

11. Metodi iterativi per i sistemi non lineari

Gran parte delle considerazioni svolte e dei risultati ottenuti nei paragrafi precedenti riguardo ai metodi iterativi per le singole equazioni possono essere estesi anche al caso di sistemi di equazioni non lineari. In tale generalizzazione saranno utilizzati metodi e concetti dell'algebra lineare, come la norma e il raggio spettrale di una matrice, per i quali si rimanda a [3].

Sia Ω un sottosinsieme aperto di \mathbf{R}^n ; date le funzioni $\mathbf{f}, \mathbf{g} : \Omega \rightarrow \mathbf{R}^n$, il sistema da risolvere può essere rappresentato nella forma

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}, \quad (48)$$

oppure nella forma

$$\mathbf{x} = \mathbf{g}(\mathbf{x}), \quad (49)$$

dove

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix}.$$

Se $\mathbf{f}(\mathbf{x}), \mathbf{g}(\mathbf{x}) \in C^1(\Omega)$, si considerano le matrici jacobiane $J(\mathbf{x})$ e $H(\mathbf{x})$ rispettivamente di \mathbf{f} e di \mathbf{g} :

$$[J(\mathbf{x})]_{rs} = \frac{\partial f_r(\mathbf{x})}{\partial x_s} \quad \text{e} \quad [H(\mathbf{x})]_{rs} = \frac{\partial g_r(\mathbf{x})}{\partial x_s}.$$

Il metodo iterativo (5) si generalizza nella forma

$$\mathbf{x}^{(i+1)} = \mathbf{g}(\mathbf{x}^{(i)}), \quad i = 0, 1, \dots \quad (50)$$

La successione (50), formata da vettori di \mathbf{R}^n , è convergente alla soluzione $\boldsymbol{\alpha} \in \Omega$ se

$$\lim_{i \rightarrow \infty} \|\mathbf{x}^{(i)} - \boldsymbol{\alpha}\| = 0,$$

per qualche norma vettoriale, e poiché le norme vettoriali su \mathbf{R}^n sono tutte topologicamente equivalenti, la convergenza non dipende dalla particolare norma considerata.

Anche nel caso dei sistemi di equazioni non lineari, analogamente al teorema 3.2, se $\mathbf{g}(\mathbf{x})$ è continua e se la successione (50) appartiene ad Ω , allora il limite della successione $\{\mathbf{x}^{(i)}\}$, se esiste, è soluzione di (49). Vale il seguente teorema, analogo al teorema 3.3.

3.37 Teorema. Sia α soluzione del sistema (49). Se esiste un intorno $S \subseteq \Omega$ di α :

$$S = \{\mathbf{x} \in \mathbf{R}^n : \|\mathbf{x} - \alpha\|_\infty \leq \sigma, \sigma > 0\},$$

tale che $\mathbf{g}(\mathbf{x}) \in C^1(S)$ e

$$\|H(\mathbf{x})\|_\infty < 1, \quad \text{per } \mathbf{x} \in S, \quad (51)$$

allora, scelto $\mathbf{x}^{(0)} \in S$, la successione $\{\mathbf{x}^{(i)}\}$ ottenuta dalla (50) converge ad α .

Dim. Posto

$$\lambda = \max_{\mathbf{x} \in S} \|H(\mathbf{x})\|_\infty,$$

dalla (51) risulta $\lambda < 1$. Si dimostra per induzione che

$$\|\mathbf{x}^{(i)} - \alpha\|_\infty \leq \lambda^i \sigma. \quad (52)$$

Per $i = 0$ la (52) è vera. Per $i > 0$, sviluppando $g_r(\mathbf{x})$, $r = 1, \dots, n$, in serie di Taylor attorno al punto α , si ha:

$$g_r(\mathbf{x}^{(i-1)}) = g_r(\alpha) + \sum_{s=1}^n \frac{\partial g_r(\xi_r)}{\partial x_s} (x_s^{(i-1)} - \alpha_s),$$

dove ξ_r è un punto del segmento che unisce $\mathbf{x}^{(i-1)}$ e α , e passando ai moduli si ha:

$$\begin{aligned} |g_r(\mathbf{x}^{(i-1)}) - g_r(\alpha)| &\leq \sum_{s=1}^n \left| \frac{\partial g_r(\xi_r)}{\partial x_s} \right| |x_s^{(i-1)} - \alpha_s| \\ &\leq \sum_{s=1}^n \left| \frac{\partial g_r(\xi_r)}{\partial x_s} \right| \|\mathbf{x}^{(i-1)} - \alpha\|_\infty \\ &\leq \|H(\xi_r)\|_\infty \|\mathbf{x}^{(i-1)} - \alpha\|_\infty. \end{aligned}$$

Per l'ipotesi induttiva è $\xi_r \in S$ e quindi vale

$$|g_r(\mathbf{x}^{(i-1)}) - g_r(\alpha)| \leq \lambda \|\mathbf{x}^{(i-1)} - \alpha\|_\infty, \quad r = 1, \dots, n,$$

e passando alle norme si ha:

$$\|\mathbf{g}(\mathbf{x}^{(i-1)}) - \mathbf{g}(\alpha)\|_\infty \leq \lambda \|\mathbf{x}^{(i-1)} - \alpha\|_\infty.$$

Dalla (50) risulta

$$\|\mathbf{x}^{(i)} - \alpha\|_\infty \leq \lambda \|\mathbf{x}^{(i-1)} - \alpha\|_\infty,$$

da cui, per l'ipotesi induttiva, segue

$$\|\mathbf{x}^{(i)} - \boldsymbol{\alpha}\|_{\infty} \leq \lambda^i \sigma,$$

e, passando al limite, risulta:

$$\lim_{i \rightarrow \infty} \|\mathbf{x}^{(i)} - \boldsymbol{\alpha}\|_{\infty} = 0. \quad \blacksquare$$

3.38 Esempio. Si consideri il sistema non lineare

$$\begin{cases} x_1 = g_1(x_1, x_2) = \frac{1}{4} (x_1^2 + x_2^2) \\ x_2 = g_2(x_1, x_2) = \sin(x_1 + 1). \end{cases}$$

Per individuare intorno delle soluzioni del sistema, si considerino le curve di equazioni $x_1 = g_1(x_1, x_2)$ (la circonferenza di centro $(2,0)$ e raggio 2) e $x_2 = g_2(x_1, x_2)$ (la sinusoide con l'argomento traslato di 1), rappresentate nella figura 3.17.

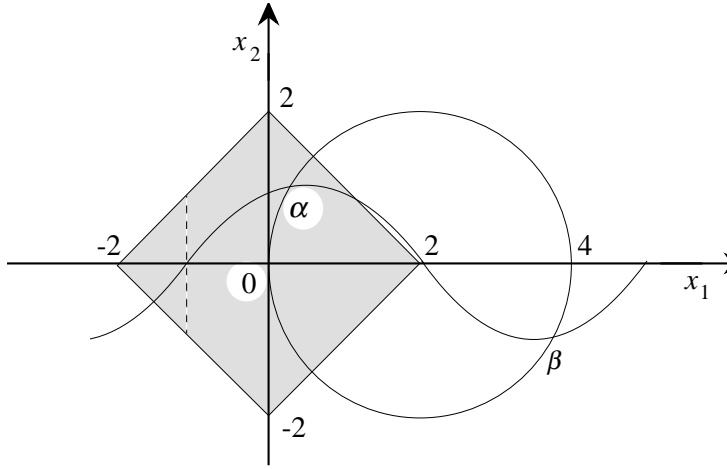


Fig. 3.17 - Grafici relativi al sistema dell'esempio 3.38.

Vi sono quindi due soluzioni reali $\boldsymbol{\alpha} \in [0, 0.5] \times [0.5, 1]$ e $\boldsymbol{\beta} \in [3, 4] \times [-1, -0.5]$. La matrice

$$H(\mathbf{x}) = \begin{bmatrix} x_1/2 & x_2/2 \\ \cos(x_1 + 1) & 0 \end{bmatrix}$$

è tale che $\|H(\mathbf{x})\|_{\infty} < 1$ se $|x_1|/2 + |x_2|/2 < 1$ e $|\cos(x_1 + 1)| < 1$. Il dominio in cui $\|H(\mathbf{x})\|_{\infty} < 1$, segnato in grigio sul grafico, è formato dai

punti appartenenti al quadrato di centro l'origine e un vertice nel punto $(0, 2)$, esclusi i punti di ascissa $x_1 = -1$ in cui $\cos(x_1 + 1) = 1$. Esiste allora un intorno S di α con le proprietà riportate nel teorema 3.37, contenuto in tale dominio, e quindi per ogni $\mathbf{x}^{(0)} \in S$ la successione $\{\mathbf{x}^{(i)}\}$ definita da

$$\begin{cases} x_1^{(i+1)} = \frac{1}{4} (x_1^{(i)^2} + x_2^{(i)^2}) \\ x_2^{(i+1)} = \sin(x_1^{(i)} + 1) \end{cases}$$

converge ad α . Assumendo $\mathbf{x}^{(0)} = (0, 1) \in S$, si ottiene la successione:

i	$x_1^{(i)}$	$x_2^{(i)}$
1	0.2500000	0.8414710
2	0.1926433	0.9489847
3	0.2344208	0.9293481
4	0.2296602	0.9439572
\vdots	\vdots	\vdots
17	0.2372839	0.9448983
18	0.2372841	0.9448983

Però anche partendo da $\mathbf{x}^{(0)} = (2, 0) \notin S$ si ottiene una successione convergente ad α . Non è invece possibile assicurare la convergenza a β . ■

Se $\mathbf{g}(\mathbf{x}) \in C^1(\Omega)$, allora la funzione \mathbf{g} è *totalmente differenziabile* in Ω , cioè per ogni $\mathbf{x} \in \Omega$ è

$$\lim_{\mathbf{x}' \rightarrow \mathbf{x}} \frac{\|\mathbf{g}(\mathbf{x}') - \mathbf{g}(\mathbf{x}) - H(\mathbf{x})(\mathbf{x}' - \mathbf{x})\|}{\|\mathbf{x}' - \mathbf{x}\|} = 0, \quad (53)$$

dove $\|\cdot\|$ è una qualsiasi norma vettoriale [9]. Il seguente teorema fornisce una condizione di convergenza che utilizza una generica norma matriciale indotta e che è quindi più generale di quella del teorema 3.37.

3.39 Teorema. Sia $\mathbf{g}(\mathbf{x}) \in C^1(\Omega)$ e sia α una soluzione del sistema (49). Si indichi con $\rho(H(\alpha))$ il raggio spettrale di $H(\alpha)$. Se $\rho(H(\alpha)) < 1$, esistono una norma vettoriale $\|\cdot\|$ e un $\sigma > 0$ tali che, posto

$$S = \{\mathbf{x} \in \mathbf{R}^n : \|\mathbf{x} - \alpha\| \leq \sigma\},$$

risulta $S \subseteq \Omega$ e la successione (50) converge per ogni $\mathbf{x}^{(0)} \in S$.

Dim. Sia ϵ tale che

$$0 < \epsilon < \frac{1 - \rho(H(\alpha))}{2}.$$

Esiste allora una norma matriciale indotta tale che

$$\|H(\boldsymbol{\alpha})\| \leq \rho(H(\boldsymbol{\alpha})) + \epsilon$$

(si veda [3], teorema 3.12). Per la (53), applicata con $\mathbf{x} = \boldsymbol{\alpha}$, esiste $\sigma > 0$ tale che

$$\|\mathbf{g}(\mathbf{x}') - \mathbf{g}(\boldsymbol{\alpha}) - H(\boldsymbol{\alpha})(\mathbf{x}' - \boldsymbol{\alpha})\| \leq \epsilon \|\mathbf{x}' - \boldsymbol{\alpha}\|$$

per ogni $\mathbf{x}' \in S$ e $S \subseteq \Omega$. Quindi per $\mathbf{x}' \in S$ risulta

$$\|\mathbf{g}(\mathbf{x}') - \mathbf{g}(\boldsymbol{\alpha})\| \leq \|\mathbf{g}(\mathbf{x}') - \mathbf{g}(\boldsymbol{\alpha}) - H(\boldsymbol{\alpha})(\mathbf{x}' - \boldsymbol{\alpha})\| + \|H(\boldsymbol{\alpha})(\mathbf{x}' - \boldsymbol{\alpha})\| \leq \lambda \|\mathbf{x}' - \boldsymbol{\alpha}\|$$

dove $\lambda = \rho(H(\boldsymbol{\alpha})) + 2\epsilon$. Poiché per la scelta di ϵ è $\lambda < 1$, si dimostra per induzione, analogamente a quanto fatto nel teorema 3.37, che per $\mathbf{x}^{(0)} \in S$ è

$$\|\mathbf{x}^{(i)} - \boldsymbol{\alpha}\| \leq \lambda^i \|\mathbf{x}^{(0)} - \boldsymbol{\alpha}\|,$$

da cui segue che la successione $\{\mathbf{x}^{(i)}\}$ è localmente convergente. ■

3.40 Esempio. Il sistema non lineare

$$\begin{cases} x_1 = \frac{1}{4} (x_1^2 + x_2^2) \\ x_2 = -x_1 + 2 \end{cases}$$

ha due soluzioni reali $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$, punti di intersezione della circonferenza di centro (2,0) e raggio 2 con la retta passante per il centro della circonferenza e per il punto (0,2) (si veda la figura 3.18).

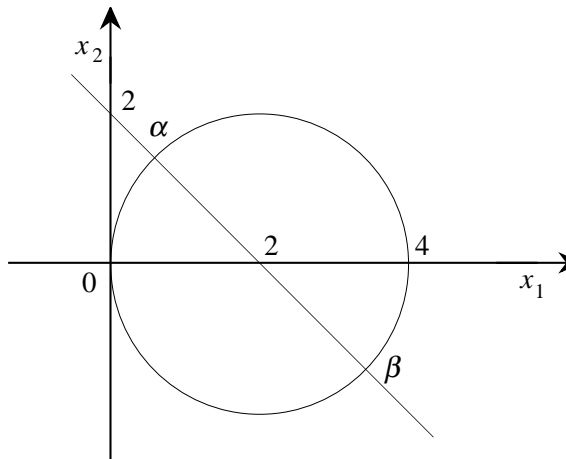


Fig. 3.18 - Grafici relativi al sistema dell'esempio 3.40.

La matrice

$$H(\mathbf{x}) = \begin{bmatrix} x_1/2 & x_2/2 \\ -1 & 0 \end{bmatrix}$$

è tale che $\|H(\mathbf{x})\|_\infty \geq 1$ per ogni \mathbf{x} e nella soluzione è $\|H(\boldsymbol{\alpha})\|_\infty = 1$. Ma

$$\rho(H(\boldsymbol{\alpha})) = \sqrt{\alpha_2/2} < 1, \quad \text{dove } \boldsymbol{\alpha} = (\alpha_1, \alpha_2),$$

e quindi, per opportune scelte di $\mathbf{x}^{(0)}$, il metodo iterativo

$$\begin{cases} x_1^{(i+1)} = \frac{1}{4} (x_1^{(i)^2} + x_2^{(i)^2}) \\ x_2^{(i+1)} = -x_1^{(i)} + 2 \end{cases}$$

converge ad $\boldsymbol{\alpha}$. Assumendo $\mathbf{x}^{(0)} = (1, 1)$, si ottiene la successione:

i	$x_1^{(i)}$	$x_2^{(i)}$
1	0.5000000	1.000000
2	0.3125000	1.500000
3	0.5869141	1.687500
4	0.7980311	1.413086
\vdots	\vdots	\vdots
77	0.5857866	1.414213
78	0.5857859	1.414213

■

12. Metodo di Newton-Raphson

Analogamente al caso delle singole equazioni, da un sistema dato nella forma (48) può essere ottenuto un sistema della forma (49) con le stesse soluzioni ponendo

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} - [C(\mathbf{x})]^{-1}\mathbf{f}(\mathbf{x}),$$

dove $C(\mathbf{x})$ è una matrice i cui elementi sono funzioni di \mathbf{x} e il cui determinante non è identicamente nullo. Se $\mathbf{f} \in C^1(\Omega)$, scegliendo $C(\mathbf{x}) = J(\mathbf{x})$ si ottiene il metodo di *Newton-Raphson*, che generalizza ai sistemi di equazioni il metodo delle tangenti:

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} - [J(\mathbf{x}^{(i)})]^{-1}\mathbf{f}(\mathbf{x}^{(i)}), \quad \det J(\mathbf{x}^{(i)}) \neq 0. \quad (54)$$

Il vettore $\mathbf{x}^{(i+1)} - \mathbf{x}^{(i)}$ è allora soluzione del sistema lineare di n equazioni in n incognite

$$J(\mathbf{x}^{(i)})[\mathbf{x}^{(i+1)} - \mathbf{x}^{(i)}] = -\mathbf{f}(\mathbf{x}^{(i)}). \quad (55)$$

Ad esempio, per la risoluzione del sistema di 2 equazioni

$$\begin{cases} f_1(x_1, x_2) = 0 \\ f_2(x_1, x_2) = 0, \end{cases} \quad (56)$$

con il metodo di Newton-Raphson si deve risolvere ad ogni iterazione il sistema lineare:

$$\begin{cases} \frac{\partial f_1(\mathbf{x}^{(i)})}{\partial x_1} \theta_1 + \frac{\partial f_1(\mathbf{x}^{(i)})}{\partial x_2} \theta_2 = -f_1(\mathbf{x}^{(i)}) \\ \frac{\partial f_2(\mathbf{x}^{(i)})}{\partial x_1} \theta_1 + \frac{\partial f_2(\mathbf{x}^{(i)})}{\partial x_2} \theta_2 = -f_2(\mathbf{x}^{(i)}), \end{cases} \quad (57)$$

in cui con θ_1 e θ_2 sono state indicate le componenti del vettore $\mathbf{x}^{(i+1)} - \mathbf{x}^{(i)}$. Del metodo di Newton-Raphson si può dare la seguente interpretazione geometrica: le equazioni (56) rappresentano due curve, intersezioni delle superfici $z = f_1(x_1, x_2)$ e $z = f_2(x_1, x_2)$ con il piano $z = 0$; fissato in questo piano il punto $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)})$, si considerano i piani tangenti alle due superfici rispettivamente nei punti $(\mathbf{x}^{(0)}, f_1(\mathbf{x}^{(0)}))$ e $(\mathbf{x}^{(0)}, f_2(\mathbf{x}^{(0)}))$:

$$\begin{aligned} z &= \frac{\partial f_1(\mathbf{x}^{(0)})}{\partial x_1} (x_1 - x_1^{(0)}) + \frac{\partial f_1(\mathbf{x}^{(0)})}{\partial x_2} (x_2 - x_2^{(0)}) + f_1(\mathbf{x}^{(0)}), \\ z &= \frac{\partial f_2(\mathbf{x}^{(0)})}{\partial x_1} (x_1 - x_1^{(0)}) + \frac{\partial f_2(\mathbf{x}^{(0)})}{\partial x_2} (x_2 - x_2^{(0)}) + f_2(\mathbf{x}^{(0)}); \end{aligned}$$

si considerano poi le rette intersezioni di questi due piani con il piano $z = 0$. Il punto $\mathbf{x}^{(1)} = (x_1^{(1)}, x_2^{(1)})$, soluzione di (57) per $i = 0$, è l'intersezione di queste due rette. Per i passi successivi si procede in modo analogo.

3.41 Teorema. Sia $\mathbf{f}(\mathbf{x}) \in C^2(\Omega)$ e sia $\boldsymbol{\alpha} \in \Omega$ soluzione del sistema (48). Se $J(\mathbf{x})$ è non singolare in Ω , esiste un intorno $S \subseteq \Omega$ di $\boldsymbol{\alpha}$ tale che, se $\mathbf{x}^{(0)} \in S$, la successione (54) converge ed inoltre per ogni norma vettoriale esiste una costante β tale che

$$\|\mathbf{x}^{(i+1)} - \boldsymbol{\alpha}\| \leq \beta \|\mathbf{x}^{(i)} - \boldsymbol{\alpha}\|^2, \quad i = 0, 1, \dots \quad (58)$$

Dim. Per il metodo di Newton-Raphson è

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} - [J(\mathbf{x})]^{-1}\mathbf{f}(\mathbf{x}),$$

da cui, ponendo $K(\mathbf{x}) = [J(\mathbf{x})]^{-1}$, si ha:

$$g_r(\mathbf{x}) = x_r - \sum_{s=1}^n k_{rs}(\mathbf{x}) f_s(\mathbf{x}), \quad r = 1, \dots, n.$$

Gli elementi della matrice $K(\mathbf{x})$, essendo funzioni razionali degli elementi di $J(\mathbf{x})$, con denominatore uguale a $\det J(\mathbf{x})$, sono derivabili con continuità. Derivando rispetto a x_t , $t = 1, \dots, n$, si ha

$$\begin{aligned} h_{rt}(\mathbf{x}) &= \frac{\partial g_r(\mathbf{x})}{\partial x_t} = \delta_{rt} - \sum_{s=1}^n \frac{\partial k_{rs}(\mathbf{x})}{\partial x_t} f_s(\mathbf{x}) - \sum_{s=1}^n k_{rs}(\mathbf{x}) \frac{\partial f_s(\mathbf{x})}{\partial x_t} \\ &= - \sum_{s=1}^n \frac{\partial k_{rs}(\mathbf{x})}{\partial x_t} f_s(\mathbf{x}), \quad \text{dove} \quad \delta_{rt} = \begin{cases} 1 & \text{per } r = t, \\ 0 & \text{altrimenti,} \end{cases} \end{aligned}$$

perché

$$\sum_{s=1}^n k_{rs}(\mathbf{x}) \frac{\partial f_s(\mathbf{x})}{\partial x_t} = \delta_{rt}.$$

Si ha allora $H(\boldsymbol{\alpha}) = O$ e quindi $\rho(H(\boldsymbol{\alpha})) = 0$, allora per il teorema 3.39 esiste un intorno S di $\boldsymbol{\alpha}$ in cui si ha convergenza. Dalla (55) si ha

$$J(\mathbf{x}^{(i)}) (\mathbf{x}^{(i+1)} - \boldsymbol{\alpha}) = J(\mathbf{x}^{(i)}) (\mathbf{x}^{(i)} - \boldsymbol{\alpha}) - \mathbf{f}(\mathbf{x}^{(i)}). \quad (59)$$

Indicate con $S_r(x)$, $r = 1, \dots, n$, le matrici hessiane delle funzioni $f_r(\mathbf{x})$, i cui elementi sono

$$(S_r(\mathbf{x}))_{st} = \frac{\partial^2 f_r(\mathbf{x})}{\partial x_s \partial x_t}, \quad (60)$$

dalla formula di Taylor si ha

$$\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{f}(\mathbf{x}^{(i)}) + J(\mathbf{x}^{(i)}) (\boldsymbol{\alpha} - \mathbf{x}^{(i)}) + \mathbf{v}, \quad (61)$$

in cui \mathbf{v} è un vettore di resti, la cui r -esima componente è

$$v_r = \frac{1}{2} (\mathbf{x}^{(i)} - \boldsymbol{\alpha})^T S_r(\boldsymbol{\xi}_r) (\mathbf{x}^{(i)} - \boldsymbol{\alpha}),$$

e $\boldsymbol{\xi}_r$ appartiene al segmento di estremi $\boldsymbol{\alpha}$ e $\mathbf{x}^{(i)}$, e quindi è

$$|v_r| \leq \frac{n}{2} \|S_r(\boldsymbol{\xi}_r)\|_{\infty} \|\mathbf{x}^{(i)} - \boldsymbol{\alpha}\|_{\infty}^2.$$

Sostituendo la (61) nella (59), poiché $\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{0}$, risulta

$$\mathbf{x}^{(i+1)} - \boldsymbol{\alpha} = K(\mathbf{x}^{(i)}) \mathbf{v},$$

da cui

$$\|\mathbf{x}^{(i+1)} - \boldsymbol{\alpha}\|_{\infty} \leq \gamma \|\mathbf{x}^{(i)} - \boldsymbol{\alpha}\|_{\infty}^2,$$

dove

$$\gamma = \frac{n}{2} \max_{\mathbf{x} \in S} \|K(\mathbf{x})\|_{\infty} \max_{\substack{\mathbf{x} \in S \\ r=1, \dots, n}} \|S_r(\mathbf{x})\|_{\infty}.$$

Per l'equivalenza topologica delle norme vettoriali ne segue che esiste una costante β per cui vale la (58) per ogni norma vettoriale. ■

La (58), che esprime la riduzione dell'errore all' i -esima iterazione, può essere considerata una generalizzazione della (20) per $p = 2$. Analogamente al caso di una singola equazione, se $\det J(\boldsymbol{\alpha}) \neq 0$ si può dire che il metodo di Newton-Raphson per risolvere sistemi di equazioni non lineari è almeno del secondo ordine.

3.42 Esempio. Per il sistema non lineare dell'esempio 3.38 si ha

$$\begin{cases} f_1(x_1, x_2) = x_1 - \frac{1}{4}(x_1^2 + x_2^2) = 0 \\ f_2(x_1, x_2) = x_2 - \sin(x_1 + 1) = 0, \end{cases}$$

e

$$J(\mathbf{x}) = \begin{bmatrix} 1 - x_1/2 & -x_2/2 \\ -\cos(x_1 + 1) & 1 \end{bmatrix}.$$

Poiché $2 \det J(\mathbf{x}) = 2 - x_1 - x_2 \cos(x_1 + 1)$, la matrice $J(\mathbf{x})$ è non singolare per i punti (x_1, x_2) tali che

$$x_2 \neq \frac{2 - x_1}{\cos(x_1 + 1)},$$

cioè i punti non appartenenti alla curva il cui grafico è rappresentato nella figura 3.19. In particolare risulta $\det J(\mathbf{x}) \neq 0$ in $\boldsymbol{\alpha}$ e in $\boldsymbol{\beta}$.

Con il metodo di Newton-Raphson, assumendo come punto iniziale $\mathbf{x}^{(0)} = (0, 1)$, si ottiene la successione convergente ad $\boldsymbol{\alpha}$:

i	$x_1^{(i)}$	$x_2^{(i)}$
1	0.2339326	0.9678653
2	0.2371014	0.9448433
3	0.2372839	0.9448981
4	0.2372841	0.9448984

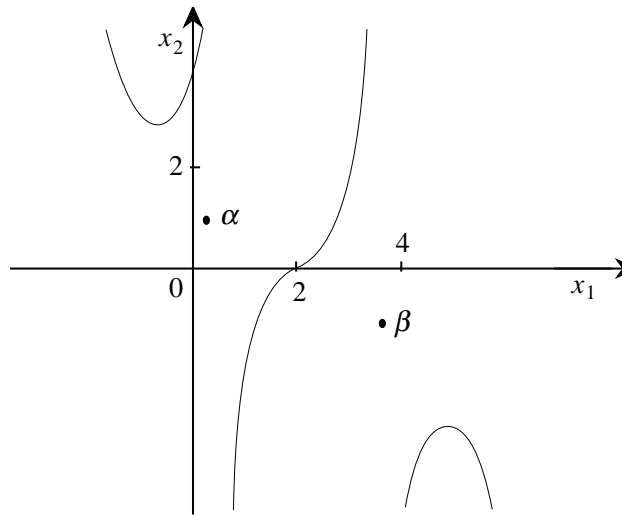


Fig. 3.19 - Grafico della funzione $x_2 = \frac{2 - x_1}{\cos(x_1 + 1)}$.

Con il metodo di Newton-Raphson si può approssimare anche la radice β . Assumendo come punto iniziale $\mathbf{x}^{(0)} = (4, -1)$, si ottiene la successione convergente:

i	$x_1^{(i)}$	$x_2^{(i)}$
1	3.732614	-1.034771
2	3.732522	-0.9997974
3	3.732164	-0.9998045
4	3.732164	-0.9998045

■

3.43 Esempio. Si consideri il sistema non lineare

$$\begin{cases} x_1 = f_1(x_1, x_2) = \frac{1}{3}(x_1 - x_2) + x_1^2 = 0 \\ x_2 = f_2(x_1, x_2) = \frac{1}{3}(-x_1 + x_2) + x_1x_2 = 0, \end{cases}$$

che ha le due soluzioni $\alpha = (0, 0)$ e $\beta = \left(-\frac{2}{3}, \frac{2}{3}\right)$, come risulta anche dalla figura 3.20, in cui sono rappresentate le curve di equazioni $f_1(\mathbf{x}) = 0$ e $f_2(\mathbf{x}) = 0$.

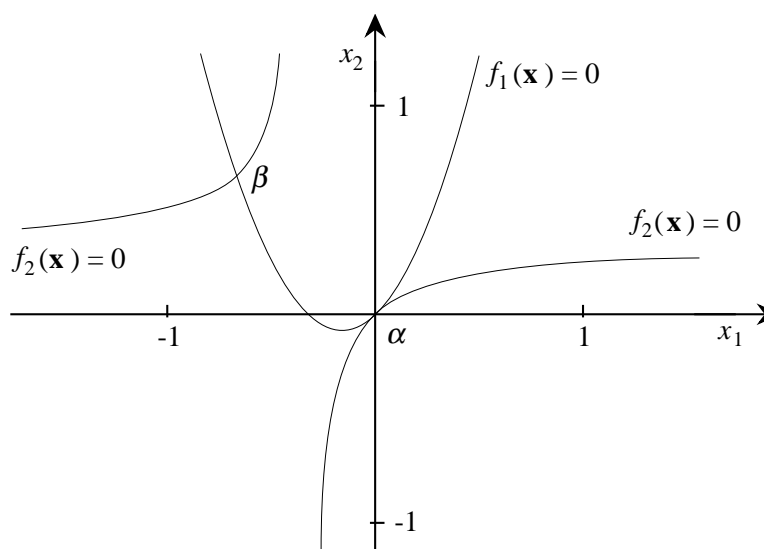


Fig. 3.20 - Grafici relativi al sistema dell'esempio 3.43.

È

$$J(\mathbf{x}) = \begin{bmatrix} \frac{1}{3} + 2x_1 & -\frac{1}{3} \\ -\frac{1}{3} + x_2 & \frac{1}{3} + x_1 \end{bmatrix},$$

quindi $\det J(\boldsymbol{\alpha}) = 0$ e $\det J(\boldsymbol{\beta}) \neq 0$. Applicando il metodo di Newton-Raphson a partire dal punto $\mathbf{x}^{(0)} = (0.5, 0.8)$, si ottiene la successione:

i	$x_1^{(i)}$	$x_2^{(i)}$
1	0.2697367	0.3289472
2	0.1399398	0.1481472
3	$0.7080776 \cdot 10^{-1}$	$0.7151115 \cdot 10^{-1}$
4	$0.3548320 \cdot 10^{-1}$	$0.3551690 \cdot 10^{-1}$
\vdots	\vdots	\vdots
18	$0.2145697 \cdot 10^{-5}$	$0.2145697 \cdot 10^{-5}$
19	$0.1072881 \cdot 10^{-5}$	$0.1072881 \cdot 10^{-5}$

che converge ad $\boldsymbol{\alpha}$. Applicando il metodo a partire dal punto $\mathbf{x}^{(0)} = (-2, 2)$, si ottiene la successione:

i	$x_1^{(i)}$	$x_2^{(i)}$
1	-1.199999	1.199999
2	-0.8307689	0.8307689
3	-0.6937348	0.6937348
4	-0.6676828	0.6676828
5	-0.6666679	0.6666679

che converge ad β . Mentre la seconda successione converge con ordine almeno 2, la prima converge molto più lentamente (linearmente) poiché per α non sono verificate le ipotesi del teorema 3.41: risulta infatti

$$\|\mathbf{x}^{(i+1)} - \alpha\|_\infty \approx 0.5 \|\mathbf{x}^{(i)} - \alpha\|_\infty. \quad \blacksquare$$

Il teorema 3.41 fornisce condizioni di convergenza locale per il metodo di Newton-Raphson. Però, analogamente al caso di una singola equazione, anche per i sistemi è possibile dare delle condizioni sufficienti di convergenza su insiemi di \mathbf{R}^n . Per questo è necessario premettere alcune definizioni e un teorema. Nel seguito le disuguaglianze fra vettori o fra matrici vanno intese componente per componente.

3.44 Definizione. Un insieme D di \mathbf{R}^n si dice *convesso* se

$$\lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \in D$$

per ogni coppia $\mathbf{x}, \mathbf{y} \in D$ e ogni $\lambda \in [0, 1]$. ■

3.45 Definizione. Sia D un insieme convesso di \mathbf{R}^n . Una funzione $\mathbf{f} : D \rightarrow \mathbf{R}^n$ si dice *convessa* su D se

$$\mathbf{f}(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \leq \lambda \mathbf{f}(\mathbf{x}) + (1 - \lambda) \mathbf{f}(\mathbf{y}), \quad (62)$$

per ogni coppia $\mathbf{x}, \mathbf{y} \in D$ e ogni $\lambda \in [0, 1]$. ■

Dalla definizione 3.45 segue che la funzione $\mathbf{f}(\mathbf{x})$ è convessa su D se e solo se tutte le funzioni $f_r(\mathbf{x})$, $r = 1, \dots, n$, sono convesse su D . Se $\mathbf{f}(\mathbf{x}) \in C^2(D)$, la condizione di convessità definita in 3.45 è equivalente alla condizione che le matrici hessiane $S_r(\mathbf{x})$, $r = 1, \dots, n$, introdotte in (60), siano semidefinite positive per $\mathbf{x} \in D$ (si veda l'esercizio 3.44).

3.46 Teorema. Sia D un insieme convesso e $\mathbf{f}(\mathbf{x}) \in C^1(D)$, allora $\mathbf{f}(\mathbf{x})$ è convessa su D se e solo se

$$\mathbf{f}(\mathbf{v}) - \mathbf{f}(\mathbf{u}) \geq J(\mathbf{u})(\mathbf{v} - \mathbf{u}) \quad (63)$$

per ogni $\mathbf{u}, \mathbf{v} \in D$.

Dim. Siano $\mathbf{x}, \mathbf{y} \in D$, $\lambda \in [0, 1]$ e $\mathbf{z}(\lambda) = \lambda\mathbf{x} + (1 - \lambda)\mathbf{y}$. Poiché D è convesso, $\mathbf{z}(\lambda) \in D$. Dalla (63), si ha:

$$\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{z}(\lambda)) \geq J(\mathbf{z}(\lambda))(\mathbf{x} - \mathbf{z}(\lambda)),$$

$$\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{z}(\lambda)) \geq J(\mathbf{z}(\lambda))(\mathbf{y} - \mathbf{z}(\lambda)).$$

Moltiplicando entrambi i membri di queste due disuguaglianze rispettivamente per λ e $(1 - \lambda)$, e sommando membro a membro, si ottiene

$$\lambda\mathbf{f}(\mathbf{x}) + (1 - \lambda)\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{z}(\lambda)) \geq J(\mathbf{z}(\lambda))(\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} - \mathbf{z}(\lambda)) = \mathbf{0},$$

da cui segue la disuguaglianza (62). Viceversa, se $\mathbf{f}(\mathbf{x})$ è convessa su D , da (62) si ha per $\mathbf{u}, \mathbf{v} \in D$, $\lambda \in [0, 1]$ e $\mathbf{w}(\lambda) = \lambda\mathbf{v} + (1 - \lambda)\mathbf{u} \in D$:

$$\mathbf{f}(\mathbf{v}) - \mathbf{f}(\mathbf{u}) \geq \frac{1}{\lambda}[\mathbf{f}(\mathbf{w}(\lambda)) - \mathbf{f}(\mathbf{u})]. \quad (64)$$

Poiché $\lim_{\lambda \rightarrow 0} \mathbf{w}(\lambda) = \mathbf{u}$, dalla (53) segue che

$$\lim_{\lambda \rightarrow 0} \frac{\|\mathbf{f}(\mathbf{w}(\lambda)) - \mathbf{f}(\mathbf{u}) - J(\mathbf{u})(\mathbf{w}(\lambda) - \mathbf{u})\|}{\|\mathbf{w}(\lambda) - \mathbf{u}\|} = 0,$$

e poiché $\|\mathbf{w}(\lambda) - \mathbf{u}\| = |\lambda|\|\mathbf{v} - \mathbf{u}\|$, si ottiene

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \|\mathbf{f}(\mathbf{w}(\lambda)) - \mathbf{f}(\mathbf{u}) - \lambda J(\mathbf{u})(\mathbf{v} - \mathbf{u})\| = 0,$$

e quindi

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} [\mathbf{f}(\mathbf{w}(\lambda)) - \mathbf{f}(\mathbf{u})] = J(\mathbf{u})(\mathbf{v} - \mathbf{u}). \quad (65)$$

Da (64) e (65) si ottiene

$$\mathbf{f}(\mathbf{v}) - \mathbf{f}(\mathbf{u}) \geq J(\mathbf{u})(\mathbf{v} - \mathbf{u}). \quad \blacksquare$$

Per il metodo di Newton-Raphson vale allora il seguente teorema.

3.47 Teorema. Siano $a_i, b_i \in \mathbf{R}$ per $i = 1, \dots, n$, e sia D l'insieme

$$D = \prod_{j=1}^n [a_j, b_j]$$

(cioè D è un intervallo di \mathbf{R}^n). Siano $\mathbf{f}(\mathbf{x}) \in C^1(D)$ una funzione convessa su D e $\boldsymbol{\alpha} \in D$ la soluzione del sistema $\mathbf{f}(\mathbf{x}) = \mathbf{0}$. Se per $\mathbf{x} \in D$, $J(\mathbf{x})$ è non singolare e $[J(\mathbf{x})]^{-1} \geq O$, per ogni scelta di $\mathbf{x}^{(0)} \in D$ tale che $\mathbf{f}(\mathbf{x}^{(0)}) \geq \mathbf{0}$, la successione (54) converge ad $\boldsymbol{\alpha}$ ed $\boldsymbol{\alpha}$ è l'unica soluzione del sistema nell'insieme D .

Dim. Si dimostra per induzione su i che

$$\boldsymbol{\alpha} \leq \mathbf{x}^{(i+1)} \leq \mathbf{x}^{(i)} \quad \text{e} \quad \mathbf{f}(\mathbf{x}^{(i+1)}) \geq \mathbf{0}, \quad i = 0, 1, \dots \quad (66)$$

Per $i = 0$, per l'ipotesi di convessità si ha

$$\mathbf{f}(\boldsymbol{\alpha}) - \mathbf{f}(\mathbf{x}^{(0)}) \geq J(\mathbf{x}^{(0)})(\boldsymbol{\alpha} - \mathbf{x}^{(0)}),$$

da cui si ha

$$-[J(\mathbf{x}^{(0)})]^{-1}\mathbf{f}(\mathbf{x}^{(0)}) \geq \boldsymbol{\alpha} - \mathbf{x}^{(0)},$$

e quindi, poiché $\mathbf{f}(\mathbf{x}^{(0)}) \geq \mathbf{0}$ per ipotesi, ne segue

$$\mathbf{x}^{(0)} \geq \mathbf{x}^{(1)} = \mathbf{x}^{(0)} - [J(\mathbf{x}^{(0)})]^{-1}\mathbf{f}(\mathbf{x}^{(0)}) \geq \boldsymbol{\alpha},$$

quindi $\mathbf{x}^{(1)} \in D$. Inoltre per l'ipotesi di convessità,

$$\mathbf{f}(\mathbf{x}^{(1)}) - \mathbf{f}(\mathbf{x}^{(0)}) \geq J(\mathbf{x}^{(0)})(\mathbf{x}^{(1)} - \mathbf{x}^{(0)}) = -\mathbf{f}(\mathbf{x}^{(0)}),$$

per cui $\mathbf{f}(\mathbf{x}^{(1)}) \geq \mathbf{0}$. Per $i > 0$, per l'ipotesi di convessità si ha

$$\mathbf{f}(\boldsymbol{\alpha}) - \mathbf{f}(\mathbf{x}^{(i)}) \geq J(\mathbf{x}^{(i)})(\boldsymbol{\alpha} - \mathbf{x}^{(i)}),$$

da cui si ha

$$-[J(\mathbf{x}^{(i)})]^{-1}\mathbf{f}(\mathbf{x}^{(i)}) \geq \boldsymbol{\alpha} - \mathbf{x}^{(i)},$$

e quindi, poiché $\mathbf{x}^{(i)} \in D$ e $\mathbf{f}(\mathbf{x}^{(i)}) \geq \mathbf{0}$ per l'ipotesi induttiva, segue che

$$\mathbf{x}^{(i)} \geq \mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} - [J(\mathbf{x}^{(i)})]^{-1}\mathbf{f}(\mathbf{x}^{(i)}) \geq \boldsymbol{\alpha}.$$

Quindi $\mathbf{x}^{(i+1)} \in D$. Inoltre per l'ipotesi di convessità si ha:

$$\mathbf{f}(\mathbf{x}^{(i+1)}) - \mathbf{f}(\mathbf{x}^{(i)}) \geq J(\mathbf{x}^{(i)})(\mathbf{x}^{(i+1)} - \mathbf{x}^{(i)}) = -\mathbf{f}(\mathbf{x}^{(i)}),$$

per cui $\mathbf{f}(\mathbf{x}^{(i+1)}) \geq \mathbf{0}$. Dalla (66) segue allora che le successioni $\{x_r^{(i)}\}$, $r = 1, \dots, n$ sono monotone e convergenti, e posto

$$\boldsymbol{\beta} = \lim_{i \rightarrow \infty} \mathbf{x}^{(i)},$$

poiché

$$\mathbf{f}(\mathbf{x}^{(i)}) = J(\mathbf{x}^{(i)})[\mathbf{x}^{(i)} - \mathbf{x}^{(i+1)}],$$

risulta $\mathbf{f}(\boldsymbol{\beta}) = \mathbf{0}$. Essendo anche $\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{0}$, dalla (63) si ha

$$\mathbf{0} = \mathbf{f}(\boldsymbol{\beta}) - \mathbf{f}(\boldsymbol{\alpha}) \geq J(\boldsymbol{\alpha})(\boldsymbol{\beta} - \boldsymbol{\alpha}),$$

$$\mathbf{0} = \mathbf{f}(\boldsymbol{\alpha}) - \mathbf{f}(\boldsymbol{\beta}) \geq J(\boldsymbol{\beta})(\boldsymbol{\alpha} - \boldsymbol{\beta}),$$

e poiché $[J(\boldsymbol{\alpha})]^{-1} \geq 0$ e $[J(\boldsymbol{\beta})]^{-1} \geq 0$, ne segue che

$$\mathbf{0} \geq \boldsymbol{\beta} - \boldsymbol{\alpha} \quad \text{e} \quad \mathbf{0} \geq \boldsymbol{\alpha} - \boldsymbol{\beta},$$

cioè $\boldsymbol{\alpha} = \boldsymbol{\beta}$. ■

Il teorema 3.47, nel caso particolare $n = 1$, assicura la convergenza del metodo delle tangenti nell'ipotesi che la funzione $f(x) \in C^1[a, b]$ sia convessa, $\alpha \in [a, b]$, e si scelga un punto x_0 tale che $f(x_0) > 0$. Se $f(x) \in C^2[a, b]$, allora la condizione di convessità è equivalente a $f''(x) > 0$, e le ipotesi del teorema 3.47 si riducono a quelle del teorema 3.25.

Teoremi analoghi al teorema 3.47 per la convergenza del metodo di Newton-Raphson possono essere dimostrati sotto le seguenti ipotesi:

- 1) $\mathbf{f}(\mathbf{x}) \in C^1(D)$ e convessa, $[J(\mathbf{x})]^{-1} \leq O$ per $\mathbf{x} \in D$. Si ha convergenza per ogni scelta di $\mathbf{x}^{(0)} \in D$ per cui $\mathbf{f}(\mathbf{x}^{(0)}) \geq \mathbf{0}$.
- 2) $\mathbf{f}(\mathbf{x}) \in C^1(D)$ e concava (cioè $-\mathbf{f}(\mathbf{x})$ convessa), $[J(\mathbf{x})]^{-1} \leq O$ oppure $[J(\mathbf{x})]^{-1} \geq O$ per $\mathbf{x} \in D$. Si ha convergenza per ogni scelta di $\mathbf{x}^{(0)} \in D$ per cui $\mathbf{f}(\mathbf{x}^{(0)}) \leq \mathbf{0}$.

3.48 Esempio. Il sistema non lineare

$$\begin{cases} f_1(x_1, x_2) = x_1^2 - \sin x_2 - 25 = 0 \\ f_2(x_1, x_2) = -\cos x_1 + x_2 = 0 \end{cases}$$

per $x_1 > 0$ ha una sola soluzione $\boldsymbol{\alpha}$ appartenente all'insieme $D = [4.6, 5.2] \times [0, 0.8]$, come risulta anche nella figura 3.21. Poiché

$$J(\mathbf{x}) = \begin{bmatrix} 2x_1 & -\cos x_2 \\ \sin x_1 & 1 \end{bmatrix},$$

è $\det(J(\mathbf{x})) = 2x_1 + \sin x_1 \cos x_2 \neq 0$, e quindi $J(\mathbf{x})$ è non singolare per $\mathbf{x} \in D$.

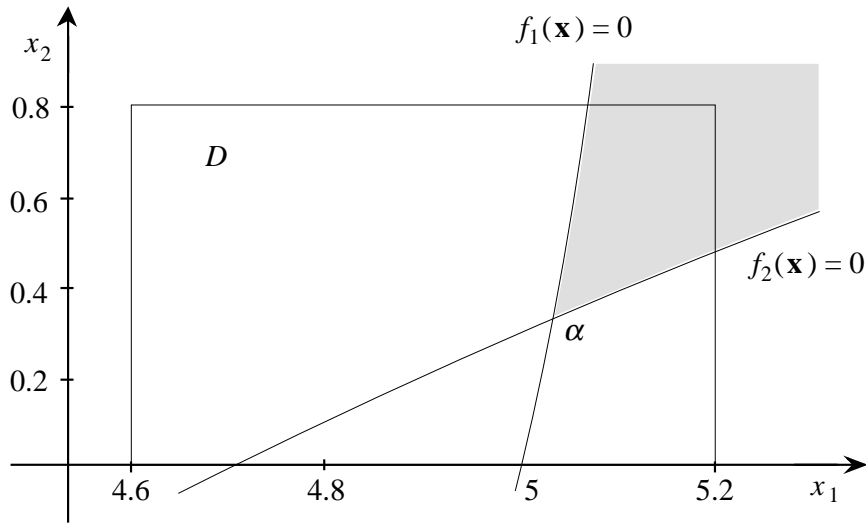


Fig. 3.21 - Grafici relativi al sistema dell'esempio 3.48.

Risulta

$$[J(\mathbf{x})]^{-1} = \frac{1}{\det(J(\mathbf{x}))} \begin{bmatrix} 1 & \cos x_2 \\ -\sin x_1 & 2x_1 \end{bmatrix} > O.$$

Inoltre le matrici hessiane sono

$$S_1 = \begin{bmatrix} 2 & 0 \\ 0 & \sin x_2 \end{bmatrix}, \quad S_2 = \begin{bmatrix} \cos x_1 & 0 \\ 0 & 0 \end{bmatrix}$$

e risultano semidefinite positive in D . Perciò in D la funzione $\mathbf{f}(\mathbf{x})$ è totalmente differenziabile e convessa e verifica quindi le ipotesi del teorema 3.47. Scegliendo come $\mathbf{x}^{(0)}$ un qualunque punto di D per cui $\mathbf{f}(\mathbf{x}^{(0)}) > \mathbf{0}$, si ottiene una successione monotona convergente ad α . Nell'insieme indicato in grigio nella figura è $\mathbf{f}(\mathbf{x}) > \mathbf{0}$. Scegliendo $\mathbf{x}^{(0)} = (5.2, 0.8)$, si ottiene la successione convergente

i	$x_1^{(i)}$	$x_2^{(i)}$
1	5.041220	0.3282420
2	5.030712	0.3129920
3	5.030693	0.3129568
4	5.030694	0.3129578

■