

LEZIONE 6: CONDIZIONAMENTO DELLE FUNZIONI, PROPAGAZIONE DEGLI ERRORI

In questa lezione ci occuperemo di due aspetti importanti del calcolo approssimato, tramite esempi: l'effetto degli errori sulla variabile x e degli errori di arrotondamento nel calcolo dei valori di una funzione $f(x)$, e l'effetto degli errori di arrotondamento in un algoritmo iterativo che costruisce i valori di una successione convergente alla quantità da approssimare.

Cominciamo col calcolo di funzioni, visto che in pratica tutte le quantità utilizzate nel calcolo sono approssimate, possiamo innanzitutto porci la seguente domanda: qual è l'effetto sul valore di f di un errore sulla variabile indipendente x ?

Come sempre, siamo maggiormente interessati agli errori relativi. Consideriamo una funzione $f : I \mapsto \mathbb{R}$ dove I è un intervallo e $x, \tilde{x} \in I$ dove $\tilde{x} \approx x, x \neq 0$ con errore relativo

$\varepsilon_x = \frac{|x - \tilde{x}|}{|x|}$ cerchiamo di stimare $\varepsilon_{f(x)}$

$$\varepsilon_{f(x)} = \frac{|f(x) - f(\tilde{x})|}{|f(x)|}, f(x) \neq 0 \quad (1)$$

supponendo f derivabile in I possiamo utilizzare la formula di Taylor al primo ordine centrata in x

$$f(\tilde{x}) \approx f(x) + f'(x) \cdot (\tilde{x} - x) \quad (2)$$

da cui ricaviamo

$$f(\tilde{x}) - f(x) \approx f'(x) \cdot (\tilde{x} - x) \iff \frac{|f(\tilde{x}) - f(x)|}{|f(x)|} \approx \frac{|f'(x)|}{|f(x)|} \cdot |\tilde{x} - x| = \frac{|f'(x)| \cdot |x|}{|f(x)|} \cdot \frac{|\tilde{x} - x|}{|x|} \quad (3)$$

che possiamo riassumere con $\varepsilon_{f(x)} \approx \text{cond}f(x) \varepsilon_x$ dove

$$\text{cond}f(x) = \frac{|f'(x)| \cdot |x|}{|f(x)|} \quad (4)$$

è l'INDICE DI CONDIZIONAMENTO di f in x ed è la quantità che permette di misurare la risposta di f ad errori su x .

Se l'indice di condizionamento è grande, la funzione risulta instabile (o "mal condizionata"), nel senso che l'errore sulla variabile viene amplificato.

Questo è il primo esempio che incontriamo di "problema instabile" (o "problema mal condizionato") che si può definire in modo empirico come un problema in cui "piccoli errori sui dati possono portare ad errori grandi sulla soluzione".

La stabilità del problema è un concetto che va tenuto ben distinto da quello di stabilità di un algoritmo che ne calcola la soluzione, come vedremo negli esempi seguenti.

Cominciamo con un esempio che ben conosciamo:

$$f(x) = \frac{(1+x) - 1}{x}, x \neq 0 \quad (5)$$

Chi è questa funzione? E' la funzione costante 1, che ovviamente risponde in modo ottimale agli errori su x , visto che ovunque vale 1, quindi $\varepsilon_{f(x)} = 0 \forall x, \tilde{x}$ e in effetti $f'(x) = 0 \forall x \Rightarrow \text{cond}f(x) = 0 \forall x$

Sappiamo però anche che in aritmetica floating-point il calcolo di f è instabile se usiamo la formula scritta sopra per $|x|$ piccolo.
Infatti per $x \rightarrow 0$ il peso

$$w_1 = \frac{|1+x|}{|(1+x)-1|} = \frac{|1+x|}{|x|} \rightarrow +\infty \quad (6)$$

nella sottrazione $(1+x)-1$ più piccolo è $|x|$, più la formula è instabile a causa della perdita di precisione nella sottrazione (sappiamo ad esempio che in Matlab $f(10^{-15})=1.11\dots$ con un errore relativo $>11\%$ spiegato dal fatto che $w_1 \approx 10^{-15}$)

ma ATTENZIONE: è la formula, cioè l'espressione di calcolo, cioè l'algoritmo, ad essere instabile invece la funzione è perfettamente stabile e potremmo calcolarla con un algoritmo stabile, ovvero semplicemente $f(x) = 1$

Chiaramente si tratta di un caso limite, adesso vedremo altri esempi meno estremi

- $f(x) = \frac{1}{x}, x \neq 0$

$$\text{cond}f(x) = \frac{\left| -\frac{1}{x^2} \right| \cdot |x|}{\left| \frac{1}{x} \right|} = 1 \quad (7)$$

questa è una funzione stabile $\forall x$, in effetti lo sappiamo già, è l'operazione di reciproco

- $f(x) = 1 - x$

$$\text{cond}f(x) = \frac{|(-1)| \cdot |x|}{|1-x|} = \frac{|x|}{|1-x|} \quad (8)$$

ora $\text{cond} f(x) \rightarrow +\infty, x \rightarrow 1$ quindi la funzione è instabile per $x \approx 1$.

Di nuovo non è una sorpresa, si tratta di una sottrazione instabile per $x \approx 1$, qui $\text{cond}f(x) = w_2$

- $f(x) = 1 - \sqrt{1-x^2}, 0 < |x| \leq 1$

$$f'(x) = \frac{-1}{2\sqrt{1-x^2}} \cdot (-2x) = \frac{x}{\sqrt{1-x^2}} \quad (9)$$

(derivata di una funzione composta)

e quindi

$$\text{cond}f(x) = \frac{|x|}{\sqrt{1-x^2}} \cdot \frac{|x|}{|f(x)|} = \quad (10)$$

$$\begin{aligned} &= \frac{x^2}{\sqrt{1-x^2}} \cdot \frac{1}{1-\sqrt{1-x^2}} = \frac{x^2}{\sqrt{1-x^2}} \cdot \frac{1+\sqrt{1-x^2}}{(1+\sqrt{1-x^2})(1-\sqrt{1-x^2})} \\ &= \frac{x^2(1+\sqrt{1-x^2})}{\sqrt{1-x^2}(1-(1-x^2))} = \frac{1+\sqrt{1-x^2}}{\sqrt{1-x^2}} \end{aligned} \quad (11)$$

Abbiamo che

$$\text{cond}f(x) \rightarrow 2, x \rightarrow 0$$

quindi f è ben condizionata per $x \approx 0$.

D'altra parte il calcolo con l'espressione

$$f(x) = 1 - \sqrt{1 - x^2}$$

è instabile per $x \approx 0$ a causa della sottrazione esterna $1 - \sqrt{1 - x^2}$ mentre la sottrazione sotto radice è stabile.

Il motivo è che nella sottrazione interna i due numeri si allontanano in termini relativi per $x \rightarrow 0$, mentre in quella esterna si avvicinano, infatti

$$w_2 = \frac{\sqrt{1 - x^2}}{1 - \sqrt{1 - x^2}} = \frac{\sqrt{1 - x^2} \cdot (1 + \sqrt{1 - x^2})}{x^2} \rightarrow +\infty \quad (12)$$

per $x \rightarrow 0$.

Cosa stiamo vedendo? La funzione è stabile per $x \approx 0$, ma l'algoritmo scelto per calcolarla non lo è (a causa della sottrazione esterna). Dai conti fatti con le equazioni di 2° grado, sappiamo però che possiamo stabilizzare la formula per il calcolo di f

$$f(x) = 1 - \sqrt{1 - x^2} = \frac{(1 - \sqrt{1 - x^2}) \cdot (1 + \sqrt{1 - x^2})}{1 + \sqrt{1 - x^2}} = \frac{1 - (1 - x^2)}{1 + \sqrt{1 - x^2}} = \frac{x^2}{1 + \sqrt{1 - x^2}} \quad (13)$$

Con questa seconda formula il calcolo in aritmetica di macchina diventa stabile per $x \approx 0$, perché tutte le operazioni coinvolte (potenza, divisione, addizione e sottrazione sotto radice) sono stabili.

La risposta di una funzione ad errori sulla variabile è una proprietà intrinseca della funzione, potremmo dire una proprietà "strutturale", che non dipende da "come" è scritta ma solo dall'indice di condizionamento $\text{cond}f(x)$.

D'altra parte "come" è scritta significa qual è l'*algoritmo* di calcolo.

Abbiamo visto che per una funzione ben condizionata possono esserci algoritmi stabili e algoritmi instabili, sta a noi sceglierne uno stabile.

Usando di nuovo una definizione non rigorosa ma empirica possiamo dire che un algoritmo è instabile quando "piccoli errori introdotti durante il processo di calcolo (ad esempio gli arrotondamenti in aritmetica floating-point) possono portare ad errori grandi sui risultati del calcolo".

Siccome nel calcolo di $f(x)$ un errore su x (misura sperimentale, arrotondamento,...) cioè sul dato del problema, è uno degli errori introdotti nel processo di calcolo, ci si aspetta che se esiste un algoritmo stabile la funzione sia ben condizionata, cioè che $\text{cond}f(x)$ non sia grande. È proprio quello

che accade nell'esempio precedente, dove la scrittura $f(x) = \frac{x^2}{1 + \sqrt{1 - x^2}}$ è stabile per $x \approx 0$, infatti

$\text{cond}f(x) \approx 2$ (cioè non è grande) per $x \approx 0$.

Viceversa, se una funzione è instabile (mal condizionata) ci aspettiamo che qualsiasi algoritmo di calcolo "erediti" questa instabilità.

È bene ribadire che tutti questi concetti (non facili) li stiamo esprimendo in modo empirico ed essenzialmente qualitativo, per non entrare nell'ambito di teorie generali della stabilità che richiedono strumenti

matematici di livello superiore a quelli che usiamo in questo corso.

Per concludere con l'esempio che $f(x) = 1 - \sqrt{1 - x^2}$ è ben condizionata per $|x| \approx 0$ ma è mal condizionata per $|x| \approx 1$ visto che $\text{cond}f(x) = \frac{1 + \sqrt{1 - x^2}}{\sqrt{1 - x^2}} \rightarrow +\infty$ per $|x| \rightarrow 1$.

Nella seconda parte della lezione facciamo un passo ulteriore verso una metodologia caratteristica del calcolo numerico, ovvero il calcolo di una successione che converge alla quantità che vogliamo approssimare, tramite un algoritmo iterativo.

Il problema che affrontiamo è il calcolo di π in un sistema floating-point a 64 bit (ad esempio in Matlab) con un errore dell'ordine della precisione di macchina (cioè con almeno 15 cifre corrette nell'interfaccia in base 10).

Consideriamo 2 successioni convergenti a π :

la prima si può ottenere dal fatto che la cosiddetta serie armonica (con esponente 2) converge e ha come somma

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6} \quad (14)$$

(risultato che accettiamo senza dimostrarlo).

Allora, detta $S_n = \sum_{k=1}^n \frac{1}{k^2}$ la somma parziale n-esima, visto che $\sum_{k=1}^{\infty} \frac{1}{k^2} = \lim_{n \rightarrow \infty} S_n = \frac{\pi^2}{6}$ si ottiene

$$\lim_{n \rightarrow \infty} \sqrt{6 \cdot S_n} = \pi$$

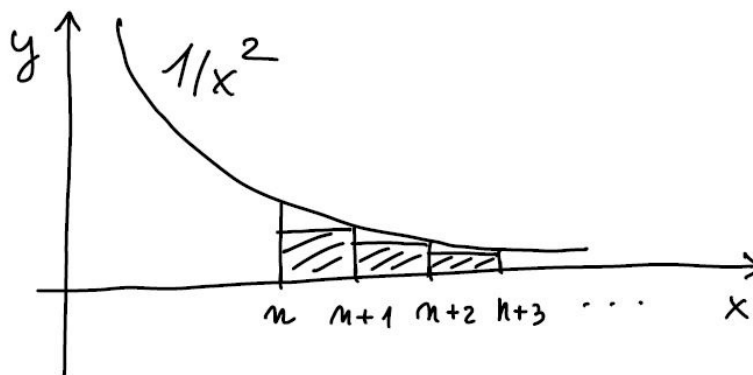
Osserviamo che il calcolo di S_n in aritmetica di macchina è stabile, perché coinvolge solo operazioni stabili (elevamento al quadrato, reciproco, addizioni essendo la serie a termini positivi) però la successione $\sqrt{6 \cdot S_n}$ è in pratica inutilizzabile per calcolare π con grande precisione, perché la successione S_n converge troppo lentamente.

Infatti si vede con il criterio di confronto integrale che il resto $\frac{\pi^2}{6} - S_n = \sum_{k=n+1}^{\infty} \frac{1}{k^2}$ si può maggiorare

nel modo seguente

$$\sum_{k=n+1}^{\infty} \frac{1}{k^2} < \int_n^{\infty} \frac{1}{x^2} dx = \frac{1}{n} \quad (15)$$

come è chiaro dal disegno schematico (non in scala)



dove si vede che l'arco che sta sotto la curva $y = 1/x^2$ tra n e ∞ (cioè $\int_n^\infty \frac{1}{x^2} dx$) è maggiore della somma delle aree dei rettangolini che hanno base 1 e altezza $1/k^2$ con $k = n+1, n+2, \dots$ cioè $\sum_{k=n+1}^\infty \frac{1}{k^2}$. Quindi l'errore che si commette approssimando $\pi^2/6$ con S_n si può stimare con $\frac{\pi^2}{6} - S_n < \frac{1}{n}$

Questo significa che per approssimare $\frac{\pi^2}{6}$ con un errore $\leq \varepsilon_M$ (precisione di macchina) basta che $\frac{\pi^2}{6} - S_n < \frac{1}{n} \leq \varepsilon_M$ cioè $n \geq \varepsilon_M^{-1} \approx 10^{16}$, quindi bisognerebbe sommare 10^{16} (10 milioni di miliardi) di termini della serie!

D'altra parte in aritmetica di macchina ad un certo punto la somma S_n diventerebbe costante, perché i termini che vanno sommati diventerebbero troppo piccoli e si comporterebbero come elementi neutri, impedendo di raggiungere la precisione richiesta (accade da $n \approx 10^8$ in poi).

Considerando invece la successione definita "per ricorrenza" da

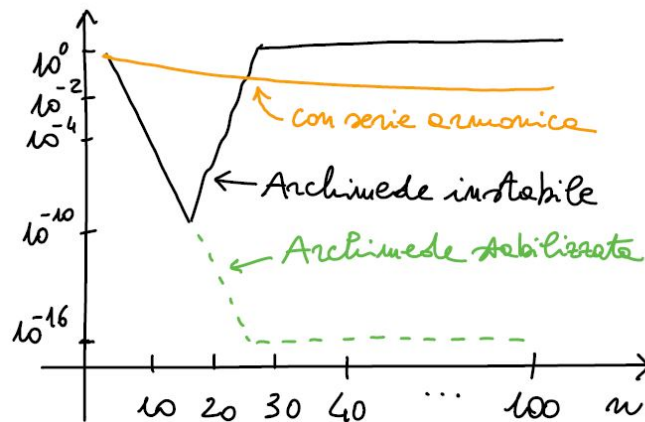
$$x_0 = 2$$

$$x_{n+1} = 2^{n-\frac{1}{2}} \cdot \sqrt{1 - \sqrt{1 - (4^{1-n})x_n^2}}$$

$$n = 2, 3, 4, \dots$$

detta "successione di Archimede" si può dimostrare con metodi geometrici (legati all'approssimazione del cerchio con poligoni regolari) che $\lim_{n \rightarrow \infty} x_n = \pi$. Se calcoliamo x_n in Matlab con un semplice ciclo

"for" e plottiamo l'errore relativo $r_n = \frac{|\tilde{x}_n - \pi|}{\pi}$ in SCALA LOGARITMICA, cioè plottiamo $\log_{10}(r_n)$, otteniamo un grafico tipo quello disegnato di seguito in nero



(dove abbiamo interpolato i valori discreti per una migliore visualizzazione). Si vede che l'errore decresce linearmente fino a circa 10^{-10} e poi comincia a crescere, cioè la successione \tilde{x}_n calcolata con lo schema descritto sopra in aritmetica di macchina da un certo n in poi comincia ad allontanarsi da π .

Si noti che fino a quel punto l'errore decade rapidamente, infatti un decadimento lineare in scala logaritmica significa che $\log_{10}(r_n) = -\alpha n + \beta$ con $\alpha > 0$ cioè in scala normale $r_n = 10^{-\alpha n + \beta} = \delta \theta^n$ con $\delta = 10^\beta$ e $\theta = 10^{-\alpha} < 1$ che è un decadimento esponenziale (tanto più veloce quanto più grande è α).

E' chiaro che l'algoritmo è instabile e che tale instabilità è in grado da un certo n in poi di distruggere la

convergenza della successione (si osservi però che comunque la convergenza finchè c'è è rapida e permette di arrivare ad approssimare π con un errore $\approx 10^{-10}$ in meno di 20 iterazioni).

A cosa è dovuta l'instabilità? All'interno della radice quadrata esterna ci sono 2 sottrazioni, $1 - \sqrt{1 - \alpha_n}$, con $\alpha_n = 4^{1-n}x_n^2$.

Visto che α_n è un infinitesimo, $\alpha_n \rightarrow 0, n \rightarrow \infty$ (perchè $4^{-n} \rightarrow 0$ e $x_n^2 \rightarrow \pi^2, n \rightarrow \infty$) la sottrazione $1 - \alpha_n$ è stabile (i numeri si allontanano in termini relativi al crescere di n).

Ma per lo stesso motivo la sottrazione $1 - \sqrt{1 - \alpha_n}$ diventa sempre più instabile al crescere di n , perchè $\sqrt{1 - \alpha_n} \rightarrow 1, n \rightarrow \infty$ e la convergenza a 1 è rapida ($|1 - \sqrt{1 - \alpha_n}|$ va a zero come 4^{-n}).

Possiamo calcolare il peso w_2 di questa sottrazione, visto che $\sqrt{1 - \alpha_n}$ è sicuramente affetto da errore (dell'ordine della precisione di macchina, visto che $\sqrt{}$ è una funzione predefinita calcolata alla precisione di macchina).

Posto $x = 1, y = -\sqrt{1 - \alpha_n}$ si ha

$$\begin{aligned} w_2 &= \frac{|y|}{|x + y|} = \frac{\sqrt{1 - \alpha_n}}{1 - \sqrt{1 - \alpha_n}} = \frac{\sqrt{1 - \alpha_n}(1 + \sqrt{1 - \alpha_n})}{(1 - \sqrt{1 - \alpha_n})(1 + \sqrt{1 - \alpha_n})} \\ &= \frac{\sqrt{1 - \alpha_n}(1 + \sqrt{1 - \alpha_n})}{X - (X - \alpha_n)} \approx \frac{2}{\alpha_n} = 2 \cdot \frac{4^{n-1}}{x_n^2} \approx \frac{1}{2\pi^2} 4^n \end{aligned} \quad (16)$$

Quindi il peso w_2 è fattore di amplificazione che cresce esponenzialmente come 4^n . Possiamo stimare l'errore $e_n = |\pi - \widetilde{x}_n|$ in questo modo:

$$e_n = |\pi - x_n + x_n - \widetilde{x}_n| \leq |\pi - x_n| + |x_n - \widetilde{x}_n| \quad (17)$$

dove $|\pi - x_n|$ è per la convergenza e $|x_n - \widetilde{x}_n|$ è per la stabilità, e dove l'errore $|\pi - x_n|$ è legato alla convergenza teorica del metodo (in generale, dire che $\lim_{x \rightarrow \infty} x_n = l$ con l limite finito è equivalente a dire che l'errore $|x_n - l| \rightarrow 0, n \rightarrow \infty$), mentre l'errore $|x_n - \widetilde{x}_n|$ è legato alla stabilità dell'algoritmo che implementa il metodo.

Dall'analisi fatta, $|x_n - \pi| = c\theta^n$, con $0 < \theta < 1$ mentre $|\widetilde{x}_n - x_n| \lesssim c'4^n \varepsilon_M$ con c e c' costanti. Quindi

$$e_n \lesssim c\theta^n + c'4^n \varepsilon_M \quad (18)$$

da cui ci aspettiamo che finché $c'4^n \varepsilon_M < c\theta^n$ si continui a vedere la convergenza, ma da quel punto in poi il termine esponenziale crescente comincia a sovrastare e in pratica fa allontanare la successione calcolata dal limite.

Si osservi che ad un certo punto l'errore diventa costante: questo accade quando α_n è così piccolo che $1 \ominus \alpha_n = 1$ e quindi $\widetilde{x}_n = 0$ forzando un errore del 100%.

È però possibile stabilizzare il calcolo di x_n eliminando la sottrazione instabile, con lo stesso trucco utilizzato per le equazioni di 2° grado.

Il cuore del problema è infatti il calcolo di $1 - \sqrt{1 - \alpha_n}$:

$$1 - \sqrt{1 - \alpha_n} = \frac{(1 - \sqrt{1 - \alpha_n})(1 + \sqrt{1 - \alpha_n})}{1 + \sqrt{1 - \alpha_n}} = \frac{1 - (1 - \alpha_n)}{1 + \sqrt{1 - \alpha_n}} = \frac{\alpha_n}{1 + \sqrt{1 - \alpha_n}} \quad (19)$$

questa riscrittura è stabile in aritmetica floating-point, perchè sono coinvolte solo operazioni stabili.

Possiamo allora riscrivere l'iterazione in forma stabile

$$x_{n+1} = 2^{n-\frac{1}{2}} \sqrt{1 - \sqrt{1 - \alpha_n}} = 2^{n-\frac{1}{2}} \sqrt{\frac{\alpha_n}{1 + \sqrt{1 - \alpha_n}}} = \frac{\sqrt{2}x_n}{\sqrt{1 + \sqrt{1 - (4^{1-n})x_n^2}}} \quad (20)$$

Nell'aritmetica dei reali le due formule sono coincidenti, mentre in aritmetica floating-point la seconda è stabile mentre la prima è fortemente instabile. Nel grafico degli errori in scala logaritmica disegnato sopra, la successione stabilizzata

$$y_2 = 2, y_{n+1} = \frac{\sqrt{2}y_n}{\sqrt{1 + \sqrt{1 - 4\frac{1-n}{y_n^2}}}}, n \geq 2 \quad (21)$$

calcolata in aritmetica di macchina, diciamola \widetilde{y}_n , fa un errore sostanzialmente coincidente con quello di \widetilde{x}_n finché per quest'ultima si innesca l'instabilità, per poi continuare a diminuire secondo la stessa retta (cioè esponenzialmente in scala normale) fino ad attestarsi all'ordine della precisione di macchina, di sotto alla quale ovviamente non si può andare in aritmetica di macchina, perché cominciano a dominare gli errori di arrotondamento, come si vede dalla stima

$$|\pi - \widetilde{y}_n| \leq |\pi - y_n| + |y_n - \widetilde{y}_n| \lesssim c\theta^n + c''\varepsilon_M \text{ con } c'' \approx \pi \quad (22)$$

per n abbastanza grande

$c\theta^n \leq c''\varepsilon_M$ e quindi ci aspettiamo che il termine $c''\varepsilon_M$ diventi dominante nell'errore.