5 Нелинеарне једначине

5.1 Проста итерација

Посматрамо проблем нумеричког решавања једначине $f(x) = 0$ (када је немогуће решење одредити тачно), где је f непрекидна функција дефинисана на интервалу $[a, b]$.

Теорема 5.1. Нека је f непрекидна функција дефинисана на интервалу $[a, b]$. Ако је $f(a) \cdot f(b) < 0$, тада постоји $\xi \in [a, b]$, такво да је $f(\xi) = 0$.

Алтернативно, довољан услов за егзистенцију решења једначине $f(x) = 0$ можемо добити записивањем ове једначине у облику $g(x) = x$, за неку функцију g. На пример, једначину $f(x) = 0$ можемо записати у облику $g(x) = x$, при чему је $g(x) = f(x) + x$.

Теорема 5.2. (Теорема о фиксној тачки) Нека је g непрекидна функција на интервалу $[a, b]$ и нека је $g(x) \in [a, b]$ за $x \in [a, b]$ (тј. $g([a, b]) \subseteq [a, b]$). Тада постоји $\xi \in [a, b]$ за које важи $\xi = g(\xi)$; број ξ је *фиксна тачка* функције g.

Доказ. Нека је $f(x) = g(x) - x$. Тада је $f(a) = g(a) - a \geqslant 0$ и $f(b) = g(b) - b \leqslant 0$, па како је $f(a) \cdot f(b) < 0$, постоји $\xi \in [a, b]$ такво да је $0 = f(\xi) = g(\xi) - \xi$. □

Пример 5.1. Нека је $f(x) = e^x - 2x - 1$ за $x \in [1, 2]$. Како је $f(1) < 0$ и $f(2) > 0$, постоји $\xi \in [1, 2]$ тако да је $f(\xi) = 0$.

Ако једначину $f(x) = 0$ запишемо у облику $x = g(x) = \ln(2x + 1)$, видимо да је $g(x) \in [1, 2]$ за све $x \in [1, 2]$, па су задовољени услови теореме о фиксној тачки. Ипак, ако дату једначину напишемо у облику $x = g(x) = (e^x - 1)/2$, тада функција g не слика интервал $[1, 2]$ у себе, па се не може применити поменута теорема.

Дефинишимо сада метод *просте итерације* за нумеричко налажење фиксне тачке.

Дефиниција 5.1. (Проста итерација) Нека је g непрекидна функција дефинисана на интервалу $[a, b]$, при чему је $g(x) \in [a, b]$ за све $x \in [a, b]$. За дато $x_0 \in [a, b]$ рекурзија

$$x_{k+1} = g(x_k), \quad k = 0, 1, 2, \ldots, \tag{5.1}$$

се назива проста итерација.

Ако низ $\{x_k\}$ дефинисан у формули (5.1) конвергира, тада је лимес фиксна тачка функције g, јер је g непрекидна на затвореном интервалу. Заиста, ако је $\xi = \lim_{k\to\infty} x_k$, тада је

$$\xi = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} g(x_k) = g(\lim_{k\to\infty} x_k) = g(\xi).$$

Довољан услов за конвергенцију низа x_k је да функција g буде *контракција*.

Дефиниција 5.2. (Контракција) Нека је g непрекидна функција дефинисана на интервалу $[a, b]$. Кажемо да је g *контракција*, ако постоји константа L, $0 < L < 1$, таква да важи

$$|g(x) - g(y)| \leqslant L|x - y|, \quad \text{за све тачке } x, y \in [a, b].$$

Теорема 5.3. (Теорема о контракцији) Нека је g контракција на интервалу $[a, b]$, при чему је $g(x) \in [a, b]$ за $x \in [a, b]$. Тада постоји јединствена фиксна тачка $\xi \in [a, b]$, $\xi = g(\xi)$. Штавише, низ $x_{k+1} = g(x_k)$ конвергира ка тачки ξ за сваку почетну вредност $x_0 \in [a, b]$.

Претпоставимо још да је g диференцијабилна на (a, b). Тада за свако $x, y \in [a, b]$ на основу теореме о средњој вредности важи

$$|g(y) - g(x)| = |g'(\eta)||y - x|, \quad \text{за неко } \eta \in [x, y].$$

Приметимо да је довољан услов да функција буде контракција да је g диференцијабилна на (a, b) и да постоји константа $L \in (0, 1)$ за коју важи $|g'(x)| < L$ за $x \in (a, b)$. Иако је овај услов јачи од услова да функција буде контракција, у пракси је лакши за проверу.

Пример 5.2. Вратимо се на пример налажења решења једначине $f(x) = e^x - 2x - 1 = 0$ на сегменту $[1, 2]$, која се може написати у облику $x = g(x) = \ln(2x + 1)$, при чему је $g([1, 2]) \subseteq [1, 2]$.

Први извод $g'(x) = 2/(2x + 1)$ монотоно опада на $[1, 2]$, па је $g'(x) \in [2/5, 3/5]$ за $x \in [1, 2]$, одакле можемо закључити да важи $|g(y) - g(x)| \leqslant L|y - x|$, за $L = \frac{2}{3}$.

На основу теореме о контракцији, низ $x_{k+1} = \ln(2x + 1)$ конвергира ка ξ за сваку почетну вредност $x_0 \in [a, b]$. Ако изаберемо $x_0 = 1$ и радимо са 6 децималних места добијамо резултате које су приказани у следећој табели. Делује да је $\xi = 1.26$ до на два децимална места.

1	2	3	4	5	6	7	8	9	10	11
1.098612	1.162283	1.201339	1.224563	1.238121	1.245952	1.250447	1.253018	1.254486	1.255323	1.255800

Теорема 5.4. Нека је g контракција на интервалу $[a, b]$, при чему је $g(x) \in [a, b]$ за $x \in [a, b]$ и нека је ξ фиксна тачка функције g, при чему g има непрекидан извод у некој околини тачке ξ за који важи $|g'(\xi)| < 1$. Тада низ $x_{k+1} = g(x_k)$ конвергира ка тачки ξ за почетну вредност $x_0 \in [a, b]$ која је довољно близу тачки ξ.

Теорема 5.5. Посматрајмо просту итерацију (5.1), при чему функција g задовољава услове теореме о контракцији на $[a, b]$. За дате $x_0 \in [a, b]$ и $\varepsilon > 0$ нека је $k_0(\varepsilon)$ најмањи позитиван број за који x_k није више од ε удаљен од фиксне тачке ξ, тј. $|x_k - \xi| \leqslant \varepsilon$ за све $k \geqslant k_0$. Тада важи

$$k_0(\varepsilon) = \left[\frac{\ln|x_1 - x_0| - \ln|\varepsilon(1 - L)|}{\ln(1/L)} \right] + 1. \tag{5.2}$$

Доказ. Знамо да важи

$$|g(x_{k-1}) - g(\xi)| = |x_k - \xi| \leqslant L|x_{k-1} - \xi|, \quad \text{па је} \quad |x_k - \xi| \leqslant L^k|x_0 - \xi|. \tag{5.3}$$

Претходни резултат за $k = 1$ нам даје

$$\begin{aligned} |x_0 - \xi| &= |x_0 - x_1 + x_1 - \xi| \\ &\leqslant |x_0 - x_1| + |x_1 - \xi| \\ &\leqslant |x_0 - x_1| + L|x_0 - \xi|, \end{aligned}$$

одакле је $|x_0 - \xi| \leqslant \frac{1}{1-L}|x_0 - x_1|$, чијом заменом у (5.3) добијамо

$$|x_k - \xi| \leqslant \frac{L^k}{1 - L}|x_1 - x_0|,$$

па из услова $|x_k - \xi| \leqslant \varepsilon$ добијамо

$$\frac{L^k}{1 - L}|x_1 - x_0| \leqslant \varepsilon, \quad \text{одакле је} \quad k \geqslant \frac{\ln|x_1 - x_0| - \ln|\varepsilon(1 - L)|}{\ln(1/L)}.$$

□

Пример 5.3. Посматрајмо итеративни процес $x_{k+1} = 0.5 + \cos(x_k)$, $x_0 = 1$. Оценити број итерација за достизање формата двоструке прецизности ($\varepsilon = 10^{-16}$), ако је тачно решење, до на 6 децималних места $x = 1.021817$.

Овде је $x_0 = 1$, $x_1 = 0.5 + \cos(1) = 1.040302$, а функција $g(x) = 0.5 + \cos x$ је контракција, при чему је $L = |g'(1.021817)| = |-\sin(1.021817)| = 0.853058$. Применом формуле (5.2) добијамо $k = [223.672] + 1 = 224$.

Може се показати и да важи

$$|x_k - \xi| \leqslant \frac{L}{1 - L}|x_k - x_{k-1}|.$$

Ако је $0 \leqslant L \leqslant \frac{1}{2}$, онда из претходне неједнакости следи да важи *критеријум поклапања двеју узастопних итерација*, тј.

$$|x_k - x_{k-1}| \leqslant \varepsilon \quad \Leftrightarrow \quad |x_k - \xi| \leqslant \varepsilon.$$

У општем случају овај критеријум не важи и то се најбоље види у примерима где је $|g'(x)|$ блиско јединици: величина $|x_k - x_{k-1}|$ може бити доста мања од величине $|x_k - \xi|$.

Из услова

$$|x_k - \xi| \leqslant \frac{L}{1-L}|x_k - x_{k-1}| \leqslant \varepsilon$$

долазимо до *критеријума заустављања*

$$|x_k - x_{k-1}| \leqslant \frac{1-L}{L}\varepsilon.$$

Пример 5.4. Методом итерације са тачношћу $\varepsilon = 10^{-4}$ израчунати приближно решење једначине $f(x) = 5x^3 - 20x + 3 = 0$ које припада одсечку $[0, 1]$.

Запишимо дату једначину у облику $x = g(x) = \frac{1}{20}(5x^3 + 3)$. Функција g слика интервал $[0, 1]$ у себе и испуњен је довољан услов да итеративни процес конвергира:

$$\max_{x\in[0,1]} |g'(x)| = \frac{3}{4} < 1.$$

Ипак, како је $L = \frac{3}{4} > \frac{1}{2}$ не важи критеријум поклапања двеју узастопних итерација. Критеријум заустављања је у овом случају

$$|x_k - x_{k-1}| \leqslant \frac{1-3/4}{3/4} \cdot 10^{-4} = 0.000033\ldots$$

За почетну вредност се може узети било која вредност из $[0, 1]$, па бирамо нпр. $x_0 = 0.5$. Рачунамо редом

$$\begin{aligned}
x_1 =& g(0.5000) = 0.18125,\\
x_2 =& g(0.18125) = 0.15149,\\
x_3 =& g(0.15149) = 0.15087,\\
x_4 =& g(0.15087) = 0.15086.
\end{aligned}$$

Видимо да је у последњој итерацији испуњен критеријум заустављања $|x_4 - x_3| = 0.00001 < 0.00003$, па је тражено приближно решење $x = 0.1509$.

Дефиниција 5.3. (Ред конвергенције низа) Ред конвергенције низа x_n који тежи ка α када $n \to \infty$ је p, ако за $L \geqslant 0$ важи

$$|x_n - \alpha| = L|x_{n-1} - \alpha|^p.$$

Нека је $p = 1$. Тада мора бити $L < 1$ и важи $|x_n - \alpha| = L^n|x_0 - \alpha|$. У овом случају кажемо да низ итерација конвергира линеарно са фактором L.

Дакле, под условом $|g'(\xi)| < 1$ низ простих итерација $x_{k+1} = g(x_k)$ ће конвергирати ка фиксној тачки ξ линеарно.

5.2 Њутнова метода тангенте

Дефиниција 5.4. (Њутнова метода) Њутнова метода тангенте за решење једначине $f(x) = 0$ је

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \dots, \tag{5.4}$$

за одговарајућу почетну вредност x_0, при чему је $f'(x_k) \neq 0$ за $k \geqslant 0$.

Њутнова метода је проста итерација за функцију $g(x) = x - \frac{f(x)}{f'(x)}$ и геометријски тангента на криву $f(x)$ у тачки $(x_k, f(x_k))$ сече x-осу у тачки x_{k+1}. До ње можемо доћи и развијањем функције f у Тејлоров ред првог реда у околини тачке x_k са грешком која се изражава преко другог извода:

$$f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{f''(\eta_k)}{2}(x - x_k)^2, \quad \text{за неко } \eta_k \text{ између } x \text{ и } x_k.$$

Ако је α решење једначине $f(x) = 0$, заменом x са α у претходној једначини добијамо

$$0 = f(\alpha) = f(x_k) + f'(x_k)(\alpha - x_k) + \frac{f''(\eta_k)}{2}(\alpha - x_k)^2, \quad \text{одакле је} \quad \alpha = x_k - \frac{f(x_k)}{f'(x_k)} - (\alpha - x_k)^2 \frac{f''(\eta_k)}{2f'(x_k)}.$$

Ако је x_{k+1} дефинисано формулом (5.6), добијамо

$$\alpha - x_{k+1} = -(\alpha - x_k)^2 \frac{f''(\eta_k)}{2f'(x_k)},$$

па ако конвергира, Њутнов метод ће конвергирати са редом два.

Теорема 5.6. Нека је $f(\alpha) = 0$, при чему је f два пута непрекидно диференцијабилна у некој околини тачке α и $f'(\alpha) \neq 0$. Ако је почетна итерација x_0 довољно близу α, тада низ x_k конвергира ка α.

У пракси, можемо користити наредну теорему као довољне услове за конвергенцију Њутновог метода.

Теорема 5.7. Нека је f два пута непрекидно диференцијабилна на сегменту $[a, b]$, при чему је
- $f(a) \cdot f(b) < 0$ и
- f' и f'' су сталног знака на $[a, b]$.

Ако за почетну итерацију x_0 важи
- $f(x_0) \cdot f''(x_0) > 0$,

тада низ x_k добијен Њутновом методом конвергира ка (простој) нули функције f.

Нека је $|f'(x)| \geqslant m_1 > 0$ и $|f''(x)| \leqslant M_1$ за $x \in [a, b]$ (тј. $m_1 = \min|f'(x)|$ и $M_2 = \max|f''(x)|$ за $x \in [a, b]$). Да би смо извели оцену грешке, приметимо да за две сустедне итерације важи

$$f(x_k) = f(x_{k-1}) + f'(x_{k-1})(x_k - x_{k-1}) + \frac{f''(\eta_k)}{2!}(x_k - x_{k-1})^2, \quad \eta_k \in [x_{k-1}, x_k].$$

Из дефиниције Њутнове методе је $f(x_{k-1}) + f'(x_{k-1})(x_k - x_{k-1}) = 0$, па је

$$f(x_k) = \frac{f''(\eta_k)}{2!}(x_k - x_{k-1})^2.$$

Ако приметимо још да важи

$$|\alpha - x_k| \leqslant \frac{|f(x_k)|}{m_1},$$

добијамо *оцену грешке Њутнове методе*

$$|\alpha - x_k| \leqslant \frac{M_2}{2m_1}(x_k - x_{k-1})^2.$$

Ако је циљ добити решење са тачношћу ε, процес рачунања итерација се зауставља када буде испуњен услов (*критеријум заустављања*)

$$|x_k - x_{k-1}| \leqslant \sqrt{\frac{2m_1\varepsilon}{M_2}}. \tag{5.5}$$

Поменимо да је модификовани Њутнов метод дефинисан итерацијама

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_0)}, \quad k = 0, 1, 2, \ldots,$$

чији је циљ да се смањи број операција у једном кораку (извод се рачуна само у почетној тачки). Очекивано, смањења је брзине конвергенције - овде је један.

На крају, ако приметимо да је

$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}},$$

заменом $f'(x_k)$ у (5.6) добијамо *метод сечице.*

Дефиниција 5.5. (Метода сечице) Метода сечице за решење једначине $f(x) = 0$ је

$$x_{k+1} = x_k - f(x_k)\left(\frac{x_k - x_{x-1}}{f(x_k) - f(x_{k-1})}\right), \quad k = 1, 2, \ldots, \tag{5.6}$$

за одговарајуће почетне вредности x_0 и x_1, при чему је $f(x_k) - f(x_{k-1}) \neq 0$ за $k \geqslant 1$.

Пример 5.5. Њутновом методом наћи позитивно решење једначине $x^3 = \sin x$ са релативном грешком мањом од 10^{-5}.

Цртањем графика функција $y_1 = x^3$ и $y_2 = \sin x$, видимо да ће позитивна нула функције $f(x) = x^3 - \sin x$ бити у нпр. интервалу $[0.7, 1]$ (проверавамо да ли је $f(a) \cdot f(b) < 0$ и сужавамо интервал нпр. методом половљења - што више сузимо интервал имаћемо бржу конвергенцију). Ако изаберемо тачку $x_0 = 1$ проверавамо да ли су задовољени услови теореме (5.7):

- $f(0.7) \cdot f(1) < 0$,
- $f'(x) = 3x^2 - \cos x \neq 0$ за $x \in [0.7, 1]$,
- $f''(x) = 6x + \sin x > 0$ за $x \in [0.7, 1]$,
- $f(1) \cdot f''(1) = 1.08457 > 0$.

Ако је релативна грешка мања од 10^{-5}, онда је апсолутна грешка на интервалу $[0.7, 1]$ мања од $0.7 \cdot 10^{-5}$, па можемо узети да је $\varepsilon = 7 \cdot 10^{-6}$. Да бисмо применили критеријум заустављања (5.5), нађимо минимум првог и маскимум другог извода функције f на интервалу $[0.7, 1]$:

$$|f'(x)| = |3x^2 - \cos x| \geqslant 3 \cdot 0.49 - 1 = 0.47 = m_1, \quad |f''(x)| = |6x + \sin x| \leqslant 7 = M_2 \quad \text{на интервалу } [0.7, 1],$$

па ће на основу формуле (5.5) тачност бити постигнута када суседне итерације буду на растојању мањем од $\sqrt{\frac{2 \cdot 0.47 \cdot 7 \cdot 10^{-6}}{7}} < 9.7 \cdot 10^{-4}$. За то су довољне 3 итерације:

$$x_1 = 0.935549, \quad x_2 = 0.928702, \quad x_3 = 0.928626,$$

па је тражено решење $x = 0.928626$.

5.3 Њутнова метода за системе нелинеалних једначина

Њутнову методу могуће је уопштити за решавања система нелинеарних једначина

$$f_i(x_1, \dots, x_n) = 0, \quad 1 \leqslant i \leqslant n, \quad \text{тј.} \quad F(x) = 0, \; F = (f_1, \dots, f_n) : \mathbb{R}^n \to \mathbb{R}^n.$$

Јакобијева матрица је

$$J = \left[\frac{\partial f_i}{\partial x_j} \right]_{ij}.$$

Њутнов метод је у овом случају

$$x^{(k+1)} = x^{(k)} - [J(x^{(k)})]^{-1} F(x^{(k)}), \quad k = 0, 1, \dots \tag{5.7}$$

Теорема 5.8. Нека је $F(\alpha) = 0$, $F \in C^2$ у некој околини тачке α и нека је $\det J(\alpha) \neq 0$. Ако је почетна итерација $x^{(0)}$ довољно близу α, низ $x^{(k)}$ ће конвергирати квадратно ка α.

Пример 5.6. Са тачношћу $\frac{1}{2} \cdot 10^{-3}$ израчунати једно решење система једначина

$$\begin{aligned} f_1(x, y) =& 2x^2 - xy - 5x + 1 = 0 \\ f_2(x, y) =& x + 3 \log x - y^2 = 0, \end{aligned}$$

ако је почтна итерација $x^{(0)} = [3.5 \; 2.2]^T$.

Нађимо прво Јакобијеву матрицу пресликавања F

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} = \begin{bmatrix} 4x - y - 5 & -x \\ 1 + \frac{3}{x \ln 10} & -2y \end{bmatrix}.$$

Према формули (5.7), при чему је $F = [f_1 \; f_2]^T$, у свакој итерацији $x^{(k)}$ рачунамо $F(x^{(k-1)})$ и $J^{-1}(x^{(k-1)})$.

Дакле

$$F(x^{(0)}) = [0.3000 \; 0.2922]^T, \quad J(x^{(0)}) = \begin{bmatrix} 6.8000 & -3.5000 \\ 1.3723 & -4.4000 \end{bmatrix}, \quad J^{-1}(x^{(0)}) = -\frac{1}{25.11695} \begin{bmatrix} -4.4000 & 3.5000 \\ -1.3723 & 6.8000 \end{bmatrix},$$

па је

$$x^{(1)} = \begin{bmatrix} 3.5000 \\ 2.2000 \end{bmatrix} + \frac{1}{25.11695} \begin{bmatrix} -4.4000 & 3.5000 \\ -1.3723 & 6.8000 \end{bmatrix} \cdot \begin{bmatrix} 0.3000 \\ 0.2922 \end{bmatrix} = \begin{bmatrix} 3.4882 \\ 2.2627 \end{bmatrix}.$$

Да би се постигла задата тачност, довољно је урадити још две итарације

$$x^{(2)} = \begin{bmatrix} 3.4874 \\ 2.2616 \end{bmatrix}, \quad x^{(3)} = \begin{bmatrix} 3.4874 \\ 2.2616 \end{bmatrix},$$

па је тражено решење $(3.487, 2.262)$.