

Facoltà di Scienze MM.FF.NN.
Corso di Studi in Matematica- A.A. 2010-2011

Corso di ANALISI NUMERICA 1:

Esempi di esercizi svolti

- 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:

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

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

- 2) *Dato un triangolo, siano a, b le lunghezze di due lati e sia θ l'ampiezza (in radianti) dell'angolo fra essi compreso. Analizzare il condizionamento del calcolo dell'area del triangolo al variare di θ .*

L'area del triangolo risulta $A = \frac{1}{2}ab \sin(\theta)$. Quindi il coefficiente di amplificazione (rispetto a θ) è:

$$\cos(\theta) \frac{\theta}{\sin(\theta)}$$

per cui il problema è malcondizionato per θ prossimo a π .

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

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

$$\frac{1}{10} = \frac{d_1}{2} + \frac{d_2}{2^2} + \dots + \frac{d_n}{2^n} + \dots =: (0.d_1d_2\dots)_2$$

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

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

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.

4) 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

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

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

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

- 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

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

5) Dato l'insieme di numeri di macchina $F(\beta, t, L, U)$ ove si operi con troncamento, mostrare che la precisione di macchina è il più 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

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

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

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

per cui

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

- 6) *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,

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

e

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

Quindi, tralasciando i termini di grado superiore al primo,

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

Analogamente

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

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

- 7) *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

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

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

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

assai maggiore degli errori relativi sui dati.

- 8) *Analizzare, al variare di x_0 sulla retta reale, la convergenza delle successioni costruite mediante il metodo di Newton per le seguenti funzioni*

$$f_1(x) = |x|^{\frac{1}{2}}, \quad f_2(x) = x^{\frac{2}{3}}, \quad f_3(x) = |x|^{\frac{1}{3}}.$$

Per ciascuna delle funzioni assegnate esiste una sola radice $\alpha = 0$. Per la funzione f_1 risulta

$$f_1'(x) = \begin{cases} \frac{1}{2\sqrt{x}} & \text{if } x > 0, \\ \frac{-1}{2\sqrt{-x}} & \text{if } x < 0, \end{cases}$$

quindi la successione generata dal metodo di Newton risulta essere un ciclo, per ogni punto iniziale x_0 , infatti

$$x_{i+1} = x_i - \frac{f_1(x_i)}{f_1'(x_i)} = x_i - 2x_i = -x_i, \quad \forall x_i \neq 0.$$

Analogamente, per $x_i \neq 0$

$$x_{i+1} = x_i - \frac{f_2(x_i)}{f_2'(x_i)} = x_i - \frac{3}{2}x_i = -\frac{1}{2}x_i;$$

$$x_{i+1} = x_i - \frac{f_3(x_i)}{f_3'(x_i)} = x_i - 3x_i = -2x_i.$$

Per cui nel secondo caso si ha convergenza a 0 per ogni punto iniziale, mentre nel terzo la successione è divergente ($|x_i| \rightarrow +\infty$).

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

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

La funzione g ammette punti fissi, ossia punti tali che $x = g(x)$, solo se $k \geq 1$.

Per $k = 1$ esiste 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 $k > 1$ esistono 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

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

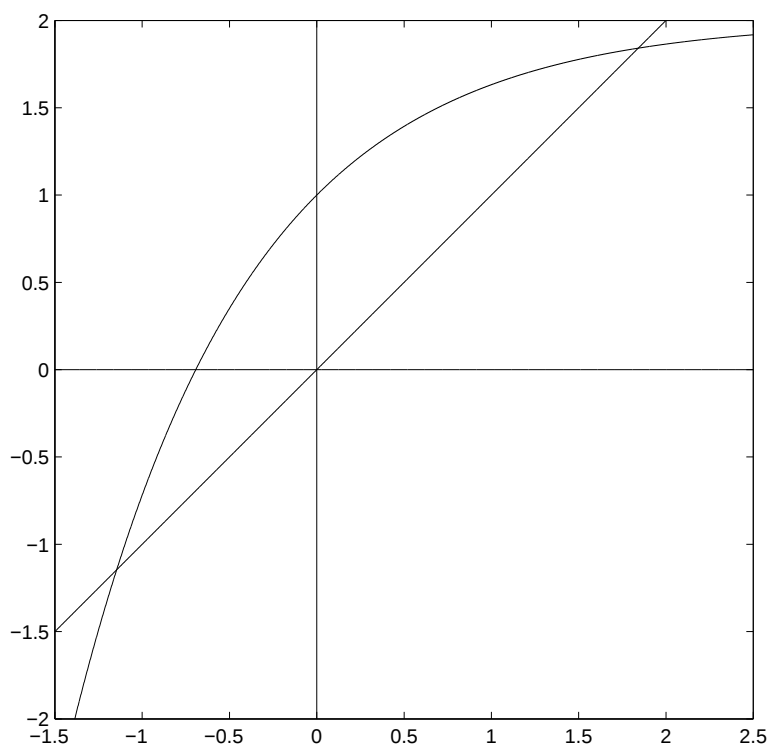


Figura 1

- 10) *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 successione convergente ad α in modo monotono (al più dopo la prima iterazione) per ogni $x_0 < 0$. Analogamente si ha una successione monotona decrescente (al più dopo la prima iterazione) convergente a β per ogni $x_0 > 0$.

L'ordine di convergenza è 1 per α in quanto si tratta di una radice doppia e 2 per β .

- 11) *Analizzare la convergenza del metodo iterativo*

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

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

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

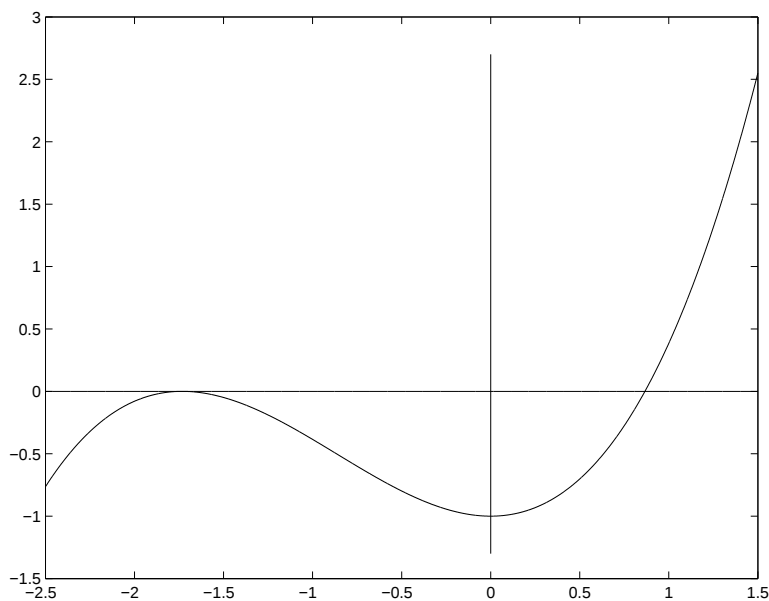


Figura 2

Se

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

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

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

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.

- 12) *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

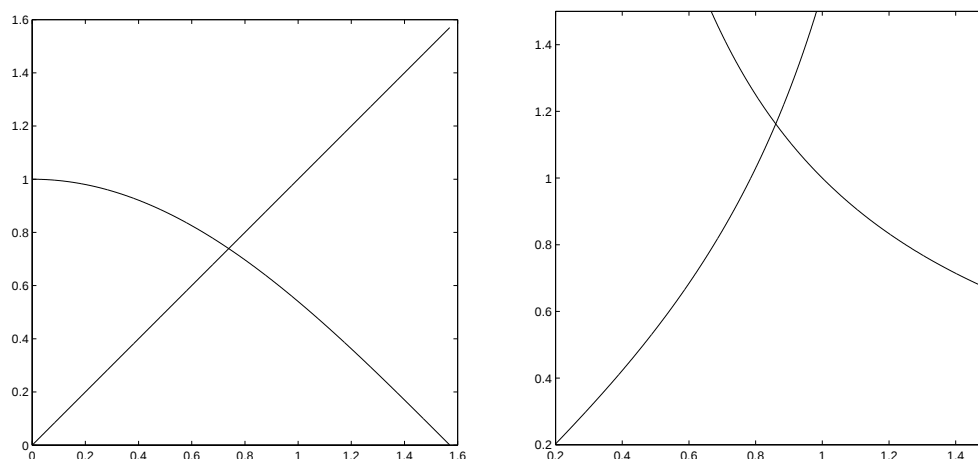


Figura 3

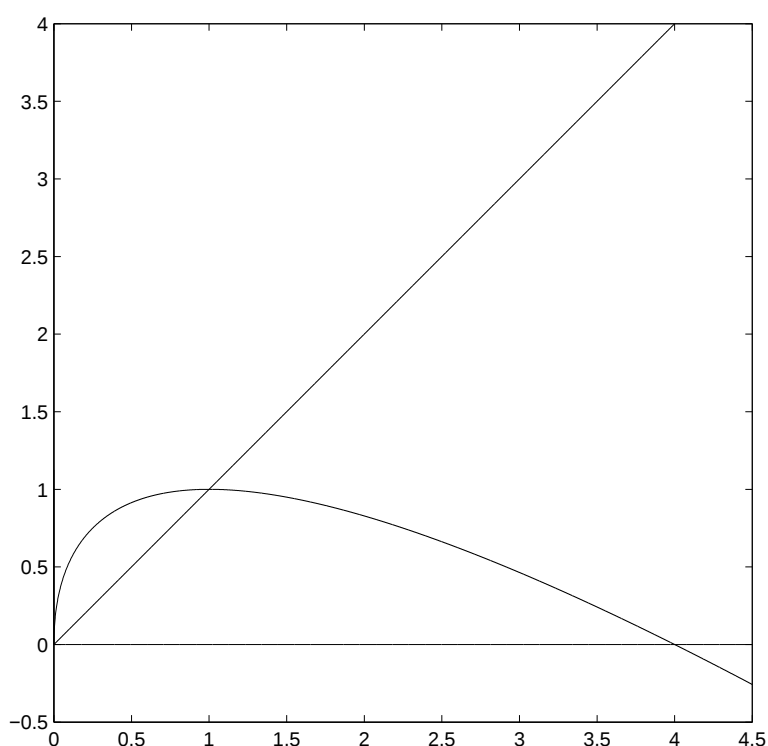


Figura 4

iterativo converge in modo monotono crescente (al più dopo la prima iterazione) ad α per ogni $0 < x_0 < 4k^2$. L'ordine di convergenza è 2 in quanto $g'(\alpha) = 0$, $g''(\alpha) \neq 0$.

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

$$f(x) = e^x - 2;$$

- 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 $\ln 2$;
- 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 è 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 è monotona decrescente al più 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

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

14) *Data la matrice*

$$\begin{pmatrix} 17 & 2 & 2 \\ 1 & 10 & 3 \\ 1 & 2 & 0 \end{pmatrix}$$

- i) stabilire se ammette autovalori complessi;*
ii) localizzarne il raggio spettrale.

- i) Il polinomio caratteristico della matrice è un polinomio di terzo grado a coefficienti reali; quindi la matrice ha almeno un autovalore reale e gli altri due autovalori, in caso siano complessi, devono essere complessi coniugati. I cerchi di Gershgorin per colonne sono disgiunti ed hanno centro sull'asse reale. Ne segue che, dal secondo teorema di Gershgorin, in ogni cerchio si trova esattamente un autovalore. La matrice non può quindi avere autovalori complessi perchè due autovalori complessi coniugati, avendo la stessa parte reale, dovrebbero appartenere ad uno stesso cerchio.
- ii) Dati i cerchi di Gershgorin per colonne, risulta che la matrice ha un solo autovalore di modulo massimo e che questo si trova nel cerchio di centro 17 e raggio 2. Quindi per il raggio spettrale della matrice valgono le seguenti limitazioni

$$15 < \rho(A) < 19.$$

Si noti che valgono le disuguaglianze strette perché, essendo la matrice irriducibile, dal terzo teorema di Gershgorin l'autovalore in questione non può appartenere alla frontiera del cerchio suddetto.

15) *Dati i sistemi lineari*

$$Ax = b, \quad A = \begin{pmatrix} 1 & -1 + 10^{-5} \\ 1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad (1)$$

$$(A + \delta A)(x + \delta x) = b, \quad (2)$$

$$\delta A = \begin{pmatrix} 0 & -10^{-5} + \epsilon \\ 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \epsilon \in \mathbb{R}, \quad 0 < \epsilon < 10^{-5};$$

- i) determinare $\mu_\infty(A)$ e valutare $\|b\|_\infty$, $\|\delta A\|_\infty$.*
ii) Calcolare la soluzione x del sistema (1) e la soluzione $x + \delta x$ del sistema (2). Valutare $\frac{\|\delta x\|_\infty}{\|x\|_\infty}$ e commentare il risultato.

i) Risulta

$$A^{-1} = 10^5 \begin{pmatrix} 1 & -1 + 10^{-5} \\ 1 & -1 \end{pmatrix},$$

quindi $\mu_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty = 2 \cdot 2 \cdot 10^5 = 4 \cdot 10^5$, mentre $\|b\|_\infty = 1$, $\|\delta A\|_\infty = 10^{-5} - \epsilon$.

ii) Si ha

$$x = \begin{pmatrix} 10^5 \\ 10^5 \end{pmatrix}, \quad x + \delta x = \begin{pmatrix} \epsilon^{-1} \\ \epsilon^{-1} \end{pmatrix}$$

quindi

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty} = \epsilon^{-1} \cdot 10^{-5} - 1. \quad (3)$$

Quindi l'errore relativo sulla soluzione può essere arbitrariamente elevato quando ϵ si avvicina a 0.

Si noti che la matrice A è malcondizionata ma l'esplosione dell'errore relativo ottenuta in (3) non può essere completamente descritta tramite la maggiorazione

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty} \leq \frac{\mu_\infty(A)}{(1 - \|A^{-1}\|_\infty \|\delta A\|_\infty)} \frac{\|\delta A\|_\infty}{\|A\|_\infty}$$

per ogni $\epsilon \in (0, 10^{-5})$ ma solo per i valori di ϵ per cui risulta

$$\|A^{-1}\|_\infty \|\delta A\|_\infty < 1$$

ossia per

$$\frac{1}{2} \cdot 10^{-5} < \epsilon < 10^{-5}.$$

- 16) *Mostrare che una matrice a diagonale dominante in senso stretto ammette la fattorizzazione LU.*

Si ricorda che, per il primo teorema di Gerschgorin, una matrice a diagonale dominante in senso stretto non può avere autovalori nulli e quindi risulta non singolare. Inoltre per una matrice a diagonale dominante in senso stretto si ha

$$|a_{i,i}| > \sum_{j=1, j \neq i}^n |a_{i,j}| \geq \sum_{j=1, j \neq i}^k |a_{i,j}|, \quad i \leq k \leq n.$$

Ne segue che, in una matrice a diagonale dominante in senso stretto, tutti i minori principali di testa sono a diagonale dominante in senso stretto ossia sono non singolari. Da cui si ha che la matrice ammette la fattorizzazione LU.

- 17) *Calcolare la fattorizzazione QR della matrice*

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ 2 & -4 & 5 \end{pmatrix},$$

Risulta

$$x = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad \|x\|_2 = 3, \quad P_1 = I - \frac{2}{v_1^T v_1} v_1 v_1^T, \quad v_1 = x + \|x\|_2 e_1 = \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix},$$

quindi

$$P_1 = I - \frac{1}{12} \begin{pmatrix} 16 & 8 & 8 \\ 8 & 4 & 4 \\ 8 & 4 & 4 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -1 & -2 & -2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix}.$$

Ne segue che

$$A^{(2)} = P_1 A^{(1)} = P_1 A = \frac{1}{3} \begin{pmatrix} -1 & -2 & -2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ 2 & -4 & 5 \end{pmatrix} = \begin{pmatrix} -3 & 3 & -3 \\ 0 & 0 & -3 \\ 0 & -3 & 3 \end{pmatrix}.$$

Quindi

$$x = \begin{pmatrix} 0 \\ -3 \end{pmatrix}, \quad \|x\|_2 = 3, \quad \tilde{P}_2 = I - \frac{2}{v_2^T v_2} v_2 v_2^T, \quad v_2 = x + \|x\|_2 e_1 = \begin{pmatrix} 3 \\ -3 \end{pmatrix}$$

ossia

$$\tilde{P}_2 = I - \frac{1}{9} \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ e } P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Quindi

$$R = A^{(3)} = P_2 A^{(2)} = \begin{pmatrix} -3 & 3 & -3 \\ 0 & -3 & 3 \\ 0 & 0 & -3 \end{pmatrix}, \quad Q = P_1 P_2 = \frac{1}{3} \begin{pmatrix} -1 & -2 & -2 \\ -2 & -1 & 2 \\ -2 & 2 & -1 \end{pmatrix}.$$

18) *Data la matrice*

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

analizzare la convergenza del metodo di Jacobi e del metodo di Gauss-Seidel per il sistema lineare $Ax = b$, $b \in \mathbb{R}^3$.

Indicata con J la matrice di iterazione del metodo di Jacobi risulta

$$J = - \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

i cui autovalori sono

$$\lambda_1 = \frac{1}{2}, \quad \lambda_2 = \frac{1}{2}, \quad \lambda_3 = -1$$

quindi $\rho(J) = 1$ ed il metodo non risulta convergente.

Indicata con GS la matrice di iterazione del metodo di Gauss-Seidel si ha

$$GS = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -1 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & \frac{1}{8} & \frac{3}{8} \end{pmatrix}$$

i cui autovalori sono

$$\lambda_1 = 0, \quad \lambda_2 = \frac{1}{16}(5 + i\sqrt{7}), \quad \lambda_3 = \frac{1}{16}(5 - i\sqrt{7})$$

quindi

$$\rho(GS) = \frac{1}{16}\sqrt{25+7} = \frac{\sqrt{2}}{4} < 1$$

ed il metodo risulta convergente. Tale risultato, riguardo alla convergenza del metodo di Gauss-Seidel, poteva essere stabilito anche osservando che la matrice A ha elementi diagonali positivi e risulta definita positiva in quanto i determinanti dei suoi minori principali di testa sono positivi.

- 19) *Analizzare la convergenza del metodo di Jacobi per il sistema lineare $Ax = b$, $b \in \mathbb{R}^n$ con*

$$A = \begin{pmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \cdots & & 0 & 1 & -2 & 1 \\ 0 & \cdots & & & 0 & 1 & -2 \end{pmatrix}.$$

La matrice A é irriducibile ed a diagonale dominante quindi il metodo di Jacobi per il sistema assegnato converge.

- 20) *Data la matrice*

$$A = \begin{pmatrix} 1 & \alpha & -\alpha \\ 1 & 1 & 1 \\ \alpha & \alpha & 1 \end{pmatrix}, \quad \alpha > 0$$

- i) analizzare la convergenza del metodo di Jacobi e del metodo di Gauss-Seidel per il sistema lineare $Ax = b$, $b \in \mathbb{R}^3$.*
ii) per i valori di α per cui entrambi i metodi convergono, si dica quale dei due, asintoticamente, converge piú velocemente.

- i) Indicata con J la matrice di iterazione del metodo di Jacobi risulta

$$J = \begin{pmatrix} 0 & -\alpha & \alpha \\ -1 & 0 & -1 \\ -\alpha & -\alpha & 0 \end{pmatrix}$$

il cui polinomio caratteristico é

$$-\lambda^3 + (-\alpha^2 + 2\alpha)\lambda$$

per cui il raggio spettrale risulta

$$\rho(J) = \sqrt{|\alpha^2 - 2\alpha|};$$

quindi il metodo di Jacobi é convergente quando

$$|\alpha^2 - 2\alpha| < 1$$

ossia per $\alpha < 1 + \sqrt{2}$, $\alpha \neq 1$ (si ricordi che $\alpha > 0$).

Analogamente, indicata con GS la matrice di iterazione del metodo di Gauss-Seidel, si ha

$$GS = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \alpha & \alpha & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -\alpha & \alpha \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\alpha & \alpha \\ 0 & \alpha & -\alpha - 1 \\ 0 & 0 & \alpha \end{pmatrix}$$

il cui raggio spettrale risulta (si ricordi che $\alpha > 0$)

$$\rho(GS) = \alpha;$$

quindi il metodo di Gauss-Seidel converge per $\alpha < 1$.

ii) Poiché per $0 < \alpha < 1$ risulta

$$\sqrt{|\alpha^2 - 2\alpha|} > \alpha$$

il metodo di Gauss-Seidel ha, asintoticamente, velocità di convergenza maggiore.

21) *Dati i punti*

$$x_0 = 0, \quad x_1 = \frac{1}{2}, \quad x_2 = 1$$

e la funzione

$$f(x) = e^x;$$

- i) determinare il polinomio, nella forma di Lagrange, che interpola la funzione data sui punti assegnati;*
- ii) determinare il polinomio, nella forma di Newton, che interpola la funzione data sui punti assegnati;*
- iii) dare una maggiorazione dell'errore che si commette nel sostituire, nell'intervallo $[0, 1]$, alla funzione data il polinomio interpolante.*

i) Il polinomio interpolante nella forma di Lagrange è

$$p_n(x) = \sum_{i=0}^n f(x_i) L_{i,n}(x)$$

ove

$$L_{i,n}(x) = \prod_{j=0, j \neq i}^n \frac{(x - x_j)}{(x_i - x_j)}.$$

Nel caso specifico è $n = 2$ e risulta

$$\begin{aligned} L_{0,2}(x) &= \frac{(x - \frac{1}{2})(x - 1)}{(-\frac{1}{2})(-1)} = 2(x - \frac{1}{2})(x - 1), \\ L_{1,2}(x) &= \frac{x(x - 1)}{\frac{1}{2}(-\frac{1}{2})} = -4x(x - 1), \\ L_{2,2}(x) &= \frac{x(x - \frac{1}{2})}{(1)\frac{1}{2}} = 2x(x - \frac{1}{2}), \end{aligned}$$

quindi

$$\begin{aligned} p_2(x) &= 2(x - \frac{1}{2})(x - 1) - 4x(x - 1)\sqrt{e} + 2x(x - \frac{1}{2})e \\ &= 1 + x(-3 + 4\sqrt{e} - e) + 2x^2(-2\sqrt{e} + e + 1). \end{aligned}$$

ii) Il polinomio interpolante nella forma di Newton è

$$p_n(x) = f[x_0] + f[x_0, x_1](x - x_0) + \cdots + f[x_0, x_1, \cdots, x_n](x - x_0) \cdots (x - x_{n-1})$$

ove $f[x_0, x_1, \cdots, x_k]$ indica la differenza divisa di ordine k di f relativa ai punti $x_0, x_1, \cdots, x_k$. Nel caso specifico risulta

$$\begin{aligned} f[x_0] &= f(x_0) = 1, \\ f[x_0, x_1] &= \frac{f(x_1) - f(x_0)}{x_1 - x_0} = 2(\sqrt{e} - 1), \\ f[x_0, x_1, x_2] &= \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = 2(\sqrt{e} - 1)^2 \end{aligned}$$

quindi

$$p_2(x) = 1 + 2(\sqrt{e}-1)x + 2(\sqrt{e}-1)^2x(x-\frac{1}{2}) = 1 + x(-3+4\sqrt{e}-e) + 2x^2(-2\sqrt{e}+e+1).$$

iii) Risulta in generale, se la funzione f é di classe C^{n+1} ,

$$f(x) - p_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x-x_0) \cdots (x-x_n)$$

nel caso specifico abbiamo quindi

$$\max_{x \in [0,1]} |e^x - p_2(x)| = \frac{1}{6} \max_{x \in [0,1]} e^\xi |x(x-\frac{1}{2})(x-1)| \leq \frac{1}{6} e \frac{\sqrt{3}}{36}.$$

21) *Determinare la retta, $r(x)$, e la parabola, $q(x)$, che meglio approssimano nel senso dei minimi quadrati i punti*

$$x_i : \quad -1 \quad -\frac{1}{2} \quad 0 \quad \frac{1}{2} \quad 1$$

$$y_i : \quad -1 \quad -\frac{1}{8} \quad 0 \quad \frac{1}{8} \quad 1.$$

La retta in questione, $r(x) = a_0 + a_1x$, si determina risolvendo il sistema

$$\begin{pmatrix} \sum_{i=0}^N 1 & \sum_{i=0}^N x_i \\ \sum_{i=0}^N x_i & \sum_{i=0}^N x_i^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^N y_i \\ \sum_{i=0}^N y_i x_i \end{pmatrix}$$

nel caso specifico $N = 4$ ed il sistema suddetto diviene

$$\begin{pmatrix} 5 & 0 \\ 0 & \frac{5}{2} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{17}{8} \end{pmatrix}$$

da cui $a_0 = 0$, $a_1 = \frac{17}{20}$, ossia

$$r(x) = \frac{17}{20}x.$$

La parabola, $q(x) = a_0 + a_1x + a_2x^2$, si determina risolvendo il sistema

$$\begin{pmatrix} \sum_{i=0}^N 1 & \sum_{i=0}^N x_i & \sum_{i=0}^N x_i^2 \\ \sum_{i=0}^N x_i & \sum_{i=0}^N x_i^2 & \sum_{i=0}^N x_i^3 \\ \sum_{i=0}^N x_i^2 & \sum_{i=0}^N x_i^3 & \sum_{i=0}^N x_i^4 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^N y_i \\ \sum_{i=0}^N y_i x_i \\ \sum_{i=0}^N y_i x_i^2 \end{pmatrix}$$

nel caso specifico $N = 4$ ed il sistema suddetto diviene

$$\begin{pmatrix} 5 & 0 & \frac{5}{2} \\ 0 & \frac{5}{2} & 0 \\ \frac{5}{2} & 0 & \frac{17}{8} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{17}{8} \\ 0 \end{pmatrix}$$

da cui $a_0 = 0$, $a_1 = \frac{17}{20}$, $a_2 = 0$ ossia

$$q(x) = r(x) = \frac{17}{20}x$$

come si poteva dedurre dal grafico.

22) *Assegnata la funzione $f(x) = 1 - x^2$ ed i nodi*

$$x_0 = -h, \quad x_1 = h, \quad 0 < h \leq 1;$$

- i) determinare il polinomio, $p(x)$, che interpola la funzione $f(x)$ sui nodi assegnati*
- ii) determinare, se possibile, il valore del parametro h per cui risulta minima la quantità*

$$\max_{x \in [-1, 1]} |f(x) - p(x)|.$$

- i) da un semplice esame grafico $p(x) = 1 - h^2$.
- ii) per quanto visto al punto precedente

$$\max_{x \in [-1, 1]} |f(x) - p(x)| = \max_{x \in [-1, 1]} |h^2 - x^2| = \max\{h^2, 1 - h^2\}$$

quindi la quantità risulta minima quando

$$h^2 = 1 - h^2, \quad \text{ossia } h = \frac{1}{\sqrt{2}}.$$