 1]^m$ a un qualunque altro compatto. È facile estenderlo con un cambio di variabile a qualunque “cubo” $[-t, t]^m$, ma per passare a un compatto qualsiasi ora serve usare anche questo risultato (con $H = [-t, t]^m$ un cubo “sufficientemente grande”):

Teorema 20 (caso \mathbb{R}^m del teorema di estensione di Tietze). *Sia $H \subseteq \mathbb{R}^m$ compatto, $K \subseteq H$ compatto in H , $f : K \rightarrow \mathbb{R}$ continua. Allora esiste $\tilde{f} : H \rightarrow \mathbb{R}$ continua che estende f .*

Il risultato è un caso particolare di un teorema più generale di topologia; non conosco una dimostrazione rapida e indolore di questo teorema nel caso \mathbb{R}^m che serve a noi².

□

Riferimenti bibliografici

[1] Bini, Bevilacqua, Capovani, Menchi. *Metodi Numerici*, Zanichelli 1992.

²Quando l'ho chiesto a un paio di analisti mi hanno assicurato che “ma sì, si dovrebbe fare con un po' di lavoro, considerando il modulo di continuità di f su K , che è uniformemente continua, per estenderla localmente e usando la compattezza...”

- [2] Bruno Iannazzo, *Dispense di Analisi Numerica II (versione preliminare)*, che mi ha spedito il prof. Bini via e-mail e che non riesco a trovare in giro su internet.
- [3] Stoer, Bulirsch, *Introduzione all'Analisi Numerica (2)*, Zanichelli 1975.
- [4] Quarteroni, Valli, *Numerical Approximation of Partial Differential Equations*, Springer series in Computational Mathematics, 1997.