Errori nel calcolo di una funzione

Calcolo Numerico - Ing. Inf. - Lezione 2

Outline

Sia $f : \mathbb{R}^n \longmapsto \mathbb{R}$ una funzione ad n variabile e a valori reali

Si vuole calcolare $f(P) = f(x_1, x_2, \ldots, x_n)$ in un punto assegnato $P_0 = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$

In generale, non si ottiene il valore $f(P_0)$ cercato, a causa delle approssimazioni che si introducono

I contribuiti all'errore sono di due tipi:

1. la funzione $f(P)$ deve essere tradotta in un algoritmo di calcolo dato dalla funzione $f_a(P)$
2. in generale non è possibile rappresentare esattamente le coordinate del punto P_0 per cui si passa ad un punto P_1 le cui componenti sono numeri di macchina (per esempio, sia $P_0 = (e, \sqrt{3})$).

Si conclude che, volendo calcolare $f(P_0)$ siamo costretti a calcolare il valore $f_a(P_1)$

In altre parole, approssimiamo il valore della funzione nel punto assegnato cambiando la funzione ed il punto in cui si vorrebbe calcolarla

Outline

Sia dato il punto $P_0 \in \mathbb{R}^n$ di coordinate $x_i^{(0)}$, $i = 1, 2, \ldots, n$, nel quale si vuole calcolare la funzione $f(P)$

Si considera l'insieme

$$D = \{\, P \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i,\ i = 1, 2, \ldots, n\}$$

dove $a_i, b_i \in \mathbb{R}$ e verificano $a_i \leq x_i^{(0)} \leq b_i$, $i = 1, 2, \ldots, n$

D si dice **insieme di indeterminazione** del punto P_0.

Il valore cercato $f(P_0)$ viene sostituito dal valore calcolato $f_a(P_1)$ dove $P_1 \in D$ con coordinate $x_i^{(1)}$, $i = 1, 2, \ldots, n$

Errore Assoluto

L'**errore assoluto** è dato da

$$\delta_f \;=\; f_a(P_1) \;-\; f(P_0)$$

Di questo errore è possibile dare una stima

Nella espressione di δ_f si somma e si sottrae $f(P_1)$ ottenendo

$$\delta_f = f_a(P_1) - f(P_1) + f(P_1) - f(P_0)$$

Da questa uguaglianza si evidenzia quello che abbiamo esposto a parole e cioè

$$\delta_f = \underbrace{f_a(P_1) - f(P_1)}_{\delta_a} + \underbrace{f(P_1) - f(P_0)}_{\delta_d}$$

Il termine $\delta_a = f_a(P_1) - f(P_1)$ si chiama **errore assoluto algoritmico** e dipende dall'aver sostituito la funzione $f(P)$ con la funzione $f_a(P)$

Il termine $\delta_d = f(P_1) - f(P_0)$ si chiama **errore assoluto trasmesso dai dati** e dipende dall'aver sostituito il punto P_0 con il punto P_1

Quindi scriviamo

$$\delta_f = \delta_a + \delta_d$$

L'errore assoluto algoritmico, una volta fissato l'algoritmo che fornisce $f_a(P)$, risulta definito e stimabile

All'errore trasmesso dai dati, nell'ipotesi che sia $f(P) \in C^1(D)$, si può dare una rappresentazione generale

Dalla formula di Taylor arrestata al primo termine e con punto iniziale P_0 si ottiene

$$f(P_1) - f(P_0) \; = \; \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_i^{(1)} - x_i^{(0)})$$

dove le derivate parziali della funzione $f(P)$ sono calcolate in un punto opportuno

Ponendo

$$\delta_{x_i} = x_i^{(1)} - x_i^{(0)}, \qquad \rho_i = \frac{\partial f}{\partial x_i}$$

si ha

$$\delta_d = \sum_{i=1}^{n} \rho_i \, \delta_{x_i}$$

I valori ρ_i sono detti **coefficienti di amplificazione** degli errori δ_{x_i}

Una limitazione per il modulo dell'errore assoluto $f_a(P_1) - f(P_0)$ è

$$|f_a(P_1) \; - \; f(P_0)| \; \leq \; E_a \; + \; E_d$$

dove si è posto

$$E_a \;\; = \;\; \text{massimo modulo dell'errore assoluto dovuto al particolare algoritmo usato}$$

$$E_d \; = \sum_{i=1}^{n} A_{x_i} \, | \, \delta_{x_i} \, |$$

con

$$A_{x_i} \; \geq \; \sup_{P \in D} \left| \frac{\partial f}{\partial x_i} \right|, \qquad i = 1, 2, \ldots, n$$

Problema Diretto – Problema Inverso

Se si conosce una stima dell'errore algoritmico e degli errori δ_{x_i} nonché le A_{x_i}, si può stabilire a posteriori un confine superiore per l'errore assoluto con cui si è calcolata la funzione nel punto desiderato; questo problema è detto **problema diretto**

Il **problema inverso** consiste nel richiedere a priori che il valore $f_a(P_1)$ sia tale che l'errore assoluto $|f_a(P_1) - f(P_0)|$ risulti minore di un valore prefissato, per cui si deve cercare sia un algoritmo $f_a(P)$ sia un opportuno punto P_1 che soddisfino la richiesta

Esempio di Problema Diretto

Si vuole calcolare la funzione

$$f(x_1, x_2) = \frac{x_1}{x_2}$$

in un punto $P_0 \in D = [1, 3] \times [4, 5]$

Si suppono di arrotondare il risultato dell'operazione alla 2^a cifra decimale e di introdurre i valori $x_1^{(0)}$ e $x_2^{(0)}$ con errori

$$|\delta_{x_1}| \leq 10^{-2}, \qquad |\delta_{x_2}| \leq 10^{-2}$$

Domanda

Quale è il massimo $|\delta_f|$?

Esempio di Problema Diretto

Arrotondando il risultato dell'operazione alla 2ª cifra decimale si ha

$$|\,\delta_a\,| \;\le\; \frac{1}{2}\,10^{-2}$$

Risultando
$$\frac{\partial}{\partial x_1} f = \frac{1}{x_2}, \qquad \frac{\partial}{\partial x_2} f = -\frac{x_1}{x_2^2}$$

si ha

$$A_{x_1} \;=\; \max_{P \in D} \left| \frac{1}{x_2} \right| \;=\; \frac{1}{4} \qquad A_{x_2} \;=\; \max_{P \in D} \left| -\frac{x_1}{x_2^2} \right| \;=\; \frac{3}{16}$$

Esempio di Problema Diretto

Conclusione

Mettendo insieme tutte le maggiorazioni ottenute risulta

$$\begin{aligned} |\delta_f| &\leq |\delta_a| + |\delta_d| \\ &\leq \frac{1}{2}10^{-2} + \frac{1}{4}10^{-2} + \frac{3}{16}10^{-2} \\ &= \frac{15}{16}10^{-2} \end{aligned}$$

Esempio di Problema Inverso

Siano $x_1^{(0)} \in [1, 2]$ e $x_2^{(0)} \in [-2, -1]$ e si consideri la funzione

$$f(x_1, x_2) = \frac{x_1}{x_2}$$

Domanda

Come si deve eseguire la divisione e con quale errore assoluto si devono introdurre x_1 e x_2 per avere

$$|\,\delta_f\,| \leq 10^{-2}\,?$$

Esempio di Problema Inverso

In questo caso si chiede come operare e come introdurre i dati affinché l'errore complessivo non superi la limitazione fissata

Essendo due i contributi dell'errore, si divide in due parti uguali

l'errore totale per cui si richiede

$$| \delta_a | \leq \frac{1}{2} 10^{-2}, \qquad | \delta_d | \leq \frac{1}{2} 10^{-2}$$

La prima limitazione ci dice che, una volta deciso come introdurre le approssimazioni di $x_1^{(0)}$ e $x_2^{(0)}$, basta arrotondare il risultato della divisione alla 2^a cifra decimale (in altre parole, calcolata la terza cifra decimale della divisione si arrotonda sulla seconda)

Esempio di Problema Inverso

L'errore assoluto trasmesso dai dati, essendo due le variabili della funzione, è la somma di due addendi maggiorabili con

$$|\,\delta_d\,| \;\leq\; A_{x_1}\,|\,\delta_{x_1}\,| + A_{x_2}\,|\,\delta_{x_2}\,|$$

Come in precedenza, si divide l'errore $|\delta_d|$ in due parti uguali

Risultando $\frac{\partial}{\partial x_1} f = \frac{1}{x_2}$, $\frac{\partial}{\partial x_2} f = -\frac{x_1}{x_2^2}$, con $D = [1,2] \times [-2,-1]$, si ha

$$A_{x_1} = \max_{P \in D} \left| \frac{1}{x_2} \right| = 1 \qquad A_{x_2} = \max_{P \in D} \left| -\frac{x_1}{x_2^2} \right| = 2$$

Esempio di Problema Inverso

Si impone quindi

$$A_{x_1} \mid \delta_{x_1} \mid \leq \frac{1}{4}\, 10^{-2}, \qquad A_{x_2} \mid \delta_{x_2} \mid \leq \frac{1}{4}\, 10^{-2}$$

ottenendo le limitazioni

$$\mid \delta_{x_1} \mid \leq \frac{1}{4}\, 10^{-2}, \qquad \mid \delta_{x_2} \mid \leq \frac{1}{8}\, 10^{-2}$$

Perché siano soddisfatte entrambe le limitazioni basta introdurre le due variabili troncando la loro rappresentazione alle 3^a cifra decimale avendo quindi

$$\mid \delta_{x_1} \mid \leq 10^{-3}, \qquad \mid \delta_{x_2} \mid \leq 10^{-3}$$

Esempio di Problema Inverso

Conclusione

Mettendo insieme tutte le maggiorazioni ottenute risulta

$$
\begin{aligned}
|\delta_f| &\leq |\delta_a| + |\delta_d| \\
&\leq \frac{1}{2}10^{-2} + 10^{-3} + 2 \cdot 10^{-3} \\
&= \frac{4}{5}10^{-2}
\end{aligned}
$$

Outline

Errore Relativo

L'**errore relativo** che si commette nel calcolo di una funzione $f(P)$ in un assegnato punto P_0 è definito da

$$\epsilon_f = \frac{f_a(P_1) - f(P_0)}{f(P_0)}$$

Si verifica facilmente che si può scrivere

$$\epsilon_f = \frac{f(P_1) - f(P_0)}{f(P_0)} + \frac{f_a(P_1) - f(P_1)}{f(P_1)} \left(1 + \frac{f(P_1) - f(P_0)}{f(P_0)}\right)$$

Indicando con

$$\epsilon_a = \frac{f_a(P_1) - f(P_1)}{f(P_1)}$$

l'**errore relativo algoritmico** e con

$$\epsilon_d = \frac{f(P_1) - f(P_0)}{f(P_0)}$$

l'**errore relativo trasmesso dai dati** si ottiene

$$\epsilon_f = \epsilon_a + \epsilon_d + \epsilon_a \, \epsilon_d$$

Nell'ultima uguaglianza si trascura il termine $\epsilon_a \, \epsilon_d$ perché (come infinitesimo) risulta di ordine superiore rispetto agli altri due addendi

L'errore relativo trasmesso dai dati è dato da

$$\epsilon_d = \frac{f(P_1) - f(P_0)}{f(P_0)} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{(x_i^{(1)} - x_i^{(0)})}{f(P_0)}$$

Definendo $\epsilon_{x_i} = \frac{x_i^{(1)} - x_i^{(0)}}{x_i^{(0)}}$, si ha

$$\epsilon_d = \sum_{i=1}^{n} \frac{x_i^{(0)}}{f(P_0)} \frac{\partial f}{\partial x_i} \epsilon_{x_i}$$

Ponendo $\gamma_i = \frac{x_i^{(0)}}{f(P_0)} \frac{\partial f}{\partial x_i}$, si ha

$$\epsilon_d = \sum_{i=1}^{n} \gamma_i \, \epsilon_{x_i}$$

dove i valori γ_i sono detti **coefficienti di amplificazione** degli errori relativi

Se i coefficienti γ_i sono tali che l'errore ϵ_d risulta dello stesso ordine degli errori ϵ_{x_i} il problema del calcolo della funzione si dice **ben condizionato** mentre si dice **mal condizionato** se a piccoli errori relativi ϵ_{x_i} corrisponde un errore ϵ_d rilevante

All'errore relativo ϵ_f concorre anche l'errore algoritmico ϵ_a
Se l'algoritmo è tale da produrre errori accettabilmente limitati nella sua applicazione, si dice **stabile**, mentre è detto **instabile** nel caso contrario

Prima di vedere un esempio di calcolo dell'errore relativo analizziamo gli errori assoluti trasmessi dai dati e gli errori relativi trasmessi dai dati nel caso delle 4 operazioni elementari

Gli errori nelle quattro operazioni

Analizzando le quattro operazioni fondamentali, per quanto riguarda gli errori trasmessi dai dati, si ottiene la tabella

operazione	δ_d	ϵ_d
$x \oplus y$	$\delta_x + \delta_y$	$\frac{x}{x+y}\,\epsilon_x + \frac{y}{x+y}\,\epsilon_y$
$x \ominus y$	$\delta_x - \delta_y$	$\frac{x}{x-y}\,\epsilon_x - \frac{y}{x-y}\,\epsilon_y$
$x \otimes y$	$y\,\delta_x + x\,\delta_y$	$\epsilon_x + \epsilon_y$
$x \oslash y$	$\frac{1}{y}\,\delta_x - \frac{x}{y^2}\,\delta_y$	$\epsilon_x - \epsilon_y$

(In questa tabella le variabili x_1, x_2 sono state sostituite con le variabili x, y per semplicità di notazione)

Gli errori nelle quattro operazioni

Osservazione 1

Dalla precedente tabella si evidenzia che le operazioni di addizione e sottrazione non danno problemi per quanto riguarda l'errore assoluto, mentre possono rendere grande l'errore relativo nel caso in cui i due termini dell'operazione siano molto vicini in valore assoluto, in quanto può accadere che i denominatori che compaiono nei coefficienti di amplificazione dell'errore relativo siano molto piccoli in valore assoluto avendo così quella che abbiamo chiamato la cancellazione

Gli errori nelle quattro operazioni

Osservazione 2

La moltiplicazione non amplifica l'errore relativo e comporta un errore assoluto che dipende dall'ordine di grandezza dei fattori.
Anche la divisione non produce amplificazione per quanto riguarda l'errore relativo, mentre l'errore assoluto diminuisce se il divisore aumenta (in valore assoluto)

Esempio

Domanda

Si vuole stimare l'errore relativo commesso nel calcolo della

funzione

$$f(x, y, z, w) = x\left(\frac{y}{z} - w\right)$$

(Anche in questo esempio utilizziamo variabili senza indice per semplicità di notazione)

Si ricorrere all'uso dei **grafi**
Si stabilisce un algoritmo di calcolo eseguendo una operazione dopo l'altra e stabilendo per ciascuna di esse l'entità degli errori relativi.

Esempio

Algoritmo

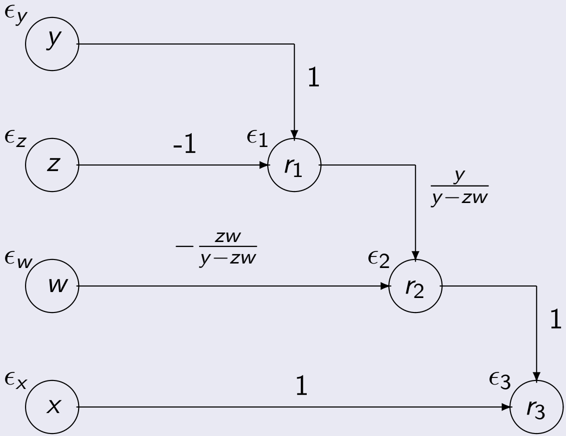
In questo esempio la sequenza delle operazioni che seguiamo è

$$r_1 = \frac{y}{z}, \qquad r_2 = r_1 - w, \qquad r_3 = x \cdot r_2$$

Il grafo della prossima slide evidenzia per ogni operazione i coefficienti di amplificazione lungo i cammini orientati. Gli errori relativi dei dati sono $\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_w$, mentre ϵ_i va inteso come l'errore relativo algoritmico per la funzione r_i

Esempio

Grafo



Esempio

Per stimare l'errore relativo totale si procede a ritroso seguendo il grafo

$$
\begin{aligned}
\epsilon_f &= \epsilon_{r_3} \\
&= \epsilon_3 + \epsilon_x + \epsilon_{r_2} \\
&= \epsilon_3 + \epsilon_x + \epsilon_2 + \frac{y}{y - zw}\epsilon_{r_1} - \frac{zw}{y - zw}\epsilon_w \\
&= \epsilon_3 + \epsilon_x + \epsilon_2 + \frac{y}{y - zw}(\epsilon_1 + \epsilon_y - \epsilon_z) - \frac{zw}{y - zw}\epsilon_w \\
&= \epsilon_3 + \epsilon_2 + \frac{y}{y - zw}\epsilon_1 + \epsilon_x + \frac{y}{y - zw}(\epsilon_y - \epsilon_z) - \frac{zw}{y - zw}\epsilon_w
\end{aligned}
$$

Esempio

Conclusione

Si ricava che l'errore relativo algoritmico è

$$\epsilon_a = \epsilon_3 + \epsilon_2 + \frac{y}{y - zw}\epsilon_1$$

mentre l'errore relativo trasmesso dai dati è

$$\epsilon_d = \epsilon_x + \frac{y}{y - zw}(\epsilon_y - \epsilon_z) - \frac{zw}{y - zw}\epsilon_w$$

Osservazione

L'errore relativo calcolato nell'Esempio dipende dall'algoritmo seguito per il calcolo della funzione $f(x, y, z, w)$
Seguendo un altro algoritmo si trova, in generale, un errore relativo diverso ma **solo ed esclusivamente** per l'errore relativo algoritmico mentre l'errore relativo trasmesso dai dati rimane lo stesso indipendentemente dall'algoritmo di calcolo seguito