Facoltá di SMFN
Corso di Studi in MAtematica- A.A. 2009-2010

Corso di ANALISI NUMERICA 1:

Esempi di esercizi svolti
Prof. Carla MANNI

1) *Determinare se il problema del calcolo delle radici reali dell'equazione*

$$x^2 - 2x + c = 0$$

risulta malcondizionato al variare di c *sulla retta reale.*

Le radici dell'equazione sono:

$$x_1(c) = 1 - \sqrt{1-c}, \quad x_2(c) = 1 + \sqrt{1-c},$$

tali radici sono reali per $c \leq 1$.
Consideriamo ad esempio $x_1(c)$. Il coefficiente di amplificazione risulta

$$\frac{c}{2\sqrt{1-c}} \frac{1}{(1-\sqrt{1-c})} = \frac{1+\sqrt{1-c}}{2\sqrt{1-c}}$$

tale espressione risulta illimitata quando $c \to 1^-$ mentre risulta limitata sia per $c \to 0$ che per $c \to -\infty$. Quindi il problema è malcondizionato per valori di c prossimi ad 1, $c \leq 1$. Analogamente si conclude per $x_2(c)$.

2) *Determinare la rappresentazione in base* 2 *di* $\frac{1}{10}$.

Vogliamo determinare $d_1, d_2, \cdots, d_n, \cdots$ tali che

$$\frac{1}{10} = \frac{d_1}{2} + \frac{d_2}{2^2} + \cdots + \frac{d_n}{2^n} + \cdots =: (0.d_1 d_2 \cdots)_2$$

Risulta

$$\begin{aligned}
\frac{1}{10}2 &= \frac{1}{5} < 1 &&\Rightarrow d_1 = 0\\
\frac{1}{5}2 &= \frac{2}{5} < 1 &&\Rightarrow d_2 = 0\\
\frac{2}{5}2 &= \frac{4}{5} < 1 &&\Rightarrow d_3 = 0\\
\frac{4}{5}2 &= \frac{8}{5} = 1 + \frac{3}{5} &&\Rightarrow d_4 = 1\\
\frac{3}{5}2 &= \frac{6}{5} = 1 + \frac{1}{5} &&\Rightarrow d_5 = 1\\
&\vdots
\end{aligned}$$

quindi

$$(0.1)_{10} = \frac{1}{10} = (0.0001100110011....)_2 = 2^{-3}(0.\overline{1100})_2 \ .$$

Ne segue che $\frac{1}{10}$ ha una rappresentazione periodica (per cui non ammette una rappresentazione con un numero finito di cifre) in base 2.

3) *Dato l'insieme di numeri di macchina* $F(2,3,-3,3)$ *ed i numeri reali* $a = \frac{1}{7}$, $b = \frac{1}{28}$, *determinare:*

i) la rappresentazione di a *e* b *in base* 2*;*

ii) le approssimazioni $\tilde{a}$, $\tilde{b}$ *di* a *e* b *in* $F(2,3,-3,3)$ *operando con troncamento;*

iii) i corrispondenti errori di rappresentazione e confrontarli con la precisione di macchina.

i) Risulta

$$\begin{aligned} &\frac{1}{7}2 = \frac{2}{7} < 1 && \Rightarrow d_1 = 0 \\ &\frac{2}{7}2 = \frac{4}{7} < 1 && \Rightarrow d_2 = 0 \\ &\frac{4}{7}2 = \frac{8}{7} = 1 + \frac{1}{7} && \Rightarrow d_3 = 1 \\ &\vdots \end{aligned}$$

quindi

$$\frac{1}{7} = (0.001001....)_2 = 2^{-2}(0.\overline{100})_2 \ .$$

Osservando che $\frac{1}{28} = \frac{1}{7}2^{-2}$ si ha inoltre

$$\frac{1}{28} = (0.001001....)_2 2^{-2} = 2^{-4}(0.\overline{100})_2 \ .$$

ii) $\tilde{a} = 2^{-2}(0.100)_2 = \frac{1}{8}$, mentre $\tilde{b} = 0$ perché si ha una condizione di underflow.

iii) La precisione di macchina risulta $\epsilon_M = 2^{-2}$. Inoltre

$$\frac{|\tilde{a} - a|}{|a|} = \frac{1}{8} < \epsilon_M, \ \frac{|\tilde{b} - b|}{|b|} = 1.$$

4) *Dato l'insieme di numeri di macchina* $F(\beta, t, L, U)$ *ove si operi con troncamento, mostrare che la precisione di macchina é il piú piccolo numero positivo,* x, *appartenente all' insieme per cui risulta*

$$1 \oplus x > 1.$$

Il valore della precisione di macchina, ϵ_M, é $\epsilon_M = \beta^{1-t}$. Quindi

$$1 \oplus \epsilon_M = \mathrm{tr}(\beta^{-1} + \beta^{-t})\beta = \beta(\beta^{-1} + \beta^{-t}) > 1$$

mentre per ogni $x \in F(\beta, t, L, U),\ 0 < x < \epsilon_M$ si ha

$$x = \beta(\frac{0}{\beta} + \cdots + \frac{0}{\beta^t} + \frac{d_1}{\beta^{t+1+k}} + \cdots + \frac{d_t}{\beta^{2t+k}}),\ \text{per qualche intero}\ \ k \geq 0$$

per cui

$$1 \oplus x = \mathrm{tr}(1 + x) = 1.$$

5) *Siano $x > y > z > 0$ numeri di macchina. Stabilire quale delle seguenti espressioni è preferibile dal punto di vista della propagazione degli errori:*

$$(x \oplus y) \oplus z, \quad x \oplus (y \oplus z).$$

Risulta, in assenza di overflow e underflow,

$$x \oplus y = (x + y)(1 + \epsilon_1),\ |\epsilon_1| < \epsilon_M$$

e

$$(x \oplus y) \oplus z = ((x + y)(1 + \epsilon_1) + z)(1 + \epsilon_2),\ |\epsilon_2| < \epsilon_M.$$

Quindi, tralasciando i termini di grado superiore al primo,

$$\left| \frac{(x \oplus y) \oplus z - (x + y + z)}{x + y + z} \right| < (1 + \frac{x + y}{x + y + z})\epsilon_M.$$

Analogamente

$$\left| \frac{x \oplus (y \oplus z) - (x + y + z)}{x + y + z} \right| < (1 + \frac{y + z}{x + y + z})\epsilon_M.$$

Ne segue che, essendo $x > y > z$, la seconda espressione è preferibile.

6) *Dato l'insieme di numeri di macchina $F(2, 5, m, M)$ in cui si opera con troncamento ed i numeri reali $a = \frac{1}{16}$, $b = \frac{1}{17}$, determinare:*

i) la rappresentazione di a e b in base 2;

ii) le approssimazioni $\tilde{a}$, $\tilde{b}$ di a e b in $F(2, 5, m, M)$;

iii) i corrispondenti errori di rappresentazione e confrontarli con la precisione di macchina;

iv) il valore $\tilde{a} \ominus \tilde{b}$ e valutare l'errore relativo rispetto ad $a - b$.

i) Risulta

$$a = \frac{1}{16} = 2^{-3}(0.1)_2, \quad \frac{1}{17} = (0.0000111100001111....)_2 = 2^{-4}(0.\overline{11110000})_2 \ .$$

ii) $\tilde{a} = a$ mentre $\tilde{b} = 2^{-4}(0.11110)_2 = \frac{1}{16}\frac{15}{16}$.

iii) La precisione di macchina risulta $\epsilon_M = 2^{-4}$. Inoltre

$$\frac{|\tilde{a} - a|}{|a|} = 0 < \epsilon_M, \quad \frac{|\tilde{b} - b|}{|b|} = 2^{-8} < \epsilon_M.$$

iv) $\tilde{a} \ominus \tilde{b} = tr((0.1 - 0.01111)2^{-3}) = 2^{-8}$ mentre $a - b = \frac{1}{272}$ per cui l'errore relativo sulla differenza risulta

$$(\frac{1}{256} - \frac{1}{272})272 = 2^{-4}$$

assai maggiore degli errori relativi sui dati.

7) *Analizzare, al variare di k sulla retta reale, la convergenza delle successioni costruite mediante il metodo iterativo*

$$x_{i+1} = g(x_i), \quad g(x) = k - \mathrm{e}^{-x}.$$

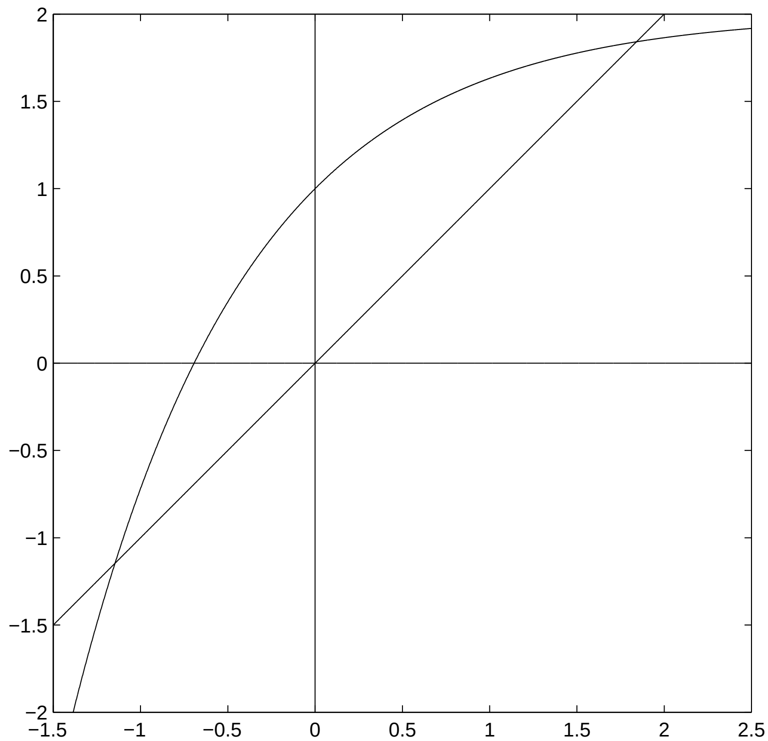

Figura 1

La funzione g ammette punti fissi, ossia punti tali che $x = g(x)$, solo se $k \geq 1$.
Per $k = 1$ esite un solo punto fisso, α e risulta $\alpha = 0$. Da un esame grafico si ha che la successione $x_{i+1} = g(x_i)$ converge ad α in modo monotono decrescente per ogni $x_0 \geq 0$ mentre risulta divergente a $-\infty$ se $x_0 < 0$.
Per $k > 1$ esitono due punti fissi, α e β con $\alpha > 0$, $\beta < 0$ (si veda la figura 1 per il caso $k = 2$.) Da un esame grafico si ha che la successione $x_{i+1} = g(x_i)$ converge ad α in maniera monotona per ogni $x_0 > \beta$ mentre risulta divergente a $-\infty$ se $x_0 < \beta$.
Si noti che, per ogni valore del parametro k, risulta

$$|g'(x)| = \mathrm{e}^{-x} < 1, \ \forall \ x > 0.$$

8) *Analizzare, la convergenza delle successioni costruite mediante il metodo di Newton per approssimare gli zeri della funzione*

$$f(x) = \frac{2}{3\sqrt{3}}x^3 + x^2 - 1.$$

La funzione ha una radice doppia in $\alpha = -\sqrt{3}$ ed una radice semplice $\beta \in (0, 1)$ (vedere figura 2).
Dal grafico risulta che il metodo di Newton origina una succesione convergente ad α in modo monotono (al più dopo la prima iterazione) per ogni $x_0 < 0$. Analogamente si ha una succesione monotona decrescente (al piú dopo la prima iterazione) convergente a β per ogni $x_0 > 0$.
L'ordine di convergenza é 1 per α in quanto si tratta di una radice doppia e 2 per β.

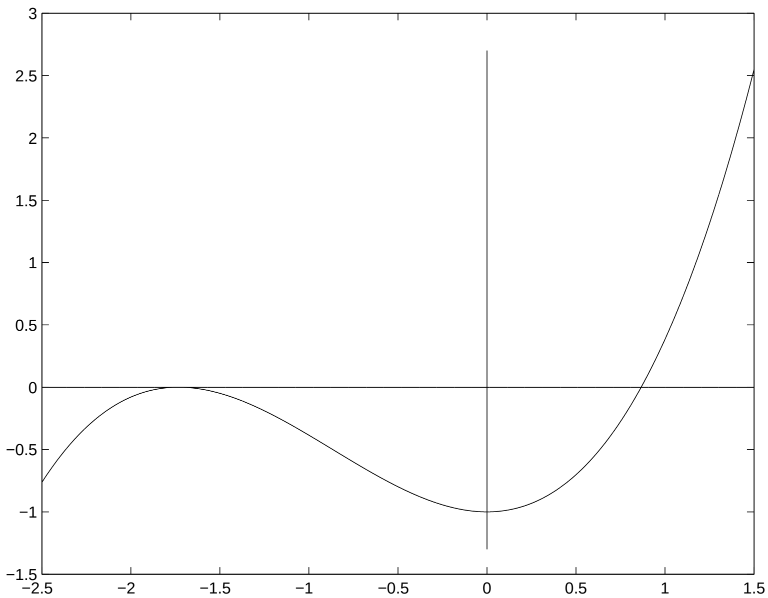

Figura 2

9) *Analizzare la convergenza del metodo iterativo*

$$x_{i+1} = g(x_i), \; g(x) = a\cos(x), \quad x \in (0, \frac{\pi}{2})$$

nei casi $a \in (0, \frac{\pi}{3})$ *ed* $a \in (\frac{2\pi}{3}, \pi)$.

Per ogni valore positivo del parametro a la funzione g ammette un unico punto fisso $\alpha \in (0, \frac{\pi}{2})$ (si veda in proposito la Figura 3 a sinistra, per il caso $a = 1$) e risulta

$$a = \frac{\alpha}{\cos(\alpha)}.$$

Se

$$|g'(\alpha)| = |-a \ \mathrm{sen}(\alpha)| = \alpha \mathrm{tg}(\alpha) < 1$$

il metodo iterativo produce successioni convergenti ad α purché il punto iniziale x_0 sia sufficientemente vicino al punto fisso. Analizziamo quindi per quali valori di $\alpha \in (0, \frac{\pi}{2})$ risulta

$$\mathrm{tg}(\alpha) < \frac{1}{\alpha}.$$

Da uno studio grafico (si veda Figura 3 a destra), risulta che tale disuguaglianza vale per $\alpha \leq \frac{\pi}{4}$ mentre non vale per $\alpha \geq \frac{\pi}{3}$. La funzione $a(\alpha) = \frac{\alpha}{\cos(\alpha)}$ é una funzione monotona crescente per $\alpha \in (0, \frac{\pi}{2})$ e risulta

$$\frac{\pi}{4}\frac{1}{\cos(\frac{\pi}{4})} = \frac{\pi}{4}\sqrt{2} > \frac{\pi}{3} \quad \text{mentre} \quad \frac{\pi}{3}\frac{1}{\cos(\frac{\pi}{3})} = \frac{2\pi}{3}.$$

Quindi il metodo iterativo converge (con ordine di convergenza 1 in quanto $g'(\alpha) \neq 0$) per $a \in (0, \frac{\pi}{3})$, mentre non converge nell'altro caso.

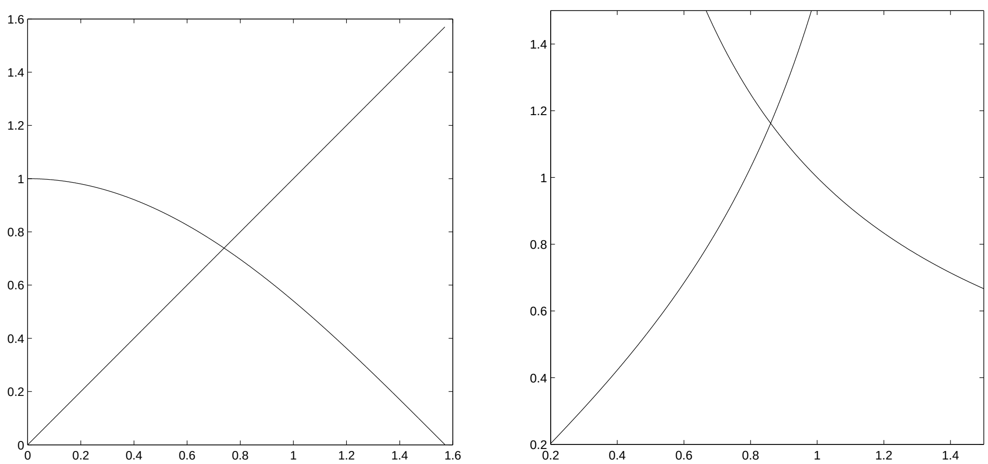

Figura 3

10) *Analizzare, al variare di k sulla retta reale, la convergenza delle successioni costruite mediante il metodo iterativo*

$$x_{i+1} = g(x_i), \quad g(x) = 2k\sqrt{x} - x, \quad x > 0.$$

La funzione g ha punti fissi positivi unicamente se $k > 0$. In tale ipotesi l'unico punto fisso positivo é $\alpha = k^2$. Inoltre $g(x) = 0$ per $x = 0,\ 4k^2$ e $g'(x) = 0$ per $x = \alpha = k^2$ (si veda la figura 4 per $k = 1$). La successione generata dal metodo iterativo converge in modo monotono crescente (al piú dopo la prima iterazione) ad α per ogni $0 < x_0 < 4k^2$. L'ordine di convergenza é 2 in quanto $g'(\alpha) = 0,\ g''(\alpha) \neq 0$.

11) *Un valore approssimato di* $\ln 2$ *si può ottenere approssimando gli zeri della funzione*

$$f(x) = \mathrm{e}^x - 2;$$

i) determinare per quali $x_0 \in \mathbb{R}$ il metodo di Newton per l'approssimazione degli zeri di f origina una successione convergente a $\ln 2$*;*

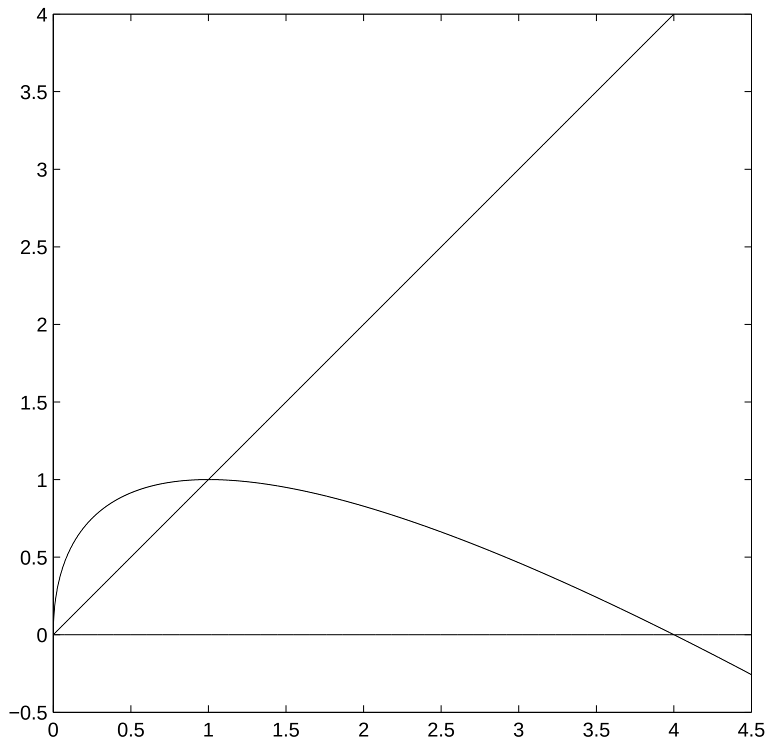

Figura 4

ii) determinare x_0 *ed* ϵ *in modo che la successione* $\{x_i\}$ *costruita tramite il metodo di Newton a partire da* x_0 *sia decrescente e tale che*

$$|x_{i+1} - x_i| < \epsilon \;\Rightarrow\; |x_i - \ln 2| < 10^{-5}$$

i) La funzione f è monotona crescente e convessa in $\mathbb{R}$, quindi esiste un solo zero di f che é appunto dato da $\ln 2$. Per ogni $x_0 \in \mathbb{R}$ il metodo di Newton origina una successione convergente a $\ln 2$. La convergenza é monotona decrescente al piú dopo la prima iterazione.

ii) Scegliendo $x_0 = 1$ si ha una successione monotona decrescente convergente a $\ln 2$. Inoltre, posto

$$g(x) = x - \frac{f(x)}{f'(x)} = x - \frac{e^x - 2}{e^x}$$

risulta

$$x_{i+1} - x_i = (g'(\xi_i) - 1)(x_i - \ln 2) = -\frac{2}{e^{\xi_i}}(x_i - \ln 2), \quad \ln 2 < \xi_i < x_i < 1.$$

Quindi

$$|x_i - \ln 2| = \frac{e^{\xi_i}}{2}|x_{i+1} - x_i| < \frac{e}{2}|x_{i+1} - x_i|.$$

Ne segue che, per aver assicurata la disuguaglianza richiesta, sarà sufficiente porre

$$\epsilon \leq \frac{2}{e}10^{-5}.$$