Soluzione della prima prova parziale di Calcolo Numerico
5 Novembre 2013

Compito A

Esercizio 1

(a) È

$$x_1 = 100_{10} = 1100100_2 = (0.11001)_2\, 2^7, \quad x_2 = 102_{10} = 1100110_2 = (0.110011)_2\, 2^7.$$

Per $t = 2$ e 3 è $\widetilde{x}_1 = \widetilde{x}_2 = (0.11)_2\, 2^7$, quindi $\widetilde{x}_1 \oslash \widetilde{x}_2 = 1$.
Per $t = 4$ è $\widetilde{x}_1 = \widetilde{x}_2 = (0.1101)_2\, 2^7$, quindi $\widetilde{x}_1 \oslash \widetilde{x}_2 = 1$.
Per $t = 5$ è $\widetilde{x}_1 = (0.11001)_2\, 2^7$ e $\widetilde{x}_2 = (0.11010)_2\, 2^7$, quindi

$$z = \frac{\widetilde{x}_1}{\widetilde{x}_2} = \frac{(0.11001)_2\, 2^7}{(0.11010)_2\, 2^7} = \frac{(0.11001)_2}{(0.11010)_2} = \frac{25_{10}}{26_{10}}.$$

Si converte 25/26 in base 2.

$$\frac{25}{26}\times 2 = 1+\frac{24}{26}, \quad \frac{24}{26}\times 2 = 1+\frac{22}{26}, \quad \frac{22}{26}\times 2 = 1+\frac{18}{26}, \quad \frac{18}{26}\times 2 = 1+\frac{10}{26},$$
$$\frac{10}{26}\times 2 = 0+\frac{20}{26}, \quad \frac{20}{26}\times 2 = 1+\frac{14}{26}, \quad \frac{14}{26}\times 2 = 1+\frac{2}{26}, \quad \ldots$$

quindi $z = (0.1111011\ldots)_2$ e $\widetilde{z} = (0.11111)_2 \neq 1$. Il minimo t per cui $\widetilde{z} \neq 1$ è 5.

(b) Per $t = 5$ si è visto che 100 e 102 hanno due rappresentazioni diverse per cui si ha $\widetilde{z} \neq 1$. Per vedere se 102 è il minimo intero maggiore di 100 per cui si ha questo risultato, proviamo con 101. Si ha

$$x_2 = 101_{10} = 1100101_2 = (0.1100101)_2\, 2^7, \quad \widetilde{x}_2 = (0.11001)_2\, 2^7 = \widetilde{x}_1.$$

Quindi $\widetilde{x}_1/\widetilde{x}_2 = 1$, per cui il minimo cercato è 2.

Esercizio 2

L'errore inerente è

$$\epsilon_{in} = c_x\, \epsilon_x, \quad \text{dove} \quad c_x = \frac{x\, f'(x)}{f(x)} = \frac{3}{x^3-1}.$$

Il problema risulta malcondizionato nell'intorno di 1. L'errore algoritmico per il primo algoritmo è

$$\epsilon_{alg}^{(1)} = \epsilon^{(7)} + \epsilon^{(6)} + \epsilon^{(5)} + \epsilon^{(4)} + \frac{x^2+x}{x^2+x+1}\,\epsilon^{(3)} - \epsilon^{(2)} - \frac{x+1}{x^2+x+1}\,\epsilon^{(1)},$$

dove $\epsilon^{(1)}$ e $\epsilon^{(2)}$ sono gli errori locali di x^2 e x^3, $\epsilon^{(3)}$ e $\epsilon^{(4)}$ sono gli errori locali delle addizioni x^2+x e x^2+x+1, $\epsilon^{(5)}$ è l'errore locale della sottrazione $x-1$, $\epsilon^{(6)}$ è l'errore locale della divisione $(x-1)/x^3$ e $\epsilon^{(7)}$ è l'errore locale dell'ultima moltiplicazione. Poiché

$$\left|\frac{x^2+x}{x^2+x+1}\right| < 1, \qquad \left|\frac{x+1}{x^2+x+1}\right| \le 1, \quad \text{per ogni } x \text{ reale},$$

maggiorando si ha

$$\left|\epsilon_{alg}^{(1)}\right| < 7u.$$

Per il secondo algoritmo si ha

$$\epsilon_{alg}^{(2)} = \epsilon^{(9)} - \frac{1}{x^3-1}\big(\epsilon^{(8)} - \epsilon^{(2)} - \epsilon^{(1)}\big),$$

dove $\epsilon^{(8)}$ è l'errore locale della divisione $1/x^3$ e $\epsilon^{(9)}$ è l'errore locale dell'ultima sottrazione. Maggiorando si ha

$$\left|\epsilon_{alg}^{(2)}\right| < u\Big(1 + \frac{3}{|x^3-1|}\Big),$$

e $\left|\epsilon_{alg}^{(2)}\right|$ non è superiormente limitato nell'intorno di 1. Confrontando le maggiorazioni dei due errori algoritmici, e tenendo conto della complessità di calcolo, il primo algoritmo appare preferibile solo nell'intorno di 1, in cui comunque il problema è malcondizionato.

Esercizio 3

(a) Poiché $f(x) = (x-1)(x^2-kx-k)$, l'equazione $f(x)=0$ ha la soluzione $\alpha = 1$ per ogni k. Le altre due soluzioni si ricavano imponendo che $x^2-kx-k=0$ e sono $\beta = k-\sqrt{k^2+4k})/2$ e $\gamma = k+\sqrt{k^2+4k})/2$. Per $k>0$ (come imposto dal testo) è $k^2+4k>0$ per cui tali soluzioni sono reali e distinte, e $\beta<0, \;\; \gamma>0$. Inoltre $\gamma<1$ per $k<1/2$, $\gamma=1$ per $k=1/2$ e $\gamma>1$ per $k>1/2$. Perciò l'equazione $f(x)=0$ ha tre soluzioni reali e distinte

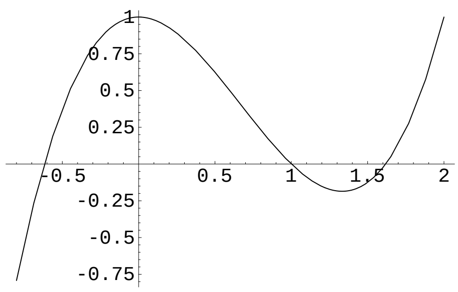
$$\beta = \frac{k-\sqrt{k^2+4k}}{2} \;<\; \alpha = 1 \;<\; \gamma = \frac{k+\sqrt{k^2+4k}}{2}$$

per $k>1/2$. Per $k=1/2$ è $\gamma=\alpha$ per cui l'equazione ha le due soluzioni $\beta=-1/2$ di molteplicità 1 e $\alpha=1$ di molteplicità 2. Non esistono valori di $k>0$ per cui l'equazione ha la sola soluzione reale $\alpha=1$.

(b) Per $k=1$ è $f(x) = x^3-2x^2+1 = (x-1)(x^2-x-1)$ e

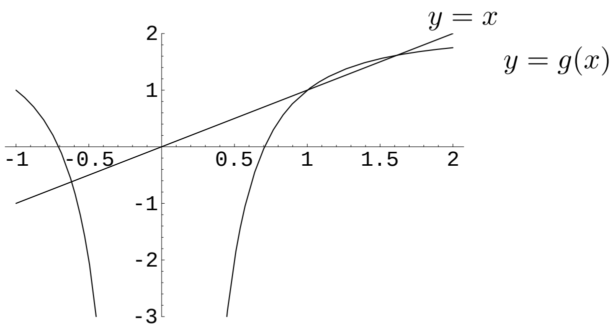
$$\beta = \frac{1-\sqrt{5}}{2} \sim -0.618, \quad \alpha = 1, \quad \gamma = \frac{1+\sqrt{5}}{2} \sim 1.618.$$

Il grafico della $f(x)$ è



Poiché $f'(x) = 3x^2 - 4x$ e $f''(x) = 6x - 4$, vi sono un punto di massimo $M = 0$, un punto di flesso $F = 2/3$ e un punto di minimo $m = 4/3$. Il metodo delle tangenti converge a β se $x_0 < 0$, converge ad α se $2/3 < x_0 < 1$ e converge a γ se $x_0 > 4/3$. Nulla si può dire a priori sulla convergenza se $0 < x_0 < 2/3$ e $1 < x_0 < 4/3$. In ogni caso l'ordine del metodo è 2.

(c) È $g(x) = 2 - 1/x^2$. I grafici di $y = x$ e $y = g(x)$ sono



Le soluzioni sono ovviamente quelle indicate al punto (b). Dal grafico risulta evidente che solo γ è punto attrattivo e che il metodo iterativo converge a γ se $x_0 > 1$. Poiché $g(-x) = g(x)$, si ha convergenza a γ anche se $x_0 < -1$. Inoltre se si sceglie $0 < x_0 < 1$ e nella successione generata non accade mai che $x_i = 0$, prima o poi verrà generato un punto $x_i < -1$, per cui si ha comunque convergenza a γ. Riassumendo, il metodo iterativo converge a γ per qualunque punto iniziale x_0 diverso da 0 e ± 1, purché non si abbia mai $x_i = 0$. L'ordine di convergenza a γ è 1.

Compito B

Lo svolgimento è sostanzialmente uguale a quello del compito A.