# Uvod u numeričku matematiku

M. Klaričić Bakula

Ožujak, 2009.

# 1 Općenito o iterativnim metodama za rješavanje nelinearnih jednadžbi

Računanje nultočaka nelinearnih funkcija jedan je od najčešćih zadataka primijenjene matematike.
Općenito, neka je zadana funkcija

$$f : I \to \mathbb{R}, \quad I \subseteq \mathbb{R}.$$

Tražimo sve one $x \in I$ za koje vrijedi

$$f(x) = 0.$$

Takve točke $x$ zovu se **rješenja** (korijeni) pripadne jednadžbe ili **nultočke** funkcije $f$.

U pravilu pretpostavljamo da je funkcija $f$ neprekidna na intervalu $I$ i da su joj nultočke izolirane. Naime, u protivnom bi postojao problem konvergencije.

Traženje nultočaka na zadanu točnost sastoji se od dvije faze:

- **izolacije jedne nultočke ili više njih**, tj. nalaženje intervala unutar kojeg se nalazi barem jedna nultočka funkcije
- **iterativno nalaženje nultočke** do na traženu točnost.

Prva faza traženja je teža i provodi se na temelju analize toka funkcije $f$.

Postoji mnoštvo metoda za nalaženje nultočaka nelinearnih funkcija do na zadanu točnost, no razlikuju se po tomu hoće li uvijek konvergirati i (ako konvergiraju) po brzini konvergencije. Općenito, brže metode nemaju sigurnu konvergenciju, dok je sporije metode imaju. Brzina konvergencije se opisuje pomoću reda konvergencije metode.

**DEFINICIJA**. Niz iteracija $(x_n)_{n\in\mathbb{N}_0}$ **konvergira** prema točki $\alpha$ s redom konvergencije $p,\ p \geq 1$, ako vrijedi

$$|\alpha - x_n| \leq c\,|\alpha - x_{n-1}|^p\,, \quad n \in \mathbb{N} \tag{1}$$

za neki $c > 0$. Ako je $p = 1$, onda kažemo da niz **konvergira linearno** prema $\alpha$. U tom slučaju nužno mora vrijediti $c < 1$ i obično se $c$ naziva **faktorom linearne konvergencije**.

Relacija $(1)$ nije uvijek prikladna za linearne iterativne algoritme, no indukcijom se može pokazati da je za slučaj $p = 1$ i $c < 1$ nejednakost $(1)$ ekvivalentna s

$$|\alpha - x_n| \leq c^n\,|\alpha - x_0|\,, \quad n \in \mathbb{N},$$

što je katkad mnogo lakše dokazati nego $(1)$. I u ovom slučaju govorimo o linearnoj konvergenciji s faktorom $c$.

# 2 Metoda polovljenja (bisekcije)

Najjednostavnija metoda nalaženja nultočaka nelinearne funkcije je **metoda polovljenja**. Ona funkcionira za sve neprekidne funkcije, no zato ima i najlošiju **ocjenu pogreške**.
Osnovna pretpostavka za primjenu algoritma polovljenja je neprekidnost funkcije $f$ na intervalu $[a,b]$ uz dodatni uvjet

$$f(a)\,f(b) < 0.$$

Ovaj uvjet nam osigurava da funkcija $f$ ima na intervalu $[a,b]$ barem jednu nultočku. Naravno, ako je

$$f(a)\,f(b) > 0,$$

to ne znači da $f$ nema nultočaka na $[a,b]$ : može ih imati paran broj ili nultočku parnog reda. U prvom slučaju možemo riješiti problem boljom separacijom intervala $[a,b]$ , no u drugom slučaju metoda polovljenja neće dati rješenje.

Algoritam polovljenja je vrlo jednostavan: označimo s $\alpha$ prvu nultočku funkcije $f$ i definiramo

$$a_0 := a, \quad b_0 := b, \quad x_0 = \frac{a_0 + b_0}{2}.$$

Neka je $n \geq 1.$ U $n$-tom koraku algoritma konstruiramo interval $[a_n, b_n]$ komu je duljina polovina duljine prethodnog intervala $[a_{n-1}, b_{n-1}]$ , ali tako da nultočka ostane unutar intervala $[a_n, b_n]$ . Konstrukcija intervala $[a_n, b_n]$ sastoji se u raspolavljanju intervala $[a_{n-1}, b_{n-1}]$ točkom $x_{n-1}$ i to tako da je

$$\begin{aligned} a_n &:= x_{n-1}, \quad b_n := b_{n-1}, \quad f(a_{n-1}) f(x_{n-1}) > 0, \\ a_n &:= a_{n-1}, \quad b_n := x_{n-1}, \quad f(a_{n-1}) f(x_{n-1}) < 0. \end{aligned}$$

Postupak zaustavljamo kad je ispunjen uvjet

$$|\alpha - x_n| \leq \varepsilon.$$

No kako ćemo znati da to vrijedi ako ne znamo $\alpha$? Jer je $x_n$ polovište intervala $[a_n, b_n]$ i $\alpha \in [a_n, b_n]$, to je

$$|\alpha - x_n| \leq |b_n - x_n| = b_n - x_n,$$

pa je u stvari dovoljno postaviti zahtjev

$$b_n - x_n \leq \varepsilon.$$

**ALGORITAM** (*Metoda polovljenja*)
$x := (a+b)/2$
**while** $b - x > \varepsilon$ **do**
    **begin**;
    **if** $f(x) * f(b) < 0.0$ **then**
        $a := x;$
    **else**
        $b := x;$
    $x := (a+b)/2;$
    **end**;

Iz konstrukcije same metode lako se ocijeni pogreška $n$-te aproksimacije nultočke $\alpha$. Vrijedi

$$|\alpha - x_n| \leq b_n - x_n = \frac{1}{2}(b_n - a_n) = \frac{1}{2^2}(b_{n-1} - a_{n-1}) = \cdots = \frac{1}{2^{n+1}}(b-a), \qquad (2)$$

odnosno

$$|\alpha - x_n| \leq \frac{1}{2^n}(b - x_0).$$

Desna strana ove nejednakosti daje nam naslutit da će konvergencija biti dosta spora. Relacija $(2)$ nam omogućava da unaprijed odredimo koliko će koraka biti potrebno da bismo postigli željenu točnost $\varepsilon$. Naime, da bi vrijedilo $|\alpha - x_n| \leq \varepsilon$ dovoljno je zahtijevati

$$\frac{1}{2^{n+1}}(b-a) \leq \varepsilon.$$

Jednostavnim računom dobijemo

$$n \geq \frac{\log(b-a) - \log\varepsilon}{\log 2} - 1, \quad n \in \mathbb{N}_0.$$

Ako je funkcija $f$ još i neprekidno derivabilna na $[a, b]$, možemo dobiti dinamičku ocjenu udaljenosti aproksimacije nultočke od prave nultočke. Po *Teoremu srednje vrijednosti* za funkciju $f$ vrijedi

$$f(x_n) = f(\alpha) + f'(\xi)(x_n - \alpha),$$

pri čemu je $\xi$ između $\alpha$ i $x_n$. Koristeći činjenicu da je $f(\alpha) = 0$ dobijemo

$$|f(x_n)| = |f'(\xi)|\,|x_n - \alpha|,$$

odakle slijedi

$$|x_n - \alpha| = \frac{|f(x_n)|}{|f'(\xi)|}.$$

Pretpostavimo li da možemo dati ocjenu

$$|f'(\xi)| \geq m_1, \quad m_1 = \min_{x \in [a,b]} |f'(x)|,$$

ako je $m_1 \neq 0$ dobijemo

$$|x_n - \alpha| \leq \frac{|f(x_n)|}{m_1}.$$

Dakle, želimo li postići da vrijedi

$$|\alpha - x_n| \leq \varepsilon,$$

dovoljno je zahtijevati

$$|f(x_n)| \leq m_1 \varepsilon.$$

# 3 Metoda pogrešnog položaja (*Regula falsi*)

Metoda koja nastaje kao prirodna posljedica ubrzavanja metode polovljenja je tzv. **regula falsi**. I ona sama je konvergentna čim se barem jedna nultočka funkcije $f$ nalazi unutar intervala $[a, b]$ .

Pretpostavimo opet da je funkcija $f$ neprekidna na intervalu $[a, b]$ i da vrijedi

$$f(a) f(b) < 0.$$

Aproksimiramo funkciju $f$ pravcem $p$ koji prolazi točkama $T_1(a, f(a))$ i $T_2(b, f(b))$ . Njegova je jednadžba

$$y - f(b) = \frac{f(a) - f(b)}{a - b}(x - b),$$

odnosno

$$y - f(a) = \frac{f(b) - f(a)}{b - a}(x - a).$$

Nultočku $\alpha$ funkcije $f$ možemo aproksimirati nultočkom tog pravca, nazovimo je $x_0$. Nakon toga pomaknemo ili točku $a$ ili točku $b$ u $x_0$, no tako da nultočka $\alpha$ ostane unutar novodobivenog intervala. Postupak ponavljamo sve dok ne dobijemo željenu točnost $\varepsilon$. Točka $x_0$ se dobije jednostavno iz jednadžbe pravca $p$ kao

$$x_0 = b - f\left(b\right) \frac{b-a}{f\left(b\right) - f\left(a\right)} = a - f\left(a\right) \frac{a-b}{f\left(a\right) - f\left(b\right)},$$

ili drugačije zapisano

$$x_0 = b - \frac{f\left(b\right)}{f\left[a,b\right]} = a - \frac{f\left(a\right)}{f\left[a,b\right]}, \tag{3}$$

gdje je

$$f\left[a,b\right] = \frac{f\left(b\right) - f\left(a\right)}{b-a}$$

prva podijeljena razlika funkcije $f$ u točkama $a$ i $b$.

Ipak, postoji nekoliko ozbiljnih problema s ovom metodom. Pogledajmo kakav je red konvergencije. Iz relacije $(3)$ imamo

$$\begin{aligned}
\alpha - x_0 &= \alpha - b + \frac{f(b)}{f[a,b]} = (\alpha - b)\left[1 + \frac{f(b)}{(\alpha - b) f[a,b]}\right] \\
&= (\alpha - b)\left[1 + \frac{f(b) - f(\alpha)}{(\alpha - b) f[a,b]}\right] \\
&= (\alpha - b)\left[1 + (b - \alpha)\frac{f[b,\alpha]}{(\alpha - b) f[a,b]}\right] \\
&= (\alpha - b)\left[1 - \frac{f[b,\alpha]}{f[a,b]}\right] = (\alpha - b)\frac{f[a,b] - f[b,\alpha]}{f[a,b]} \\
&= -(\alpha - b)(\alpha - a)\frac{f[a,b,\alpha]}{f[a,b]}.
\end{aligned}$$

Podsjetimo se da je

$$f[a,b,\alpha] = \frac{f[b,\alpha] - f[a,b]}{\alpha - a}.$$

Ako je funkcija $f$ neprekidno derivabilna, onda po Teoremu srednje vrijednosti imamo

$$f[a,b] = \frac{f(b) - f(a)}{b - a} = f'(\xi), \quad \xi \in [a,b].$$

Slično, ako je funkcija $f$ još i dvaput neprekidno derivabilna, onda po Teoremu srednje vrijednosti imamo

$$f[a,b,\alpha] = \frac{f[b,\alpha] - f[a,b]}{\alpha - a} = \frac{1}{2} f''(\eta), \quad \eta \in [m,M],$$

pri čemu je

$$m = \min\{a,b,\alpha\}, \quad M = \max\{a,b,\alpha\}.$$

Iskoristimo li ovo dobijemo ocjenu

$$\alpha - x_0 = -(\alpha - b)(\alpha - a)\frac{f''(\eta)}{2f'(\xi)}.$$

Sada ćemo na jednom slučaju pokazati kako se provodi analiza konvergencije ove metode.

Pretpostavimo da je $\alpha$ jedini korijen funkcije $f$ unutar $[a, b]$ i da vrijedi $f'(\alpha) \neq 0$. Također pretpostavimo da je $f''(x) \geq 0$ za sve $x \in [a, b]$ (tj. da je $f$ konveksna na $[a, b]$). Ako je $f'(x) > 0$ za sve $x \in [a, b]$, onda je $f$ konveksna rastuća funkcija, a spojnica točaka $T_1(a, f(a))$ i $T_2(b, f(b))$ je uvijek iznad grafa funkcije $f$. Iz početnih uvjeta dobijemo da je

$$\alpha - x_0 = -(\alpha - b)(\alpha - a)\frac{f''(\eta)}{2f'(\xi)} > 0,$$

pa će se u sljedećem koraku pomaknuti $a$. Isto će se dogoditi i u svim narednim koracima. Dakle, $b$ je fiksan, a $\alpha$ je stalno desno od aproksimacije $x_n$. To znači da vrijedi

$$\alpha - x_n = -(\alpha - b)(\alpha - a_n)\frac{f''(\eta_n)}{2f'(\xi_n)}.$$

Uzimanjem apsolutnih vrijednosti s desna i lijeva dobijemo da je konvergencija metode *regula falsi* u ovom slučaju linearna. Lako se vidi da je moguće naći primjere kod kojih metoda polovljenja konvergira brže nego metoda *regula falsi*.

# 4 Metoda sekante

Ako, slično kao kod *regule falsi*, graf funkcije $f$ aproksimiramo sekantom, ali pri tom ne zahtijevamo da nultočka funkcije $f$ ostane "zatvorena" unutar poljednje dvije iteracije, dobit ćemo **metodu sekante**. Time smo izgubili svojstvo sigurne konvergencije, ali se nadamo da će metoda konvergirati brže nego *regula falsi*.

U ovoj metodi počinjemo s dvije početne točke $x_0$ i $x_1$, te povlačimo sekantu kroz točke $T_0\left(x_0, f\left(x_0\right)\right)$ i $T_1\left(x_1, f\left(x_1\right)\right)$. Ta sekanta siječe os $x$ u točki $x_2$. Postupak nastavljamo provlačenjem sekante kroz točke $T_1\left(x_1, f\left(x_1\right)\right)$ i $T_2\left(x_2, f\left(x_2\right)\right)$. Formule za metodu sekante dobiju se iteriranjem početne formule za *regulu falsi*, pa tako dobivamo

$$x_{n+1} = x_n - f\left(x_n\right)\frac{x_n - x_{n-1}}{f\left(x_n\right) - f\left(x_{n-1}\right)}.$$

Iskoristimo li za svaki $n \in \mathbb{N}_0$ ocjenu

$$\alpha - x_n = -\left(\alpha - b_n\right)\left(\alpha - a_n\right)\frac{f''\left(\eta_n\right)}{2f'\left(\xi_n\right)},$$

dobit ćemo red konvergencije metode sekante uz odgovarajuće pretpostavke. Vrijedi

$$\alpha - x_{n+1} = -\left(\alpha - x_n\right)\left(\alpha - x_{n-1}\right)\frac{f''\left(\eta_n\right)}{2f'\left(\xi_n\right)}. \tag{4}$$

**TEOREM**. Neka su $f, f'$ i $f''$ neprekidne na nekom intervalu $I \subset \mathbb{R}$, pri čemu $I$ sadrži jednostruku nultočku $\alpha$. Ako su početne iteracije $x_0$ i $x_1$ izabrane dovoljno blizu nultočke $\alpha$, onda niz iteracija $\left(x_n\right)_{n\in\mathbb{N}_0}$ dobiven metodom sekante konvergira prema $\alpha$ s redom konvergencije $p$, gdje je

$$p = \frac{1+\sqrt{5}}{2} \approx 1.618.$$

**DOKAZ**. Primijetimo da jednostrukost nultočke $\alpha$ osigurava ispunjenje uvjeta $f'(\alpha) \neq 0$. Također, za neki $\varepsilon > 0$ postoji okolina $O = [\alpha - \varepsilon, \alpha + \varepsilon] \subset I$ nultočke $\alpha$, takva da je $f'(x) \neq 0$ za sve $x \in O$. U tom slučaju je dobro definiran broj

$$M = \frac{\max_{x \in O} |f''(x)|}{2 \min_{x \in O} |f'(x)|}.$$

Zbog relacije $(4)$ za sve $x_0, x_1 \in O$ vrijedi

$$|\alpha - x_2| \leq M |\alpha - x_1| |\alpha - x_0| .$$

Da bismo skratili zapis označimo s

$$e_n = \alpha - x_n.$$

grešku $n$-te iteracije aproksimacije nultočke $\alpha$. Sada prethodnu nejednakost možemo nakon množenja s $M$ pisati kao

$$M |e_2| \leq M^2 |e_1| |e_0| .$$

Pretpostavimo da su $x_0, x_1 \in O$ izabrani toliko blizu nultočke $\alpha$ da vrijedi

$$\delta = \max\{M\,|e_1|\,, M\,|e_0|\} < 1,$$

iz čega odmah slijedi

$$M\,|e_2| \le \delta^2 < \delta,$$

pa je

$$|e_2| < \frac{\delta}{M} = \max\{|e_1|\,, |e_0|\} \le \varepsilon,$$

odnosno

$$x_2 \in [\alpha - \varepsilon, \alpha + \varepsilon] = O.$$

Primjenimo li ovaj argument induktivno, dobijemo

$$\begin{aligned} M\,|e_3| &\le M^2\,|e_2|\,|e_1| \le \delta^2 \cdot \delta = \delta^3, \\ M\,|e_4| &\le M^2\,|e_3|\,|e_2| \le \delta^3 \cdot \delta^2 = \delta^5, \end{aligned}$$

i općenito

$$M\,|e_{n+1}| \le M^2\,|e_n|\,|e_{n-1}| \le \delta^{q_n} \cdot \delta^{q_{n-1}} = \delta^{q_{n+1}},$$

pri čemu je

$$\begin{aligned} q_{n+1} &= q_n + q_{n-1}, \quad n \ge 1, \\ q_0 &= q_1 = 1. \end{aligned}$$

Dakle, vidimo da se radi o rekurziji za Fibonaccijeve brojeve, pa se lako izračuna eksplicitno rješenje

$$q_n = c_0 r_0^n + c_1 r_1^n.$$

Pri tom je

$$\begin{aligned} r_0 &= \frac{1+\sqrt{5}}{2}, \quad r_1 = \frac{1-\sqrt{5}}{2}, \\ c_0 &= \frac{1}{\sqrt{5}} r_0, \quad c_1 = -\frac{1}{\sqrt{5}} r_1. \end{aligned}$$

Dakle, dobili smo

$$q_n = \frac{1}{\sqrt{5}} \left[ \left( \frac{1+\sqrt{5}}{2} \right)^{n+1} - \left( \frac{1-\sqrt{5}}{2} \right)^{n+1} \right], \quad n \geq 0.$$

Kako je $r_0 \approx 1.618$ i $r_1 \approx -0.618$, to vidimo da $r_1$ teži k nuli kada $n$ teži prema beskonačnosti, pa je za velike $n$

$$q_n \approx \frac{1}{\sqrt{5}} \, (1.618)^{n+1}.$$

No vratimo se na grešku $e_n$. Vidjeli smo da je

$$M\left|e_n\right| \leq \delta^{q_n}, \quad n \geq 0,$$

pa budući da je $0 < \delta < 1$ i za velike $n$ broj $q_n$ teži k beskonačnosti, to vrijedi $|e_n| \to 0$, tj.

$$x_n \to \alpha, \quad n \to \infty \qquad \blacksquare$$

Napominjemo da je ovaj dokaz bitno pojednostavljen i nije posve korektan!

Mane ove metode su da ona bitno ovisi o dobrom odabiru početnih aproksimacija, te da se lako može javiti poznati problem "*kraćenja*" u brojniku i (posebno) nazivniku kvocijenta

$$\frac{x_n - x_{n-1}}{f\left(x_n\right) - f\left(x_{n-1}\right)}$$

kada $x_n \to \alpha$. Također, budući da iteracije s obje strane ne zatvaraju nultočku $\alpha$ nije lako reći kada treba zaustaviti proces.

# 5 Metoda tangente (Newtonova metoda)

Ako graf funkcije $f$ aproksimiramo tangentom umjesto sekantom, dobili smo **metodu tangente** ili **Newtonovu metodu**. Slično kao i kod metode sekante time smo izgubili sigurnu konvergenciju, no nadamo se da će metoda brzo konvergirati.

Pretpostavimo da je zadana početna točka $x_0$. Ideja metode je povući tangentu u točki $T_0\left(x_0, f\left(x_0\right)\right)$ i definirati novu aproksimaciju $x_1$ u točki gdje tangenta siječe os $x$. Općenito bi to išlo ovako: u točki $x_n$ napiše se jednadžba tangente

$$y - f\left(x_n\right) = f'\left(x_n\right)\left(x - x_n\right).$$

Nultočka joj je

$$x = x_n - \frac{f\left(x_n\right)}{f'\left(x_n\right)},$$

pa stavimo $x_{n+1} := x.$

Primijetimo da je ova metoda usko vezana uz metodu sekante jer je

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

Do Newtonove metode možemo doći i korištenjem razvoja u Taylorov red funkcije $f$ oko točke $x_n$ uz pretpostavku da je $f$ dvaput neprekidno derivabilna na u nekoj okolini nultočke $\alpha$. Vrijedi

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{f''(\xi_n)}{2}(x - x_n)^2,$$

pri čemu je $\xi_n$ između $x$ i $x_n$. Uvrštavanjem $x = \alpha$ dobivamo

$$0 = f(\alpha) = f(x_n) + f'(x_n)(\alpha - x_n) + \frac{f''(\xi_n)}{2}(\alpha - x_n)^2.$$

Uz pretpostavku da je $f'(x_n) \neq 0$ dobijemo

$$\alpha = x_n - \frac{f(x_n)}{f'(x_n)} - (\alpha - x_n)^2 \frac{f''(\xi_n)}{2f'(x_n)}.$$

Kako je

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

to dobivamo

$$\alpha - x_{n+1} = -(\alpha - x_n)^2 \frac{f''(\xi_n)}{2f'(x_n)}, \quad n \in \mathbb{N}_0.$$

Iz ove zadnje jednakosti odmah vidimo da je Newtonova metoda, kada konvergira, kvadratno konvergentna. Ipak, takav zaključak vrijedi samo ako $f'(x_n)$ ne teži k nuli tijekom procesa, tj. ako je $f'(\alpha) \neq 0$ ili drugim riječima ako je nultočka $\alpha$ jednostruka.

**TEOREM**. Neka su $f, f'$ i $f''$ neprekidne na intervalu $I \subset \mathbb{R}$, pri čemu $I$ sadrži jednostruku nultočku $\alpha$ funkcije $f$. Ako je početna iteracija $x_0$ izabrana dovoljno blizu nultočke $\alpha$, onda niz iteracija $(x_n)_{n\in\mathbb{N}_0}$ dobiven Newtonovom metodom konvergira prema $\alpha$ s redom konvergencije $p = 2$. Štoviše, vrijedi

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{(\alpha - x_n)^2} = -\frac{f''(\alpha)}{2f'(\alpha)}.$$

**DOKAZ**. Izaberimo okolinu $O = [\alpha - \varepsilon, \alpha + \varepsilon] \subset I$ nultočke $\alpha$ i neka je

$$M = \frac{\max_{x\in O} |f''(x)|}{2\min_{x\in O} |f'(x)|}.$$

Za sve $x_0 \in O$ korištenjem jednakosti

$$\alpha - x_{n+1} = -\left(\alpha - x_n\right)^2 \frac{f''\left(\xi_n\right)}{2f'\left(x_n\right)}, \quad n \in \mathbb{N}_0$$

u posebnom slučaju $n = 0$ dobijemo

$$\left|\alpha - x_1\right| \leq M \left|\alpha - x_0\right|^2,$$

odnosno

$$M \left|\alpha - x_1\right| \leq \left(M \left|\alpha - x_0\right|\right)^2.$$

Izaberimo $x_0 \in O$ tako da zadovoljava uvjet $M \left|\alpha - x_0\right| < 1$. Tada vrijedi

$$M \left|\alpha - x_1\right| \leq M \left|\alpha - x_0\right|,$$

odnosno

$$\left|\alpha - x_1\right| \leq \left|\alpha - x_0\right| \leq \varepsilon,$$

pa i $x_1$ leži u okolini $O$.

Induktivnom primjenom istog argumenta dobivamo

$$|\alpha - x_n| \leq \varepsilon, \quad M\,|\alpha - x_n| < 1$$

za sve $n \in \mathbb{N}$.

Da bismo dokazali konvergenciju iskoristimo opet jednakost

$$\alpha - x_{n+1} = -\left(\alpha - x_n\right)^2 \frac{f''\left(\xi_n\right)}{2f'\left(x_n\right)}, \quad n \in \mathbb{N}_0.$$

Imamo

$$M\,|\alpha - x_{n+1}| \leq \left(M\,|\alpha - x_n|\right)^2, \quad n \in \mathbb{N}_0,$$

pa indukcijom lako pokažemo

$$M\,|\alpha - x_n| \leq \left(M\,|\alpha - x_0|\right)^{2^n}, \quad n \in \mathbb{N},$$

odnosno

$$|\alpha - x_n| \le \frac{1}{M}\left(M\,|\alpha - x_0|\right)^{2^n}, \quad n \in \mathbb{N}.$$

Budući vrijedi

$$M\,|\alpha - x_0| < 1,$$

odmah dobijemo $x_n \to \alpha$ kada $n \to \infty$,a kako $\xi_n$ leži između $x_n$ i $\alpha$ slijedi i da $\xi_n \to \alpha$ kada $n \to \infty$.

Zbog toga je

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{\left(\alpha - x_n\right)^2} = -\lim_{n\to\infty} \frac{f''(\xi_n)}{2f'(x_n)} = -\frac{f''(\alpha)}{2f'(\alpha)} \qquad \blacksquare$$

Prethodni teorem nam daje dovoljne uvjete za **lokalnu konvergenciju** Newtonove metode prema jednostrukoj nultočki $\alpha$. Konvergencija je lokalna jer je postavljen uvjet da početna aproksimacija $x_0$ mora biti dovoljno blizu nultočke $\alpha$. Veličina $\varepsilon$ određena je uvjetom

$$M\left|\alpha - x_0\right| < 1$$

koji osigurava konvergenciju metode.
Da bismo izveli ocjenu greške opet ćemo iskoristiti Taylorov teorem. Za dvije susjedne iteracije u Newtonovoj metodi vrijedi

$$f\left(x_n\right) = f\left(x_{n-1}\right) + f'\left(x_{n-1}\right)\left(x_n - x_{n-1}\right) + \frac{f''\left(\xi_{n-1}\right)}{2}\left(x_n - x_{n-1}\right)^2,$$

pri čemu je $\xi_{n-1}$ između $x_{n-1}$ i $x_n$. Po definiciji iteracija u Newtonovoj metodi vrijedi

$$f\left(x_{n-1}\right) + f'\left(x_{n-1}\right)\left(x_n - x_{n-1}\right) = 0,$$

pa je

$$f\left(x_n\right) = \frac{f''\left(\xi_{n-1}\right)}{2}\left(x_n - x_{n-1}\right)^2.$$

Promotrimo sada slučaj $I = [a, b]$ . Pod pretpostavkom da su $f''$ i $f'$ neprekidne na $[a, b]$ one zasigurno na njemu postižu i svoj minimum i svoj maksimum. Označimo li

$$M = \max_{x \in [a,b]} \left| f''(x) \right|, \quad m = \min_{x \in [a,b]} \left| f'(x) \right|,$$

dobijemo

$$f(x_n) \leq \frac{M}{2} (x_n - x_{n-1})^2 ,$$

a kao i kod metode bisekcije ako je $m \neq 0$ iz Teorema srednje vrijednosti dobijemo i ocjenu

$$|x_n - \alpha| = \frac{|f(x_n)|}{|f'(\xi)|} \leq \frac{|f(x_n)|}{m}.$$

Kombinacijom ovih dviju ocjena dobivamo

$$|x_n - \alpha| \leq \frac{M}{2m} (x_n - x_{n-1})^2 .$$

Ako je $\varepsilon$ gornja ograda za apsolutnu grešku (tj. tražena točnost), onda test

$$\frac{M}{2m}\left(x_n - x_{n-1}\right)^2 \leq \varepsilon,$$

ili (u drugoj formi zapisa)

$$|x_n - x_{n-1}| \leq \sqrt{\frac{2m\varepsilon}{M}},$$

garantira da je

$$|x_n - \alpha| \leq \varepsilon.$$

Naravno, možemo koristiti i prije spomenuti test

$$|x_n - \alpha| \leq \frac{|f\left(x_n\right)|}{m} \leq \varepsilon.$$

U prethodnim ocjenama za lokalnu konvergenciju koristili smo pretpostavke da su $f, f'$ i $f''$ neprekidne sve $x \in [a, b]$ i da je $m \neq 0$. No kako je $m = \min_{x\in[a,b]} |f'(x)|$, to znači da je

$$|f'(x)| > 0, \quad x \in [a, b],$$

pa je $f$ strogo monotona na $[a, b]$. Ako još i druga derivacija $f''$ ima svojstvo da joj je predznak isti na cijelom segmentu $[a, b]$, onda možemo dobiti i uvjete za **globalnu konvergenciju** Newtonove metode.

**TEOREM**. Neka su $f, f'$ i $f''$ neprekidne na segmentu $[a, b] \subset \mathbb{R}$, pri čemu je $f(a) f(b) < 0$, i neka $f'$ i $f''$ nemaju nultočaka u $[a, b]$ (tj. imaju stalan predznak na $[a, b]$). Ako polazna iteracija $x_0 \in [a, b]$ zadovoljava uvjet

$$f(x_0) f''(x_0) > 0,$$

onda niz iteracija dobiven Newtonovom metodom konvergira prema (jedinstvenoj jednostrukoj) nultočki $\alpha$ funkcije $f$.

**DOKAZ**. Uočimo najprije da uvjeti teorema osiguravaju postojanje i jedinstvenost jednostruke nultočke $\alpha$.

Pretpostavimo, na primjer, da je za sve $x \in [a, b]$ ispunjeno $f'(x) > 0$ i $f''(x) > 0$. Tada je $f$ monotono rastuća i zbog početnog uvjeta $f(a) f(b) < 0$ mora vrijediti $f(a) < 0$ i $f(b) > 0$. Jer je $f''$ pozitivna i vrijedi $f(x_0) f''(x_0) > 0$, to za početnu iteraciju $x_0$ mora vrijediti $f(x_0) > 0$. U praksi možemo uzeti $x_0 = b$, jer je to jedina točka za koju sigurno znamo da ispunjava taj uvjet.

Neka je $(x_n)_{n \in \mathbb{N}_0}$ niz iteracija generiran Newtonovom metodom iz startne točke $x_0$ za koju vrijedi $f(x_0) > 0$. Dakle imamo

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

i znamo da je $x_0 > \alpha$ (zbog monotonosti funkcije $f$). Tvrdimo da vrijedi $\alpha < x_n \leq x_0,\ n \in \mathbb{N}_0$. Dokaz provodimo matematičkom indukcijom čiju bazu već imamo. Pretpostavimo da tvrdnja vrijedi za $n \in \mathbb{N}_0$. Po pretpostavci je $f(x_n) > 0$ i $f'(x_n) > 0$, pa je $x_{n+1} < x_n$,što pokazuje da naš niz iteracija monotono pada. Posebno iz toga slijedi $x_{n+1} \leq x_0$ jer je po pretpostavci indukcije $x_n \leq x_0$.

Dokažimo još i lijevu nejednakost u dvostrukoj nejednakosti

$$\alpha < x_{n+1} \leq x_0, \quad n \in \mathbb{N}_0.$$

Po Taylorovoj formuli vrijedi

$$0 = f(\alpha) = f(x_n) + f'(x_n)(\alpha - x_n) + \frac{f''(\xi_n)}{2}(\alpha - x_n)^2,$$

pri čemu je $\xi_n \in (\alpha, x_n) \subset [a, b]$. Zbog $f''(\xi_n) > 0$ imamo

$$f(x_n) + f'(x_n)(\alpha - x_n) < 0,$$

tj.

$$\alpha < x_n - \frac{f(x_n)}{f'(x_n)},$$

odakle slijedi

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} > \alpha.$$

Time je dokazan korak indukcije, pa i sama tvrdnja.

Uočimo da smo usput dokazali i monotonost niza $(x_n)_{n\in\mathbb{N}_0}$. Dokažimo sada da je limes ovog niza upravo nultočka $\alpha$.
Kako je, dakle, niz $(x_n)_{n\in\mathbb{N}_0}$ monotono padajući i omeđen odozdo (s $\alpha$), to postoji

$$\alpha' := \lim_{n\to\infty} x_n$$

za kojeg vrijedi

$$\alpha \leq \alpha' \leq x_0,$$

pa je $\alpha' \in [a,b]$ . Prijelazom na limes u formuli za Newtonove iteracije dobijemo

$$\alpha' = \alpha' - \frac{f(\alpha')}{f'(\alpha')},$$

odakle zbog $f'(\alpha') \neq 0$ slijedi $f(\alpha') = 0$. Kako je $\alpha$ jedina nultočka funkcije $f$ iz $[a,b]$ slijedi

$$\alpha' = \alpha.$$

Preostali slučajevi s obzirom na predznake $f'$ i $f''$ dokažu se analogno $\blacksquare$

Napomenimo da uvjet

$$f(x_0)\, f''(x_0) > 0$$

ima vrlo jednostavno geometrijsko značenje: gledajući graf funkcije $f$ početnu iteraciju $x_0$ trebamo izabrati na "*strmijoj*" strani funkcije.

Računanje korištenjem Newtonove metode može trajati duže nego računanje po metodi sekante iako Newtonova metoda ima veći red konvergencije nego metoda sekante. Objašnjenje leži u činjenici da se za svaki korak Newtonove metode mora izračunati i vrijednost funkcije i vrijednost derivacije u danoj iteraciji, dok se u metodi sekante računa samo vrijednost funkcije.
Metode koje nemaju sigurnu konvergenciju katkad se kombiniraju s metodom polovljenja na sljedeći način:

- izračunamo novu iteraciju po bržoj metodi i ako nije izašla iz danog intervala nastavimo dalje;
- u protivnom napravimo jedan iteracijski korak metodom polovljenja, a zatim se opet vratimo na bržu metodu.

# 6 Metoda jednostavne iteracije

Neka je $g : D \to \mathbb{R}$ neka zadana funkcija. Pretpostavimo da tražimo rješenje jednadžbe

$$x = g(x)$$

koje ćemo označiti s $\alpha$. Definirajmo **jednostavnu iteracijsku funkciju** (jednostavnu u smislu da "pamti" samo jednu prethodnu iteraciju) s

$$x_{n+1} = g(x_n), \quad n \in \mathbb{N}_0,$$

pri čemu je neki $x_0 \in D$ na neki način odabrana prva iteracija (početna aproksimacija za $\alpha$). Primijetimo da Newtonova metoda pripada klasi jednostavnih iteracija jer je u tom slučaju funkcija $g$ definirana s

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

Točke za koje vrijedi $x = g(x)$ nazivaju se **čvrstim** (**fiksnim**) točkama funkcije $g$. Mi smo najčešće zainteresirani za rješavanje jednadžbe oblika $f(x) = 0$, no lako je uočiti da se iz problema $f(x) = 0$ jednostavno prelazi na problem $x = g(x)$.

**PRIMJER**. Pretpostavimo da želimo riješiti jednadžbu

$$x^2 - a = 0, \quad a > 0,$$

ali prelaskom na rješavanje jednadžbe oblika $x = g(x)$. To možemo napraviti na više načina:

- $x = x^2 + x - a,$
- $x = a/x, \quad x \neq 0,$
- $x = \frac{1}{2}\left(x + \frac{a}{x}\right).$

Prirodno se nameće pitanje kako se ponašaju razne jednostavne iteracije.

**LEMA**. Neka je funkcija $g$ neprekidna na $[a, b]$ i neka je

$$(\forall x \in [a, b]) \qquad a \leq g(x) \leq b,$$

ili kraće $g([a, b]) \subseteq [a, b]$ . Tada jednostavna iteracija $x = g(x)$ ima barem jedno rješenje na $[a, b]$ .

**DOKAZ**. Za neprekidnu funkciju $G : [a, b] \to \mathbb{R}$ definiranu s $G(x) = g(x) - x$ vrijedi

$$G(a) \geq 0, \quad G(b) \leq 0,$$

pa ona nužno mijenja predznak na $[a, b]$ . To znači da $G$ ima nultočku na $[a, b]$ , pa postoji rješenje jednadžbe $x = g(x)$ na $[a, b]$ ■

**LEMA**. Neka je funkcija $g$ neprekidna na $[a,b]$ i neka je $g([a,b]) \subseteq [a,b]$. Nadalje, neka postoji konstanta $\lambda \in (0,1)$ takva da vrijedi

$$(\forall x, y \in [a,b]) \ |g(x) - g(y)| \leq \lambda |x - y|$$

(drugim riječima pretpostavimo da je $g$ Lipshitzova za $\lambda \in (0,1)$, odnosno da je **kontrakcija**). Tada jednostavna iteracija $x = g(x)$ ima jedinstveno rješenje $\alpha$ na $[a,b]$. Također, niz iteracija $x_n = g(x_{n-1})$, $n \in \mathbb{N}$, konvergira prema $\alpha$ za proizvoljni $x_0 \in [a,b]$.

**DOKAZ**. Prema prethodnoj lemi znamo da postoji barem jedno rješenje $\alpha \in [a,b]$ jednadžbe $x = g(x)$. Pokažimo da je ono jedinstveno. Dokaz ćemo provesti kontradikcijom: pretpostavimo da postoje dva različita rješenja $\alpha$ i $\beta$. Za njih tada vrijedi $\alpha = g(\alpha)$ i $\beta = g(\beta)$.

Iz ovoga i pretpostavke da je $g$ kontrakcija slijedi

$$|\alpha - \beta| = |g(\alpha) - g(\beta)| \leq \lambda |\alpha - \beta|,$$

odnosno

$$(1 - \lambda) |\alpha - \beta| \leq 0.$$

Budući je po pretpostavci $\lambda \in (0, 1)$, tj. $1 - \lambda > 0$, to mora biti $|\alpha - \beta| = 0$, što je u kontradikciji s pretpostavkom da je $\alpha \neq \beta$. Dakle, ne mogu postojati dva različita rješenja.
Dokažimo još konvergenciju jednostavnih iteracija za proizvoljni $x_0 \in [a, b]$. Uočimo da $x_{n-1} \in [a, b]$ povlaći $x_n = g(x_{n-1}) \in [a, b]$. Dalje, vrijedi

$$|\alpha - x_n| = |g(\alpha) - g(x_{n-1})| \leq \lambda |\alpha - x_{n-1}|,$$

iz čega indukcijom slijedi

$$|\alpha - x_n| \leq \lambda^n |\alpha - x_0|, \quad n \in \mathbb{N}.$$

Ako pustimo da $n$ teži u beskonačno, onda zbog $\lambda \in (0, 1)$ slijedi $\lambda^n \to 0$, pa $x_n \to \alpha$■

No pogledajmo što se događa ako $g$ ima još neke "lijepe" osobine. Ako je $g$ derivabilna na $(a,b)$, onda po *Teoremu srednje vrijednosti* za bilo koje $x,y \in [a,b]$ postoji odgovarajući $\xi \in (a,b)$ takav da vrijedi

$$g(x) - g(y) = g'(\xi)(x-y).$$

Definirajmo

$$\lambda = \max_{x\in(a,b)} |g'(x)|.$$

Tada možemo pisati

$$(\forall x,y \in [a,b]) \ \ |g(x) - g(y)| \le \lambda |x-y|,$$

tj. $g$ je Lipshitzova za tako odabrani $\lambda$. No primijetimo da takav $\lambda$ može biti i veći ili jednak $1$, tj. $g$ ne mora biti kontrakcija.

**TEOREM**. Neka je funkcija $g$ neprekidno derivabilna na $(a, b)$ , neka je ispunjeno $g\left([a, b]\right) \subseteq [a, b]$ , te neka je

$$\lambda = \max_{x \in (a,b)} \left| g'\left(x\right)\right| < 1.$$

Tada vrijedi:

1. Jednadžba $x = g\left(x\right)$ ima točno jedno rješenje $\alpha \in [a, b]$ .
2. Za proizvoljni $x_0 \in [a, b]$ niz jednostavnih iteracija $x_n = g\left(x_{n-1}\right),\ n \in \mathbb{N}$, konvergira prema $\alpha$ i vrijedi

$$\left|\alpha - x_n\right| \leq \lambda^n \left|\alpha - x_0\right|, \quad n \in \mathbb{N},$$

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{\alpha - x_n} = g'\left(\alpha\right).$$

**DOKAZ**. Sve tvrdnje slijede direktno iz prethodnoga osim tvrdnje o brzini konvergencije.

Vrijedi

$$\alpha - x_{n+1} = g(\alpha) - g(x_n) = g'(\xi_n)(\alpha - x_n), \quad n \in \mathbb{N}_0,$$

pri čemu je $\xi_n$ odabran između $\alpha$ i $x_n$. Budući da $x_n \to \alpha$ kada $n \to \infty$, to onda i $\xi_n \to \alpha$ (jer se interval "steže"). Dakle, vrijedi

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{\alpha - x_n} = \lim_{n\to\infty} g'(\xi_n) = g'(\alpha) \quad \blacksquare$$

**TEOREM**. Neka je $\alpha$ rješenje jednadžbe $x = g(x)$ i neka je $g$ neprekidno derivabilna na nekoj okolini $O$ točke $\alpha$. Nadalje, neka je ispunjeno $|g'(\alpha)| < 1$ i neka je početna iteracija $x_0$ izabrana dovoljno blizu $\alpha$. Tada vrijede sve tvrdnje prethodnog teorema.

**DOKAZ**. Definirajmo interval $I = [a, b] = [\alpha - \varepsilon, \alpha + \varepsilon] \subset O$, gdje je $\varepsilon > 0$ odabran tako je $x_0 \in I$ i da vrijedi

$$\max_{x \in [\alpha - \varepsilon, \alpha + \varepsilon]} |g'(x)| = \lambda < 1,$$

(ovo je moguće ispuniti zbog uvjeta $|g'(\alpha)| < 1$ i zbog pretpostavke da je početna iteracija $x_0$ izabrana dovoljno blizu $\alpha$). Tada je $g(I) \subseteq I$. Naime, ako je $x \in I$, tj. $|\alpha - x| \le \varepsilon$, onda za neki $\xi$ između $\alpha$ i $x$ slijedi

$$|\alpha - g(x)| = |g(\alpha) - g(x)| = |g'(\xi)| \, |\alpha - x| \le \lambda \, |\alpha - x| \le \lambda\varepsilon \le \varepsilon,$$

pa je $g(x) \in I$. Dakle, zaista se može primijeniti prethodni teorem, pa odmah slijede i njegove tvrdnje ■

Vratimo se na primjer $x^2 - a = 0,\ a > 0.$

- U slučaju da smo izabrali zapis $x = x^2 + x - a$, imali bismo $g(x) = x^2 + x - a$, pa bi imali $g'(x) = 2x + 1$. Znamo da je pozitivna nultočka ove funkcije $\alpha = \sqrt{a}$, te bi imali $g'(\sqrt{a}) = 2\sqrt{a} + 1 > 1$. Dakle, proces ne bi konvegirao.

- U slučaju $x = a/x$ imali bismo $g'(x) = -a/x^2$, pa je $g'(\sqrt{a}) = -1$ i opet nema konvergencije.

- U slučaju $x = \frac{1}{2}\left(x + \frac{a}{x}\right)$ imali bismo $g'(x) = \frac{1}{2}\left(1 - \frac{a}{x^2}\right)$ , pa je $g'(\sqrt{a}) = 0$, što je dobro.

I na kraju promotrimo kako postići da jednostavne iteracijske funkcije budu višeg reda konvergencije.

**TEOREM**. Neka je $\alpha$ rješenje jednadžbe $x = g(x)$ i neka je funkcija $g$ $p$ $(p \geq 2)$ puta neprekidno derivabilna na nekoj okolini $O$ točke $\alpha$. Nadalje, pretpostavimo da vrijedi

$$g'(\alpha) = g''(\alpha) = \cdots = g^{(p-1)}(\alpha) = 0.$$

Ako je početna iteracija $x_0$ izabrana dovoljno blizu $\alpha$, onda iteracijska funkcija

$$x_{n+1} = g(x_n), \quad n \in \mathbb{N}_0,$$

ima red konvergencije $p$ i vrijedi

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{(\alpha - x_n)^p} = (-1)^{p-1} \frac{g^{(p)}(\alpha)}{p!}.$$

**DOKAZ**. Uočimo najprije da su uvjeti ovog teorema jači nego uvjeti prethodnog teorema, pa odmah možemo zaključiti da niz jednostavnih iteracija konvergira prema $\alpha$. Razvijmo sada funkciju $g$ u red okolini točke $\alpha$ do uključivo potencije reda $(p-1)$ i napišimo ostatak. Uvrstimo za argument $x = x_n$. Dobijemo

$$\begin{aligned} x_{n+1} &= g\left(x_n\right) \\ &= g\left(\alpha\right) + g'\left(\alpha\right)\left(x_n - \alpha\right) + \cdots + \frac{g^{(p-1)}\left(\alpha\right)}{(p-1)!}\left(x_n - \alpha\right)^{p-1} + \frac{g^{(p)}\left(\xi_n\right)}{p!}\left(x_n - \alpha\right)^p, \end{aligned}$$

gdje je $\xi_n$ neka točka između $x_n$ i $\alpha$. Iskoristimo li činjenicu da je $\alpha$ čvrsta točka i pretpostavku $g'\left(\alpha\right) = \cdots = g^{(p-1)}\left(\alpha\right) = 0$, dobijemo

$$x_{n+1} = \alpha + \frac{g^{(p)}\left(\xi_n\right)}{p!}\left(x_n - \alpha\right)^p,$$

odnosno

$$\alpha - x_{n+1} = -\frac{g^{(p)}\left(\xi_n\right)}{p!}\left(x_n - \alpha\right)^p .$$

Kako $x_n \to \alpha$ kada $n \to \infty$, to onda i $\xi_n \to \alpha$ (jer se interval "steže"), pa dobijemo

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{\left(\alpha - x_n\right)^p} = (-1)^{p-1}\,\frac{g^{(p)}\left(\alpha\right)}{p!}.$$

Iz gornje relacije se vidi da je ta konvergencija reda $p$ ■

Korištenjem prethodnog teorema možemo analizirati i Newtonovu metodu za koju je

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

Deriviranjem slijedi

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

i

$$g''(x) = \frac{(f'(x))^2 f''(x) + f(x)\, f'(x)\, f'''(x) - 2f(x)\,(f''(x))^2}{(f'(x))^4} f'(x),$$

pa je uz pretpostavku da je $\alpha$ nultočka funkcije $f$ i da je $f'(\alpha) \neq 0$ (što osigurava jednostrukost) ispunjeno

$$g'(\alpha) = 0, \quad g''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}.$$

Ako je $f''(\alpha) \neq 0$, onda je red konvergencije Newtonove metode jednak 2. No ako je $f'(\alpha) \neq 0$ i $f''(\alpha) = 0$, onda je red konvergencije barem 3.

# 7 Općenito o sustavima nelinearnih jednadžbi

Pretpostavimo da rješavamo sustav nelinearnih jednadžbi

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, \ldots, x_n) &= 0, \end{aligned} \tag{5}$$

pri čemu je $n \geq 2$ i $f_i : \mathbb{R}^n \to \mathbb{R}$ za sve $i \in \{1, \ldots, n\}$ . Ako uvedemo vektorske oznake

$$F(\boldsymbol{x}) = \begin{bmatrix} f_1(\boldsymbol{x}) \\ f_2(\boldsymbol{x}) \\ \vdots \\ f_n(\boldsymbol{x}) \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

gdje je $F : \mathbb{R}^n \to \mathbb{R}^n$, onda sustav $(5)$ možemo zapisati kao

$$F(\boldsymbol{x}) = \mathbf{0}.$$

Prisjetimo se da vrijedi

$$f'(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\boldsymbol{x}) \\ \frac{\partial f}{\partial x_2}(\boldsymbol{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\boldsymbol{x}) \end{bmatrix},$$

te da je funkcija $f$ neprekidno diferencijabilna na otvorenom skupu $D \subset \mathbb{R}^n$ ako je $f'$ (koja očito mora i postojati) neprekidna u svakoj točki $\boldsymbol{x}$ iz skupa $D$. Ako je $f$ neprekidno diferencijabilna na otvorenom i konveksnom skupu $D \subset \mathbb{R}^n$, onda za sve $\boldsymbol{x},\ \boldsymbol{x} + \boldsymbol{p} \in D$ vrijedi sljedeći identitet

$$\begin{aligned} f(\boldsymbol{x} + \boldsymbol{p}) &= f(\boldsymbol{x}) + \int_0^1 \langle \nabla f(\boldsymbol{x} + t\boldsymbol{p}), \boldsymbol{p} \rangle \,\mathrm{d}t \\ &= f(\boldsymbol{x}) + \int_{\boldsymbol{x}}^{\boldsymbol{x}+\boldsymbol{p}} \nabla f(\boldsymbol{z}) \,\mathrm{d}\boldsymbol{z}. \end{aligned}$$

Također znamo da je

$$F'(\boldsymbol{x}) = J(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\boldsymbol{x}) & \cdots & \frac{\partial f_n}{\partial x_n}(\boldsymbol{x}) \end{bmatrix}.$$

Ako je $F$ neprekidno diferencijabilna na otvorenom konveksnom skupu $D \subset \mathbb{R}^n$, onda za sve $\boldsymbol{x},\ \boldsymbol{x} + \boldsymbol{p} \in D$ vrijedi sljedeći identitet

$$F(\boldsymbol{x} + \boldsymbol{p}) = F(\boldsymbol{x}) + \int_0^1 J(\boldsymbol{x} + t\boldsymbol{p})\, \boldsymbol{p} \mathrm{d}t = F(\boldsymbol{x}) + \int_{\boldsymbol{x}}^{\boldsymbol{x}+\boldsymbol{p}} J(\boldsymbol{z})\, \mathrm{d}\boldsymbol{z}.$$

Podsjetimo se još i kada kažemo da je neka funkcija $G : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ Lipschitzova na $D$ (pišemo $G \in$Lip$_\gamma(D)$): ako postoji konstanta $\gamma$ takva da za sve $\boldsymbol{x}, \boldsymbol{y} \in D$ vrijedi

$$\|G(\boldsymbol{y}) - G(\boldsymbol{x})\| \leq \gamma \|\boldsymbol{y} - \boldsymbol{x}\|.$$

Za daljnji rad trebat će nam sljedeća tehnička lema.

**LEMA**. Neka je $F : \mathbb{R}^n \to \mathbb{R}^n$ neprekidno diferencijabilna na otvorenom konveksnom skupu $D \subset \mathbb{R}^n$ i neka je $J = F'$ Lipschitzova na $D$. Tada za sve $\boldsymbol{x},\ \boldsymbol{x} + \boldsymbol{p} \in D$ vrijedi

$$\left\| F\left(\boldsymbol{x} + \boldsymbol{p}\right) - F\left(\boldsymbol{x}\right) - J\left(\boldsymbol{x}\right) \boldsymbol{p} \right\| \leq \frac{\gamma}{2} \left\| \boldsymbol{p} \right\|^2 .$$

**DOKAZ**. Vrijedi

$$\begin{aligned} & F\left(\boldsymbol{x} + \boldsymbol{p}\right) - F\left(\boldsymbol{x}\right) - J\left(\boldsymbol{x}\right) \boldsymbol{p} \\ = & \int_0^1 J\left(\boldsymbol{x} + t\boldsymbol{p}\right) \boldsymbol{p} \mathrm{d}t - J\left(\boldsymbol{x}\right) \boldsymbol{p} \\ = & \int_0^1 \left(J\left(\boldsymbol{x} + t\boldsymbol{p}\right) - J\left(\boldsymbol{x}\right)\right) \boldsymbol{p} \mathrm{d}t, \end{aligned}$$

pa je

$$
\begin{aligned}
& \left\| F\left(\boldsymbol{x}+\boldsymbol{p}\right) - F\left(\boldsymbol{x}\right) - J\left(\boldsymbol{x}\right)\boldsymbol{p} \right\| \\
\leq \; & \int_0^1 \left\| J\left(\boldsymbol{x}+t\boldsymbol{p}\right) - J\left(\boldsymbol{x}\right) \right\| \left\| \boldsymbol{p} \right\| \mathrm{d}t \\
\leq \; & \gamma \int_0^1 \left\| t\boldsymbol{p} \right\| \left\| \boldsymbol{p} \right\| \mathrm{d}t = \gamma \left\| \boldsymbol{p} \right\|^2 \int_0^1 t\mathrm{d}t \\
= \; & \frac{\gamma}{2} \left\| \boldsymbol{p} \right\|^2 .
\end{aligned}
$$

čime je tvrdnja leme dokazana. ■

# 8 Newtonova metoda za rješavanje sustava nelinearnih jednadžbi

Newtonova metoda je jedna od najpoznatijih metoda za rješavanje sustava nelinearnih jednadžbi. Lako je uočiti da je ona u stvari poopćenje Newtonove metode za rješavanje nelinearnih jednadžbi. Ideja metode je u tome da se u relaciji

$$F(\boldsymbol{x}+\boldsymbol{p}) = F(\boldsymbol{x}) + \int_{x}^{x+p} J(\boldsymbol{z})\,\mathrm{d}z.$$

integral aproksimira s $J(\boldsymbol{x})\,\boldsymbol{p}$, pa dobijemo

$$F(\boldsymbol{x}+\boldsymbol{p}) \approx F(\boldsymbol{x}) + J(\boldsymbol{x})\,\boldsymbol{p}.$$

Izjednačimo li $F(\boldsymbol{x}+\boldsymbol{p})$ s nul-vektorom dobijemo

$$\boldsymbol{p} \approx -J^{-1}(\boldsymbol{x})\,F(\boldsymbol{x})$$

iz čega slijedi

$$\boldsymbol{x}+\boldsymbol{p} \approx \boldsymbol{x} - J^{-1}(\boldsymbol{x})\,F(\boldsymbol{x})\,.$$

Zamislimo li da su $\boldsymbol{x}+\boldsymbol{p}$ i $\boldsymbol{x}$ dvije susjedne iteracije u nekom iterativnom procesu dobit ćemo

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - J^{-1}\left(\boldsymbol{x}^{(k)}\right) F\left(\boldsymbol{x}^{(k)}\right). \tag{6}$$

No što računamo ovakvim iterativnim procesom? Pretpostavka

$$F(\boldsymbol{x}+\boldsymbol{p}) = \mathbf{0}$$

koju smo iskoristili znači da mi u stvari tražimo rješenje sustava $F(\boldsymbol{x}) = \mathbf{0}$. Relacijom (6) definirana je Newtonova metoda za rješavanje sustava nelinearnih jednadžbi. U praksi se ova metoda realizira po shemi:

1) riješite linearni sustav

$$J\left(\boldsymbol{x}^{(k)}\right) \boldsymbol{s}^{(k)} = -F\left(\boldsymbol{x}^{(k)}\right),$$

2) stavite

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \boldsymbol{s}^{(k)}.$$

Označimo s

$$K(\boldsymbol{x}^*, r) = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}^* - \boldsymbol{x}\| < r\}$$

otvorenu kuglu u $\mathbb{R}^n$ radiusa $r$ sa središtem u $\boldsymbol{x}^*$ i prisjetimo se činjenice iz linearne algebre koja glasi: ako su $I, E$ matrice takve da je $\|I\| = 1$ i $\|E\| < 1$ onda postoji $(I - E)^{-1}$ i vrijedi

$$\left\|(I - E)^{-1}\right\| \leq \frac{1}{1 - \|E\|}.$$

Iz ovoga slijedi da ako je $A$ regularna matrica i $B$ matrica koja ispunjava uvijet $\left\|A^{-1}(B - A)\right\|$ $1$ onda je i $B$ regularna i vrijedi

$$\left\|B^{-1}\right\| \leq \frac{\left\|A^{-1}\right\|}{1 - \left\|A^{-1}(B - A)\right\|}. \tag{7}$$

Sada imamo sve što nam treba za dokazati teorem o lokalnoj konvergenciji Newtonove metode.

**TEOREM**. Neka je $F:\mathbb{R}^n \to \mathbb{R}^n$ neprekidno diferencijabilna na otvorenom konveksnom skupu $D \subset \mathbb{R}^n$. Neka su $\boldsymbol{x}^* \in \mathbb{R}^n,\ r, \beta \in \mathbb{R}_+$ takvi da vrijedi:

$1)\ K\left(\boldsymbol{x}^*, r\right) \subset D,$
$2)\ F\left(\boldsymbol{x}^*\right) = 0,$
$3)\ J^{-1}\left(\boldsymbol{x}^*\right)$ postoji i $\left\|J^{-1}\left(\boldsymbol{x}^*\right)\right\| \leq \beta,$
$4)\ J \in$Lip$_\gamma\left(K\left(\boldsymbol{x}^*, r\right)\right).$

Tada postoji $\varepsilon \in (0, r)$ takav da je za svaki $\boldsymbol{x}^{(0)} \in K\left(\boldsymbol{x}^*, \varepsilon\right)$ niz $\left(\boldsymbol{x}^{(k)}\right)_{\mathbb{N}_0}$ generiran Newtonovom metodom dobro definiran i konvergira k $\boldsymbol{x}^*$, a pri tom zadovoljava ocjenu

$$\left\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^*\right\| \leq \beta\gamma \left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\right\|^2.$$

**DOKAZ**. Ideja je vrijednost $\varepsilon \in (0, r)$ izabrati tako da matrica $J(\boldsymbol{x})$ bude regularna za svaki $\boldsymbol{x} \in K(\boldsymbol{x}^*, \varepsilon)$ (što je moguće postići zbog uvjeta br. 3) i da svaka sljedeća iteracija dobivena Newtonovom metodom leži u istoj kugli i to upola bliže točnom rješenju $\boldsymbol{x}^*$. U tom cilju stavimo

$$\varepsilon = \min\left\{r, \frac{1}{2\beta\gamma}\right\}.$$

Dokazat ćemo da za tako izabrani $\varepsilon$ vrijedi

$$\left\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^*\right\| \leq \beta\gamma \left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\right\|^2,$$

te da je

$$\left\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^*\right\| \leq \frac{1}{2} \left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\right\|,$$

iz čega zaključujemo da pretpostavka $\boldsymbol{x}^{(k)} \in K(\boldsymbol{x}^*, \varepsilon)$ povlači $\boldsymbol{x}^{(k+1)} \in K(\boldsymbol{x}^*, \varepsilon)$ za sve $k \in \mathbb{N}_0$.

Najprije moramo postići da iz pretpostavke da je matrica $J\left(\boldsymbol{x}^{(k)}\right)$ regularna slijedi i da je matrica $J\left(\boldsymbol{x}^{(k+1)}\right)$ regularna, čime bi niz iteracija $\left(\boldsymbol{x}^{(k)}\right)_{\mathbb{N}_0}$ bio dobro definiran čim je ispunjen uvjet da je matrica $J\left(\boldsymbol{x}^{(0)}\right)$ regularna.
Iz pretpostavki da $J^{-1}\left(\boldsymbol{x}^{*}\right)$ postoji te da vrijedi $\left\|J^{-1}\left(\boldsymbol{x}^{*}\right)\right\| \leq \beta$, da je $\boldsymbol{x}^{(0)} \in K\left(\boldsymbol{x}^{*}, \varepsilon\right)$ i $J \in$Lip$_\gamma\left(K\left(\boldsymbol{x}^{*}, r\right)\right)$ dobijemo

$$\begin{aligned}
&\left\|J^{-1}\left(\boldsymbol{x}^{*}\right)\left(J\left(\boldsymbol{x}^{(0)}\right)-J\left(\boldsymbol{x}^{*}\right)\right)\right\| \\
\leq\ &\left\|J^{-1}\left(\boldsymbol{x}^{*}\right)\right\|\left\|J\left(\boldsymbol{x}^{(0)}\right)-J\left(\boldsymbol{x}^{*}\right)\right\| \\
\leq\ &\beta\gamma\left\|\boldsymbol{x}^{(0)}-\boldsymbol{x}^{*}\right\| \leq \beta\gamma\varepsilon \leq \frac{1}{2}<1.
\end{aligned}$$

Iskoristimo li (7) za $A=J\left(\boldsymbol{x}^{*}\right)$ i $B=J\left(\boldsymbol{x}^{(0)}\right)$ dobijemo da je i $J\left(\boldsymbol{x}^{(0)}\right)$ regularna te da vrijedi

$$\left\|J^{-1}\left(\boldsymbol{x}^{(0)}\right)\right\| \leq \frac{\left\|J^{-1}\left(\boldsymbol{x}^{*}\right)\right\|}{1-\left\|J^{-1}\left(\boldsymbol{x}^{*}\right)\left(J\left(\boldsymbol{x}^{(0)}\right)-J\left(\boldsymbol{x}^{*}\right)\right)\right\|} \leq 2\left\|J^{-1}\left(\boldsymbol{x}^{*}\right)\right\| \leq 2\beta.$$

Zbog činjenice da je $J\left(\boldsymbol{x}^{(0)}\right)$ regularna $\boldsymbol{x}^{(1)}$ je dobro definirano i vrijedi

$$\begin{aligned}
\boldsymbol{x}^{(1)}-\boldsymbol{x}^{*} &= \boldsymbol{x}^{(0)}-J^{-1}\left(\boldsymbol{x}^{(0)}\right)F\left(\boldsymbol{x}^{(0)}\right)-\boldsymbol{x}^{*} \\
&= \boldsymbol{x}^{(0)}-\boldsymbol{x}^{*}-J^{-1}\left(\boldsymbol{x}^{(0)}\right)\left(F\left(\boldsymbol{x}^{(0)}\right)-F\left(\boldsymbol{x}^{*}\right)\right) \\
&= J^{-1}\left(\boldsymbol{x}^{(0)}\right)\left[F\left(\boldsymbol{x}^{*}\right)-F\left(\boldsymbol{x}^{(0)}\right)-J\left(\boldsymbol{x}^{(0)}\right)\left(\boldsymbol{x}^{*}-\boldsymbol{x}^{(0)}\right)\right].
\end{aligned}$$

Zbog $\left\|J^{-1}\left(\boldsymbol{x}^{(0)}\right)\right\| \leq 2\beta$ i prethodne leme slijedi

$$\begin{aligned}
&\left\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{*}\right\| \\
\leq\ &\left\|J^{-1}\left(\boldsymbol{x}^{(0)}\right)\right\|\left\|F\left(\boldsymbol{x}^{*}\right)-F\left(\boldsymbol{x}^{(0)}\right)-J\left(\boldsymbol{x}^{(0)}\right)\left(\boldsymbol{x}^{*}-\boldsymbol{x}^{(0)}\right)\right\| \\
\leq\ &2\beta\frac{\gamma}{2}\left\|\boldsymbol{x}^{*}-\boldsymbol{x}^{(0)}\right\|^{2}=\beta\gamma\left\|\boldsymbol{x}^{(0)}-\boldsymbol{x}^{*}\right\|^{2},
\end{aligned}$$

čime je tvrdnja teorema dokazana za $k=0$.

Kako je $\boldsymbol{x}^{(0)} \in K(\boldsymbol{x}^*, \varepsilon)$ vrijedi

$$\left\| \boldsymbol{x}^* - \boldsymbol{x}^{(0)} \right\| \leq \varepsilon \leq \frac{1}{2\beta\gamma},$$

pa je i

$$\begin{aligned} \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(1)} \right\| &\leq \beta\gamma \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(0)} \right\|^2 = \beta\gamma \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(0)} \right\| \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(0)} \right\| \\ &\leq \beta\gamma \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(0)} \right\| \frac{1}{2\beta\gamma} = \frac{1}{2} \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(0)} \right\| \leq \frac{1}{4\beta\gamma}. \end{aligned}$$

Iz gornjega slijedi da je $\boldsymbol{x}^{(1)} \in K(\boldsymbol{x}^*, \varepsilon)$ i da je, štoviše, upola bliži središtu kugle (točnom rješenju) $\boldsymbol{x}^*$ od $\boldsymbol{x}^{(0)}$. Stoga se postupak može nastaviti induktivno zamjenama $0 \to k$ i $1 \to k+1$, no radi islustracije pokazat ćemo ipak još jedan korak.

Iz pretpostavki teorema i prethodno dokazanih nejednakosti imamo:

$$\begin{aligned}
&\left\| J^{-1}\left(\boldsymbol{x}^*\right)\left(J\left(\boldsymbol{x}^{(1)}\right)-J\left(\boldsymbol{x}^*\right)\right)\right\| \\
\leq\ & \left\| J^{-1}\left(\boldsymbol{x}^*\right)\right\| \left\| J\left(\boldsymbol{x}^{(1)}\right)-J\left(\boldsymbol{x}^*\right)\right\| \\
\leq\ & \beta\gamma \left\| \boldsymbol{x}^{(1)}-\boldsymbol{x}^*\right\| \leq \beta\gamma\frac{1}{4\beta\gamma} \leq \frac{1}{4} < 1.
\end{aligned}$$

Iskoristimo li (7) za $A = J\left(\boldsymbol{x}^*\right)$ i $B = J\left(\boldsymbol{x}^{(1)}\right)$ dobijemo da je i $J\left(\boldsymbol{x}^{(1)}\right)$ regularna te da vrijedi

$$\left\| J^{-1}\left(\boldsymbol{x}^{(1)}\right)\right\| \leq \frac{\left\| J^{-1}\left(\boldsymbol{x}^*\right)\right\|}{1-\left\| J^{-1}\left(\boldsymbol{x}^*\right)\left(J\left(\boldsymbol{x}^{(1)}\right)-J\left(\boldsymbol{x}^*\right)\right)\right\|} \leq \frac{4}{3}\left\| J^{-1}\left(\boldsymbol{x}^*\right)\right\| \leq \frac{4}{3}\beta.$$

Zbog činjenice da je $J\left(\boldsymbol{x}^{(1)}\right)$ regularna $\boldsymbol{x}^{(2)}$ je dobro definiran i vrijedi

$$\begin{aligned}\boldsymbol{x}^{(2)}-\boldsymbol{x}^{*} &= \boldsymbol{x}^{(1)}-J^{-1}\left(\boldsymbol{x}^{(1)}\right)F\left(\boldsymbol{x}^{(1)}\right)-\boldsymbol{x}^{*}\\ &= \boldsymbol{x}^{(1)}-\boldsymbol{x}^{*}-J^{-1}\left(\boldsymbol{x}^{(1)}\right)\left(F\left(\boldsymbol{x}^{(1)}\right)-F\left(\boldsymbol{x}^{*}\right)\right)\\ &= J^{-1}\left(\boldsymbol{x}^{(1)}\right)\left[F\left(\boldsymbol{x}^{*}\right)-F\left(\boldsymbol{x}^{(1)}\right)-J\left(\boldsymbol{x}^{(1)}\right)\left(\boldsymbol{x}^{*}-\boldsymbol{x}^{(1)}\right)\right].\end{aligned}$$

Zbog $\left\|J^{-1}\left(\boldsymbol{x}^{(1)}\right)\right\| \leq \frac{4}{3}\beta$ i prethodne leme slijedi

$$\begin{aligned}&\left\|\boldsymbol{x}^{(2)}-\boldsymbol{x}^{*}\right\|\\ \leq\ &\left\|J^{-1}\left(\boldsymbol{x}^{(1)}\right)\right\|\left\|F\left(\boldsymbol{x}^{*}\right)-F\left(\boldsymbol{x}^{(1)}\right)-J\left(\boldsymbol{x}^{(1)}\right)\left(\boldsymbol{x}^{*}-\boldsymbol{x}^{(1)}\right)\right\|\\ \leq\ &\frac{4}{3}\beta\frac{\gamma}{2}\left\|\boldsymbol{x}^{*}-\boldsymbol{x}^{(1)}\right\|^{2}=\frac{2}{3}\beta\gamma\left\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{*}\right\|^{2}\\ \leq\ &\beta\gamma\left\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{*}\right\|^{2},\end{aligned}$$

čime je tvrdnja teorema dokazana za $k=1$.

Kako vrijedi

$$\left\| \boldsymbol{x}^* - \boldsymbol{x}^{(1)} \right\| \leq \frac{1}{4\beta\gamma},$$

dobijemo još i

$$\begin{aligned} \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(2)} \right\| &\leq \beta\gamma \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(1)} \right\|^2 = \beta\gamma \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(1)} \right\| \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(1)} \right\| \\ &\leq \beta\gamma \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(1)} \right\| \frac{1}{4\beta\gamma} = \frac{1}{2} \left\| \boldsymbol{x}^* - \boldsymbol{x}^{(1)} \right\| \leq \frac{1}{8\beta\gamma}. \end{aligned}$$

Iz gornjega slijedi da je $\boldsymbol{x}^{(2)} \in K\left(\boldsymbol{x}^*, \varepsilon\right)$ i da je, štoviše, upola bliži središtu kugle (točnom rješenju) $\boldsymbol{x}^*$ od $\boldsymbol{x}^{(1)}$ ■