

Riassunto di Calcolo Numerico

Indice

1	Introduzione	2
2	Domande del Syllabus	3
2.1	Precisione di macchina	3
2.2	Stabilità delle operazioni	4
2.3	Convergenza del metodo di bisezione	5

1 Introduzione

Questo è un breve riassunto che ho scritto dopo aver fallito per innumerevoli volte la prova scritta di calcolo numerico. Per ogni domanda non riporterò la dimostrazione esatta che bisogna scrivere all'esame, ma concetti che aiutano a capire il senso (e quindi a memorizzare gli argomenti).

Prima di leggere questo documento consiglio di rivedere la teoria di analisi, in particolare gli argomenti che vengono trattati anche nel corso di calcolo.

2 Domande del Syllabus

2.1 Precisione di macchina

Si parte da un numero reale, scritto in notazione floating point:

$$x = \text{sign}(x)(0, d_1 d_2 \dots dt \dots) \cdot b^p$$

È importante tenere a mente che la 1° cifra dopo la virgola (d_1) è diversa da zero.

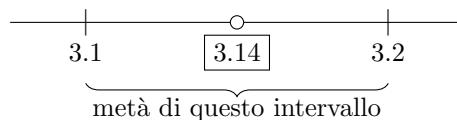
Quindi si scrive il numero arrotondato a t cifre di mantissa, e si scrive la definizione formale di arrotondamento. Poi diciamo che la precisione di macchina è l'errore relativo di arrotondamento, e scriviamo l'errore di arrotondamento:

$$\frac{|x - fl^t(x)|}{|x|}, |x| \neq 0$$

Quindi stimiamo questo rapporto in due parti: prima il numeratore.

Le cifre dei due numeri sono uguali fino alla $t-1$, cambia dalla t poi (per l'arrotondamento), e con un esempio si può vedere che questa quantità è $\leq \frac{b^{-t}}{2}$

ESEMPIO: se $x = 3.14$ e voglio arrotondare ai decimi, l'errore che compio è \leq mezzo decimo (in questo caso di 0.04):



Ora stimiamo il denominatore. Sappiamo che è una quantità positiva (c'è il valore assoluto), è $\neq 0$, e la prima cifra decimale (d_1) deve essere diversa da zero. Questo serve ad evitare che ci siano più rappresentazioni dello stesso numero.

Quindi $|x|$ è almeno (\geq) $0.1 \cdot b^p$ (non ci serve uno specifico p), ovvero $b^{-1} \cdot b^p \Rightarrow b^{p-1}$.

Siccome nella definizione di errore relativo $|x|$ "sta sotto", dobbiamo scrivere il reciproco:

$$\frac{1}{|x|} \leq \frac{1}{b^{p-1}}$$

Da notare il fatto che è cambiato il verso della disequazione.

Quindi uniamo i due pezzi, e otteniamo che:

$$\frac{|x - fl^t(x)|}{|x|} \leq \frac{\frac{b^{-t}}{2}}{b^{p-1}} = \frac{b^{p-t+1-p}}{2} = \frac{b^{1-t}}{2}$$

Che è la nostra precisione di macchina.

2.2 Stabilità delle operazioni

Questa è la più lunga di tutte le dimostrazioni, ma:

- all'esame il prof ne chiede sempre metà (solitamente moltiplicazione e divisione oppure somma algebrica)
- Non serve impararsi tutto a memoria: gran parte del testo sono operazioni algebriche. Si utilizza la disuguaglianza triangolare e si moltiplica/divide per una certa quantità (questo nel caso della somma algebrica, nella moltiplicazione si somma/sottrae).
- Si inizia dalla definizione, ovvero l'errore relativo che ho su un'operazione con numeri approssimati:

$$\varepsilon_{x \star y} = \frac{|x \star y - \tilde{x} \star \tilde{y}|}{|x \star y|}$$

Dove al posto di \star si mette una delle operazioni, e \tilde{x} , \tilde{y} (con la tilde) sono i numeri approssimati.

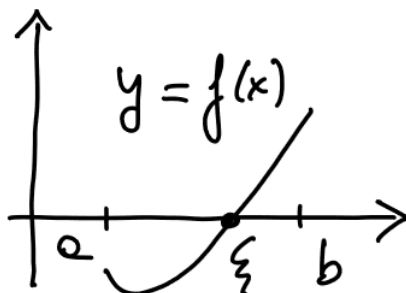
- La sottrazione (ovvero la somma algebrica nel caso in cui il segno dei due numeri è diverso) è l'unica operazione instabile (quando x e y sono vicini in termini relativi)

2.3 Convergenza del metodo di bisezione

Questa è la prima domanda in cui il prof inizia a utilizzare la lettera ξ (csi), che semplicemente indica la soluzione che stiamo cercando. Invece con x (oppure x_n) indica la "soluzione" che abbiamo trovato a una certa iterazione. In generale, più iterazioni abbiamo e più la nostra x si avvicina a ξ .

Con convergenza di un metodo (in questo caso la bisezione) si intende il fatto che, facendo iterazioni, ci avviciniamo sempre di più alla soluzione ξ , lo "zero" della funzione.

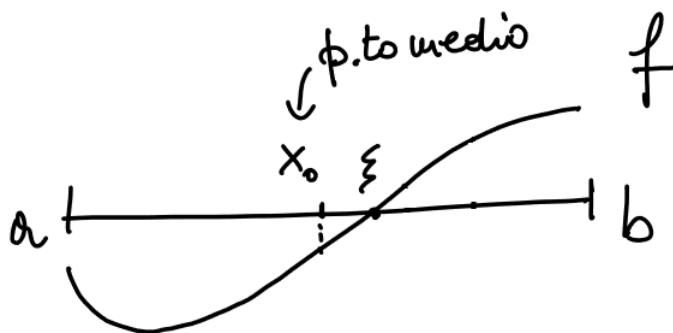
Il metodo di bisezione funziona applicando iterativamente il teorema degli zeri (o teorema di Bolzano).



Se ho una funzione continua in un intervallo $[a, b]$ e $f(a) \cdot f(b) \leq 0$ (ovvero hanno segno opposto), allora esiste un punto ξ , con $f(\xi) = 0$, che appartiene all'intervallo (a, b) .

Il metodo consiste nell'individuare il punto medio, e continuare il procedimento sulla metà dell'intervallo in cui vale la condizione che i due estremi hanno segno opposto.

Esempio:



In questa funzione abbiamo come estremi a, b . Individuiamo il punto medio x_0 , e la successiva iterazione la facciamo sul semintervallo $[a, x]$, perché $f(a)f(x_0) \leq 0$.

Il punto medio è $x_n = \frac{a_n + b_n}{2}$.

Poi si individuano tre successioni: a_n, b_n, x_n , tali che:

- $|\xi - a_n| \leq b_n - a_n = \frac{b-a}{2^n}$
- $|\xi - b_n| \leq b_n - a_n = \frac{b-a}{2^n}$
- $|\xi - x_n| \leq \frac{b_n - a_n}{2} = \frac{b-a}{2^{n+1}}$

$\frac{b-a}{2^n}$ è la lunghezza dell'intervallo $b_n - a_n$, mentre $\frac{b-a}{2^{n+1}}$ è la distanza tra il punto medio e ξ , e non può superare metà dell'intervallo $b_n - a_n$.

Quindi si dimostra che le tre successioni convergono ad uno zero, utilizzando il teorema dei carabinieri:

$$0 \leq |\xi - a_n|, |\xi - b_n| < \frac{b-a}{2^n} \rightarrow 0, n \rightarrow \infty \xrightarrow{\text{teor. carabinieri}} |\xi - a_n|, |\xi - b_n| \rightarrow 0, n \rightarrow \infty$$

Qui semplicemente diciamo che $|\xi - a_n|, |\xi - b_n|$ sono due quantità positive (per via dei valori assoluti), e da quello che abbiamo detto prima sono \leq di $\frac{b-a}{2^n}$.

Calcolando il $\lim_{n \rightarrow \infty} \frac{b-a}{2^n}$ otteniamo che è zero. Quindi possiamo utilizzare il teorema dei carabinieri per dire che le due successioni tendono a zero.

Lo stesso ragionamento lo utilizziamo per la terza successione:

$$0 \leq |\xi - x_n| < \frac{b-a}{2^{n+1}} \implies |\xi - x_n| \rightarrow 0, n \rightarrow \infty$$