Numerička matematika

13. predavanje

Saša Singer i Nela Bosner

PMF - Matematički odsjek, Zagreb

Sadržaj predavanja

Rješavanje nelinearnih jednadžbi

Derivabilne iteracijske funkcije i metode višeg reda

Newtonova metoda za višestruke nultočke

Primjeri za jednostruke nultočke

Primjeri za višestruke nultočke

Rješavanje sustava nelinearnih jednadžbi

Newtonova metoda

Metoda sekante — Broydenova metoda

Primjer rješavanja sustava nelinearnih jednadžbi

Jednostavne derivabilne iteracijske funkcije

Pretpostavimo sad da je φ neprekidno derivabilna na $[a, b]$.

Po teoremu srednje vrijednosti, za bilo koje $x, y \in [a, b]$, vrijedi

$$\varphi(x) - \varphi(y) = \varphi'(\xi)(x - y),$$

gdje je ξ između x i y, tj. vrijedi $\xi \in [a, b]$. Definiramo

$$q := M_1 = \max_{x \in [a,b]} |\varphi'(x)|.$$

Onda možemo pisati

$$|\varphi(x) - \varphi(y)| \leq q|x - y|, \quad \forall x, y \in [a, b].$$

Primijetite da q, općenito, može biti i veći od 1. No, ako je $q < 1$ i još vrijedi $\varphi([a, b]) \subseteq [a, b]$, onda je φ kontrakcija.

Jednostavne derivabilne iteracijske funkcije

Teorem. Neka je $\varphi \in C^1[a,b]$, takva da je $\varphi([a,b]) \subseteq [a,b]$. Ako je

$$q := M_1 = \max_{x\in[a,b]} |\varphi'(x)| < 1,$$

onda jednadžba $x = \varphi(x)$ ima točno jedno rješenje $\alpha \in [a,b]$.

- Za proizvoljnu startnu točku $x_0 \in [a,b]$ i niz jednostavnih iteracija definiran s $x_{n+1} = \varphi(x_n)$, za $n \geq 0$, vrijedi

$$\lim_{n\to\infty} x_n = \alpha,$$

$$|\alpha - x_n| \leq q^n |\alpha - x_0|, \quad n \geq 0,$$

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

Jednostavne derivabilne iteracijske funkcije

Dokaz. Sve tvrdnje ovog teorema su dokazane u prethodnom teoremu, osim zadnje tvrdnje o linearnoj brzini konvergencije.

Po teoremu srednje vrijednosti, imamo

$$\alpha - x_{n+1} = \varphi(\alpha) - \varphi(x_n) = \varphi'(\xi_n)(\alpha - x_n), \quad n \geq 0,$$

gdje je ξ_n neki broj između α i x_n.

Budući da $x_n \to \alpha$, onda i $\xi_n \to \alpha$. Zbog neprekidnosti derivacije φ' u fiksnoj točki α, onda vrijedi i

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

∎

Bitna pretpostavka $q < 1$

Pretpostavka $q < 1$ u prethodnom teoremu je ključna.

Kontraprimjer. Pretpostavimo "samo" da je $|\varphi'(\alpha)| > 1$, u fiksnoj točki α funkcije φ.

Za neku startnu točku $x_0 \in [a, b]$, generiramo niz jednostavnih iteracija $x_{n+1} = \varphi(x_n)$. Zbog $\alpha = \varphi(\alpha)$, onda vrijedi

$$\alpha - x_{n+1} = \varphi(\alpha) - \varphi(x_n) = \varphi'(\xi_n)(\alpha - x_n).$$

Za bilo koji x_n dovoljno blizu α, mora biti $|\varphi'(\xi_n)| > 1$. Ako je $x_n \neq \alpha$, onda je

$$|\alpha - x_{n+1}| > |\alpha - x_n|.$$

Dakle, konvergencija metode nije moguća! Upravo suprotno, imamo divergenciju (iteracije se udaljuju od α, ako su blizu α).

Pojednostavljeni — “lokalni” teorem

Prethodni teorem možemo izreći i u “lokalnoj” formi — oko α.

Teorem. Neka je α rješenje jednostavne iteracije $x = \varphi(x)$ i neka je φ neprekidno diferencijabilna na nekoj okolini od α. Ako je $|\varphi'(\alpha)| < 1$, onda vrijede svi rezultati prethodnog teorema — uz pretpostavku da je start x_0 dovoljno blizu α.

Dokaz. Postoji $\varepsilon > 0$ takav da za interval $I = [\alpha - \varepsilon, \alpha + \varepsilon]$ vrijedi

$$\max_{x \in I} |\varphi'(x)| \leq q < 1.$$

Tada je $\varphi(I) \subseteq I$ (dovoljno je $q \leq 1$), jer $|\alpha - x| \leq \varepsilon$ povlači

$$|\alpha - \varphi(x)| = |\varphi(\alpha) - \varphi(x)| = |\varphi'(\xi)||\alpha - x| \leq q|\alpha - x| \leq \varepsilon.$$

Sada možemo primijeniti prethodni teorem za $[a, b] = I$. ∎

Primjer iteracijskih funkcija za $\sqrt{a}$

Primjer. Za problem $x^2 - a = 0$, gdje je $a > 0$, definirali smo tri jednostavne iteracijske funkcije:

1. $x = x^2 + x - a$, ili općenitije, $x = x + c(x^2 - a)$, za $c \neq 0$,
2. $x = a/x$,
3. $x = \frac{1}{2}\,(x + a/x)$.

Ispitajte konvergenciju ovih iteracijskih funkcija oko $\alpha = \sqrt{a}$.

1. Za $\varphi(x) = x^2 + x - a$, izlazi $\varphi'(x) = 2x + 1$. U $x = \sqrt{a}$ je

$$\varphi'(\sqrt{a}) = 2\sqrt{a} + 1 > 1,$$

   pa ta iteracijska funkcija neće konvergirati. Baš suprotno, za bilo koji start x_0, ove iteracije divergiraju!

Primjer iteracijskih funkcija za $\sqrt{a}$

1. Općenito, $\varphi(x) = x + c(x^2 - a)$, pa je $\varphi'(x) = 1 + 2cx$ i
$$\varphi'(\sqrt{a}) = 1 + 2c\sqrt{a}.$$
Da bismo osigurali lokalnu konvergenciju, mora biti
$$-1 < 1 + 2c\sqrt{a} < 1,$$
odnosno,
$$-\frac{1}{\sqrt{a}} < c < 0.$$
2. Za $\varphi(x) = a/x$, dobivamo $\varphi'(x) = -a/x^2$, pa je
$$\varphi'(\sqrt{a}) = -1,$$
što znači da iteracijska funkcija neće konvergirati. Niz iteracija je $x_0,\ a/x_0,\ x_0,\ a/x_0,\ \ldots$ (periodički niz).

Primjer iteracijskih funkcija za $\sqrt{a}$

3. Za $\varphi(x) = \frac{1}{2}\,(x + a/x)$, izlazi $\varphi'(x) = \frac{1}{2}\,(1 - a/x^2)$, pa je

$$\varphi'(\sqrt{a}) = 0.$$

   Zato ova iteracijska funkcija konvergira u okolini $\alpha = \sqrt{a}$.

Posljednja iteracijska funkcija je baš Newtonova metoda za jednadžbu $x^2 - a = 0$, a poznavali su ju još Babilonci.

Vidimo da metoda jednostavne iteracije može imati

- lokalnu konvergenciju koja je brža od linearne.

Stvarni "krivac" za kvadratnu konvergenciju je $\varphi'(\alpha) = 0$.

Slično tome, jednostavne iteracijske funkcije mogu poslužiti za konstrukciju iterativnih metoda proizvoljno visokog reda p.

Iterativne metode višeg reda konvergencije

Teorem. Neka je α rješenje jednadžbe $x = \varphi(x)$ i neka je φ

- p puta neprekidno diferencijabilna za sve x u okolini α, za neki $p \geq 2$.

Nadalje, pretpostavimo da je

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

Ako je startna točka x_0 dovoljno blizu α, onda niz iteracija

$$x_{n+1} = \varphi(x_n), \quad n \geq 0,$$

konvergira prema α s redom konvergencije (barem) p i vrijedi

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

Iterativne metode višeg reda konvergencije

Dokaz. Funkciju φ razvijemo, u okolini od α, u Taylorov polinom stupnja p, s tim da najviši član predstavlja ostatak. Zatim uvrstimo $x = x_n$, pa dobijemo

$$\begin{aligned}\varphi(x_n) &= \varphi(\alpha) + \varphi'(\alpha)(x_n - \alpha) + \cdots \\ &\qquad + \frac{\varphi^{(p-1)}(\alpha)}{(p-1)!}\,(x_n - \alpha)^{p-1} + \frac{\varphi^{(p)}(\xi_n)}{p!}\,(x_n - \alpha)^p,\end{aligned}$$

za neki ξ_n između x_n i α.

Sad iskoristimo da je $\varphi(\alpha) = \alpha$, $\varphi(x_n) = x_{n+1}$ i pretpostavku da za derivacije vrijedi $\varphi^{(k)}(\alpha) = 0$, za $k = 1, \ldots, p-1$. Izlazi

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

Iterativne metode višeg reda konvergencije

odnosno,

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

Zbog $\varphi'(\alpha) = 0 < 1$, iz “lokalnog” teorema slijedi da

- niz iteracija x_n konvergira prema α, za svaku startnu točku x_0 koja je dovoljno blizu α (lokalna konvergencija).

Iz $x_n \to \alpha$ slijedi i $\xi_n \to \alpha$. Na kraju, u gornjoj relaciji, na limesu $n \to \infty$, iskoristimo neprekidnost $\varphi^{(p)}$ u α. Izlazi

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

■

Za $p = 1$, ovaj rezultat odgovara ranijem “lokalnom” teoremu!

Primjer — analiza Newtonove metode

Primjer. Primjenom prethodnog teorema, analizirajmo red konvergencije Newtonove metode u okolini jednostruke nultočke α funkcije f. Pripadna iteracijska funkcija φ je

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

Deriviranjem dobivamo da je

$$\varphi'(x) = 1 - \frac{[f'(x)]^2 - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2}.$$

Nultočka α je jednostruka, pa je $f(\alpha) = 0$ i $f'(\alpha) \neq 0$. Onda je

$$\varphi'(\alpha) = 0,$$

tj. Newtonova metoda konvergira (barem) kvadratno oko α.

Primjer — analiza Newtonove metode

Za detaljniju analizu, pogledajmo drugu derivaciju $\varphi''(\alpha)$.
Deriviranjem $\varphi'(x)$ u produktnom obliku, dobivamo

$$\begin{aligned}\varphi''(x) &= \left(f(x) \frac{f''(x)}{[f'(x)]^2} \right)' = f'(x) \frac{f''(x)}{[f'(x)]^2} + f(x) \left(\frac{f''(x)}{[f'(x)]^2} \right)' \\ &= \frac{f''(x)}{f'(x)} + f(x) \left(\frac{f''(x)}{[f'(x)]^2} \right)'.\end{aligned}$$

Zbog $f(\alpha) = 0$ i $f'(\alpha) \neq 0$, odmah izlazi

$$\varphi''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}.$$

Ako je $f''(\alpha) \neq 0$, red konvergencije Newtonove metode je 2.

Ako je $f''(\alpha) = 0$, red konvergencije je barem 3.

Multiplicitet nultočke funkcije

Definicija (Multiplicitet nultočke). Neka je $f(\alpha) = 0$. Ako postoji prirodni broj m, takav da je

$$f(x) = (x - \alpha)^m g(x), \quad g(\alpha) \neq 0,$$

gdje je g neprekidna funkcija u okolini od α, onda nultočka α ima multiplicitet (višestrukost, kratnost ili red) m. ■

Teorem. Pretpostavimo da funkcija f ima neprekidnu m-tu derivaciju u okolini od α, tj. f je klase C^m oko α. Onda je

- α nultočka od f multipliciteta m, ako i samo ako vrijedi

$$f(\alpha) = f'(\alpha) = \cdots = f^{(m-1)}(\alpha) = 0, \quad f^{(m)}(\alpha) \neq 0.$$

Dokaz "$\Leftarrow$" ide direktno iz Taylorovog razvoja do stupnja m, za funkciju f oko α (slično kao malo prije — u teoremu za φ).

Nultočka funkcije i njezinih derivacija

Obrat "$\Rightarrow$" je nešto složeniji, jer g ne mora biti derivabilna. Zato ne "ide" Leibnizovo pravilo za derivaciju produkta!

Početak: Za $k = 0, \ldots, m$, definiramo funkciju g_k oko α, kao

$$g_k(x) = \frac{f^{(k)}(x)}{(x - \alpha)^{m-k}}, \quad x \neq \alpha.$$

Uočimo da je $g_0(x) = g(x)$. Po pretpostavci, g_0 je neprekidna u α i vrijedi $f(\alpha) = 0$, $g_0(\alpha) \neq 0$ (baza indukcije, zbog $m \geq 1$).

Korak: Neka je $k \geq 0$ i $k < m$, i uzmimo da za k vrijedi

$$f^{(k)}(\alpha) = 0, \quad \text{i postoji} \quad \lim_{x \to \alpha} g_k(x) = g_k(\alpha) \neq 0,$$

gdje je $g_k(\alpha)$ definiran proširenjem po neprekidnosti, tako da je g_k neprekidna u α, pa onda i oko α (iz definicije za $x \neq \alpha$).

Nultočka funkcije i njezinih derivacija

Onda je

$$\frac{f^{(k)}(x) - f^{(k)}(\alpha)}{x - \alpha} = \frac{f^{(k)}(x)}{x - \alpha} = (x - \alpha)^{m-k-1} g_k(x).$$

Desna strana je neprekidna u α, pa prijelazom na limes $x \to \alpha$ slijedi

$$f^{(k+1)}(\alpha) = \begin{cases} 0, & \text{za } k + 1 < m, \\ g_k(\alpha), & \text{za } k + 1 = m. \end{cases}$$

Treba još pokazati da postoji $\lim\limits_{x \to \alpha} g_{k+1}(x)$ i da je $g_{k+1}(\alpha) \neq 0$.

Ako je $k + 1 = m$, onda je (po definiciji) $g_m(x) = f^{(m)}(x)$, pa neprekidnost $f^{(m)}$ u α daje

$$g_m(\alpha) = g_{m-1}(\alpha) \neq 0.$$

Nultočka funkcije i njezinih derivacija

Za $k+1 < m$, iz definicije dobivamo neodređeni oblik $0/0$, kojeg računamo “obratnim” L'Hospitalovim pravilom, tj.

- integriramo brojnik i nazivnik, a ne deriviramo.

Izlazi

$$\begin{aligned}\lim_{x\to\alpha} g_{k+1}(x) &= \lim_{x\to\alpha} \frac{f^{(k+1)}(x)}{(x-\alpha)^{m-k-1}} = \{\text{ L'Hospital }\} \\ &= \lim_{x\to\alpha} \frac{f^{(k)}(x)}{\frac{1}{m-k}(x-\alpha)^{m-k}} \\ &= (m-k) \lim_{x\to\alpha} g_k(x) = (m-k)g_k(\alpha) \neq 0.\end{aligned}$$

Na kraju, vidimo da je $g_k(\alpha) = m(m-1)\cdots(m-k+1)g(\alpha)$, za $k = 1, \ldots, m$. Posebno, vrijedi $g_1(\alpha) = mg(\alpha)$.

■

Veza nultočke funkcije i derivacije

U nastavku, radi jednostavnosti, pretpostavljamo da je i funkcija g dovoljno glatka oko α, tako da ju možemo derivirati koliko puta treba. Onda prethodni obrat ide puno lakše.

Na primjer, uzmimo da je g' neprekidna oko α. Pokažimo da

- ako funkcija f ima nultočku multipliciteta m u α,
- onda derivacija f' ima nultočku multipliciteta $m - 1$ u α.

Dokaz. Napišimo f u obliku

$$f(x) = (x - \alpha)^m g(x), \quad g(\alpha) \neq 0.$$

Onda je

$$\begin{aligned} f'(x) &= m(x - \alpha)^{m-1} g(x) + (x - \alpha)^m g'(x) \\ &= (x - \alpha)^{m-1} \Big(mg(x) + (x - \alpha)\, g'(x) \Big). \end{aligned}$$

Veza nultočke funkcije i derivacije

Zatim, definiramo funkciju g_1 na okolini od α, kao drugi faktor iz prethodnog produkta

$$g_1(x) := mg(x) + (x - \alpha)\, g'(x),$$

tako da je

$$f'(x) = (x - \alpha)^{m-1}\, g_1(x).$$

Iz pretpostavke da je g' neprekidna u okolini od α, slijedi da su f' i g_1, također, neprekidne oko α. U točki α je

$$g_1(\alpha) = mg(\alpha) + (\alpha - \alpha)\, g'(\alpha) = mg(\alpha) \neq 0,$$

što pokazuje da f' ima $(m - 1)$-struku nultočku u α. ∎

Ovu formulu za $g_1(\alpha)$ koristimo više puta u nastavku.

Dodatno, uzimamo da je g klase C^2 oko α, iako ide i bez toga.

Konvergencija Newtona za višestruku nultočku

Analizirajmo što će se dogoditi s konvergencijom Newtonove metode, ako funkcija f ima višestruku nultočku u α.

Newtonovu metodu promatramo kao jednostavnu iteraciju,

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

Pretpostavimo da

- f ima m-struku nultočku u α, za neki $m \geq 2$, i da je
- f dovoljno glatka na okolini od α — barem klase C^{m+1}, tako da je iteracijska funkcija φ barem klase C^m.

Onda je

$$\begin{aligned} f(x) &= (x - \alpha)^m \, g(x), & g(\alpha) &\neq 0, \\ f'(x) &= (x - \alpha)^{m-1} \, g_1(x), & g_1(\alpha) &\neq 0. \end{aligned}$$

Konvergencija Newtona za višestruku nultočku

Kad uvrstimo ove relacije za f i f', dobivamo

$$\varphi(x) = x - \frac{(x-\alpha)^m\, g(x)}{(x-\alpha)^{m-1}\, g_1(x)} = x - (x-\alpha)\frac{g(x)}{g_1(x)}.$$

Deriviranjem izlazi

$$\varphi'(x) = 1 - \frac{g(x)}{g_1(x)} - (x-\alpha)\left(\frac{g(x)}{g_1(x)}\right)'.$$

U nultočki α je

$$\varphi'(\alpha) = 1 - \frac{g(\alpha)}{g_1(\alpha)} = 1 - \frac{g(\alpha)}{mg(\alpha)} = 1 - \frac{1}{m}.$$

Za $m \geq 2$, vrijedi $\varphi'(\alpha) \neq 0$. Prema ranijem teoremu, to znači

- da Newtonova metoda onda konvergira samo linearno!

Konvergencija Newtona za višestruku nultočku

Dodatno, faktor linearne konvergencije je $\varphi'(\alpha) = 1 - 1/m$, što je vrlo sporo. U prosjeku, to je

- podjednako brzo kao bisekcija, za $m = 2$,
- ili čak lošije od bisekcije, za $m \geq 3$.

Newtonovu metodu možemo popraviti na dva načina:

- ako unaprijed točno znamo red m nultočke,
- ako ne znamo red (višestruke) nultočke.

Ako znamo m, onda u okolini m-struke nultočke definiramo iteracijsku funkciju

$$\varphi(x) = x - m\,\frac{f(x)}{f'(x)}.$$

Newtonova metoda kad znamo red nultočke

Na isti način kao malo prije, redom, dobivamo

$$\varphi(x) = x - m(x-\alpha)\frac{g(x)}{g_1(x)},$$

$$\varphi'(x) = 1 - m\frac{g(x)}{g_1(x)} - m(x-\alpha)\left(\frac{g(x)}{g_1(x)}\right)',$$

pa je

$$\varphi'(\alpha) = 1 - m\frac{g(\alpha)}{g_1(\alpha)} = 1 - m\frac{g(\alpha)}{mg(\alpha)} = 1 - m\frac{1}{m} = 0.$$

To pokazuje da ova modifikacija Newtonove metode,

- s m-strukom korekcijom,
- osigurava barem kvadratnu konvergenciju, za bilo koji m.

Newtonova metoda kad ne znamo red nultočke

Što ćemo napraviti ako unaprijed ne znamo m? Primijetimo da funkcija u — to je korekcija u običnoj Newtonovoj metodi,

$$u(x) := \frac{f(x)}{f'(x)} = \frac{(x-\alpha)^m\, g(x)}{(x-\alpha)^{m-1}\, g_1(x)} = (x-\alpha)\frac{g(x)}{g_1(x)},$$

uvijek ima jednostruku nultočku u α, jer je $g(\alpha)/g_1(\alpha) \neq 0$.

Dakle, obična Newtonova metoda, primijenjena na u (a ne f)

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

gdje je

$$u'(x) = \frac{[f'(x)]^2 - f(x)f''(x)}{[f'(x)]^2} = 1 - \frac{f''(x)}{f'(x)}\, u(x),$$

konvergira barem kvadratno, iako ne znamo red nultočke!

Ostale metode kad ne znamo red nultočke

Sasvim isto vrijedi i za sve ostale metode, koje imaju

- red konvergencije $p > 1$, u okolini jednostruke nultočke α.

Ako se metoda "uspori" u okolini višestruke nultočke,

- treba metodu primijeniti na $u = f/f'$, umjesto na f.

Na primjer, za višestruke nultočke, metodu sekante treba primijeniti na funkciju u,

$$x_{n+1} = x_n - u(x_n) \frac{x_n - x_{n-1}}{u(x_n) - u(x_{n-1})}.$$

Dodatna "cijena" = računanje još jedne derivacije više.

Na primjer, u Newtonovoj metodi, za u' treba računati i f'', a u metodi sekante, za u treba računati i f'.

Numerički red konvergencije

U praksi se može sasvim dobro

- numerički procijeniti red konvergencije iterativne metode.

Kako se to radi?

Red konvergencije niza iteracija $(x_n, n \in \mathbb{N}_0)$, koji konvergira prema nultočki α, je najveći eksponent p, uz $p \geq 1$, za koji vrijedi

$$\lim_{n\to\infty} \frac{|\alpha - x_{n-1}|}{|\alpha - x_{n-2}|^p} = c > 0.$$

Ovdje je x_n niz iteracija generiran nekom metodom, uz neki start dovoljno blizu nultočke (ranije su indeksi bili n i $n-1$).

Ovako dobiveni p i c su “teorijske” vrijednosti koje vrijede asimptotski — na limesu $n \to \infty$.

Numerički red konvergencije

Definicijsku relaciju ne možemo direktno iskoristiti za računanje p i c, jer ne znamo nultočku α.

Praktični pogled. Ako su iteracije x_k dovoljno blizu α, onda limes možemo "zaboraviti", pa je

$$|\alpha - x_k| \approx c|\alpha - x_{k-1}|^p,$$

za dovoljno velike k. Umjesto α, uzmemo aproksimaciju za α!

Dakle, ako smo dovoljno blizu nultočke, onda možemo uzeti da je $\alpha \approx x_n$, a za k uzmemo $k = n - 1, n - 2$. Onda vrijedi

$$\begin{aligned} |x_n - x_{n-1}| &= c_n |x_n - x_{n-2}|^{p_n}, \\ |x_n - x_{n-2}| &= c_n |x_n - x_{n-3}|^{p_n}, \end{aligned}$$

s tim da očekujemo da je $p_n \approx p$ i $c_n \approx c$, za dovoljno velike n.

Numerički red konvergencije

Dijeljenjem dobivamo

$$\frac{|x_n - x_{n-1}|}{|x_n - x_{n-2}|} = \left(\frac{|x_n - x_{n-2}|}{|x_n - x_{n-3}|} \right)^{p_n},$$

pa logaritmiranjem izlazi tzv. numerički red konvergencije p_n

$$p_n = \frac{\log(\,|x_n - x_{n-1}|/|x_n - x_{n-2}|\,)}{\log(\,|x_n - x_{n-2}|/|x_n - x_{n-3}|\,)}.$$

Nakon toga, c_n možemo dobiti (recimo) iz prve relacije, kao

$$c_n = \frac{|x_n - x_{n-1}|}{|x_n - x_{n-2}|^{p_n}}.$$

Ovaj račun možemo provesti tek za $n \geq 3$, a vrijednosti p_n i c_n ovise o n — pa treba pratiti njihovo ponašanje kroz iteracije.

Jednostavni primjer

Primjer. Za uspoređivanje metoda za nalaženje nultočaka, izračunajmo $\sqrt[3]{1.5}$. Problem možemo interpretirati kao traženje realne, pozitivne nultočke funkcije (polinoma)

$$f(x) = x^3 - 1.5.$$

Jednostavna lokacija nultočke je $\alpha \in [1, 2]$. Iz neprekidnosti funkcije f i

$$f(1) = -0.5 < 0, \quad f(2) = 6.5 > 0,$$

znamo da sigurno postoji nultočka $\alpha \in [1, 2]$. To je i jedina realna nultočka funkcije f, jer f strogo raste na $\langle -\infty, 0\rangle$ (tamo je $f < 0$) i na $\langle 0, \infty\rangle$.

Svo računanje provedeno je u `extended` tipu ($u \approx 5.4 \cdot 10^{-20}$).

Metoda bisekcije

Pokažimo kako se ponaša metoda bisekcije ako je $[a, b] = [1, 2]$, a tražena točnost je

- $\varepsilon = 10^{-8}$,
- $\varepsilon = 10^{-18}$.

Na sljedeće dvije stranice, sa z_n je označena veličina

$$z_n := f(a_n) \cdot f(x_n).$$

Metoda bisekcije, točnost $\varepsilon = 10^{-8}$ (početak)

n	a_n	b_n	x_n	$f(x_n)$	z_n
0	1.000000000	2.000000000	1.500000000	1.875000000	< 0
1	1.000000000	1.500000000	1.250000000	0.453125000	< 0
2	1.000000000	1.250000000	1.125000000	−0.076171875	> 0
3	1.125000000	1.250000000	1.187500000	0.174560547	< 0
4	1.125000000	1.187500000	1.156250000	0.045806885	< 0
5	1.125000000	1.156250000	1.140625000	−0.016017914	> 0
6	1.140625000	1.156250000	1.148437500	0.014684200	< 0
7	1.140625000	1.148437500	1.144531250	−0.000719249	> 0
8	1.144531250	1.148437500	1.146484375	0.006969355	< 0
9	1.144531250	1.146484375	1.145507813	0.003121776	< 0
10	1.144531250	1.145507813	1.145019531	0.001200444	< 0
11	1.144531250	1.145019531	1.144775391	0.000240393	< 0
12	1.144531250	1.144775391	1.144653320	−0.000239479	> 0
13	1.144653320	1.144775391	1.144714355	0.000000444	< 0

Metoda bisekcije, točnost $\varepsilon = 10^{-8}$ (nastavak)

n	a_n	b_n	x_n	$f(x_n)$	z_n
14	1.144653320	1.144714355	1.144683838	-0.000119521	> 0
15	1.144683838	1.144714355	1.144699097	-0.000059539	> 0
16	1.144699097	1.144714355	1.144706726	-0.000029548	> 0
17	1.144706726	1.144714355	1.144710541	-0.000014552	> 0
18	1.144710541	1.144714355	1.144712448	-0.000007054	> 0
19	1.144712448	1.144714355	1.144713402	-0.000003305	> 0
20	1.144713402	1.144714355	1.144713879	-0.000001431	> 0
21	1.144713879	1.144714355	1.144714117	-0.000000493	> 0
22	1.144714117	1.144714355	1.144714236	-0.000000025	> 0
23	1.144714236	1.144714355	1.144714296	0.000000210	< 0
24	1.144714236	1.144714296	1.144714266	0.000000092	< 0
25	1.144714236	1.144714266	1.144714251	0.000000034	< 0
26	1.144714236	1.144714251	1.144714244	0.000000005	< 0
27	1.144714236	1.144714244	1.144714240	-0.000000010	

Metoda bisekcije, točnost $\varepsilon = 10^{-8}, 10^{-18}$

Izračunata rješenja su (uz ispis na 18 znamenki):

- za točnost 10^{-8} , rješenje je $x_{27} = 1.14471423998475075$,
- za točnost 10^{-18}, rješenje je $x_{60} = 1.14471424255333187$.

Iz ovih rezultata vidi se

- spora konvergencija metode bisekcije — broj vodećih nula u $f(x_n)$ se, uglavnom, linearno povećava.

Ponegdje, kao u x_{13}, dolazi do čudnog povećanja broja vodećih nula u $f(x_n)$. Objašnjenje:

- Slučajno smo "pogodili" blizu nultočke, iako je duljina intervala još uvijek prevelika za zaustavljanje iteracija.

Uočite broj iteracija potreban za odgovarajuću točnost!

Newtonova metoda, točnost $\varepsilon = 10^{-18}$

Nije teško pokazati da će Newtonova metoda na $[1, 2]$ sigurno konvergirati, ako krenemo sa strmijeg ruba (to je $x_0 = 2$), jer su f' i f'' fiksnog znaka na $[1, 2]$.

n	x_n	$f(x_n)$	korekcija
0	2.000000000000000	6.500000000000000	0.541666666666667
1	1.458333333333333	1.601490162037037	0.251009070294785
2	1.207324263038549	0.259834330619978	0.059419284771986
3	1.147904978266562	0.012578134527712	0.003181874908824
4	1.144723103357739	0.000034833084971	0.000008860735819
5	1.144714242621919	0.000000000269625	0.000000000068587
6	1.144714242553332	0.000000000000000	0.000000000000000
7	1.144714242553332	0.000000000000000	

Izračunata nultočka je $x_7 = 1.1447142425533187$ (sve točno).

Newtonova metoda, kvadratna konvergencija

Gledamo li korekcije napisane u znanstvenoj notaciji, vidimo područje kvadratne konvergencije — gdje se broj točnih znamenki u x_n, u svakom koraku, udvostručava.

n	x_n	$f(x_n)$	korekcija
0	2.0000000000000000	6.50000000000E+00	5.416666667E−01
1	1.4583333333333333	1.60149016204E+00	2.510090703E−01
2	1.20732426303854875	2.59834330620E−01	5.941928477E−02
3	1.14790497826656245	1.25781345277E−02	3.181874909E−03
4	1.14472310335773870	3.48330849709E−05	8.860735819E−06
5	1.14471424262191933	2.69625000386E−10	6.858746179E−11
6	1.14471424255333187	1.08420217249E−19	2.758003708E−20
7	1.14471424255333187	1.08420217249E−19	

Isti smo zaključak mogli vidjeti i bez znanstvene notacije — pogledajte kako se povećava broj vodećih nula u korekciji.

Newtonova m., numerički red konvergencije

Prema formuli s početka ovog odjeljka, možemo izračunati i numerički red konvergencije p_n za Newtonovu metodu.

n	x_n	p_n	c_n
0	2.0000000000000000	—	—
1	1.4583333333333333	—	—
2	1.2073242630385487	—	—
3	1.1479049782665624	1.63738	4.03440E−01
4	1.1447231033577387	1.84894	5.34225E−01
5	1.1447142426219193	1.97750	7.64767E−01
6	1.1447142425533318	1.99937	8.67206E−01
7	1.1447142425533318	—	—

Završne vrijednosti p_5 i p_6 su vrlo blizu očekivanog teorijskog reda konvergencije $p = 2$!

Metoda sekante, točnost $\varepsilon = 10^{-18}$

Za metodu sekante potrebne su dvije startne točke — to su $x_0 = 2$ i $x_1 = 1.5$. Izračunata nultočka x_{10} ima sve znamenke jednake aproksimaciji x_7 iz Newtonove metode.

n	x_n	$f(x_n)$	korekcija
0	2.000000000000000	6.500000000000000	
1	1.500000000000000	1.875000000000000	0.202702702702703
2	1.297297297297297	0.683325765502537	0.116233090792778
3	1.181064206504520	0.147481413791545	0.031991044609771
4	1.149073161894749	0.017200732670849	0.004223722208545
5	1.144849439686204	0.000531537856385	0.000134683664909
6	1.144714756021295	0.000002018501023	0.000000513407324
7	1.144714242613971	0.000000000238377	0.000000000060639
8	1.144714242553332	0.000000000000000	0.000000000000000
9	1.144714242553332	0.000000000000000	0.000000000000000
10	1.144714242553332	0.000000000000000	

Metoda sekante, numerički red konvergencije

Red konvergencije metode sekante je $p = (1 + \sqrt{5})/2 \approx 1.6180$, a numerički red konvergencije p_n ga dobro aproksimira.

n	x_n	p_n	c_n
0	2.00000000000000000	—	—
1	1.50000000000000000	—	—
2	1.29729729729729730	—	—
3	1.18106420650451962	1.07039	3.94966E−01
4	1.14907316189474910	1.77904	9.54986E−01
5	1.14484943968620389	1.49493	6.02649E−01
6	1.14471475602129474	1.63923	9.97717E−01
7	1.14471424261397050	1.60467	8.29830E−01
8	1.14471424255333190	1.62274	9.74858E−01
9	1.14471424255333187	1.61618	8.86501E−01
10	1.14471424255333187	—	—

Primjer (ne)konvergencije Newtonove metode

Primjer. Jedina nultočka funkcije

$$f(x) = \mathrm{arctg}(x)$$

je $x = 0$. Međutim, Newtonova metoda ne konvergira iz svake startne točke x_0.

- Sigurnu konvergenciju (po ranijem teoremu) ne možemo osigurati, jer f'' mijenja znak baš u nultočki (infleksija).

Naći ćemo točku β za koju vrijedi:

$$\begin{cases} |x_0| < |\beta| \;\Rightarrow \text{Newtonova metoda sa startom } x_0 \text{ konvergira,} \\ |x_0| > |\beta| \;\Rightarrow \text{Newtonova metoda sa startom } x_0 \text{ divergira,} \\ |x_0| = |\beta| \;\Rightarrow \text{Newtonova metoda sa startom } x_0 \text{ “ciklira”.} \end{cases}$$

Primjer (ne)konvergencije Newtonove metode

Kako ćemo naći točku "cikliranja" (ili "kruženja") β?

Funkcija $f(x) = \text{arctg} x$ je neparna, pa je dovoljno da

- tangenta na graf funkcije f u točki $(\beta, f(\beta))$ presiječe os x u točki $-\beta$ — dobijemo neparnu simetriju oko nule.

Jednadžba tangente na arctg u točki β je

$$y - \text{arctg}\beta = \frac{1}{1+\beta^2}\,(x-\beta),$$

pa će tangenta sjeći os x u $-\beta$ (tada je $y = 0$), ako je

$$\text{arctg}\beta = \frac{2\beta}{1+\beta^2}.$$

Dobili smo nelinearnu jednadžbu za β, koju treba riješiti.

Primjer (ne)konvergencije Newtonove metode

Zbog neparnosti, postoje dva rješenja, suprotnih predznaka. Možemo ih izračunati, recimo, metodom bisekcije i dobijemo

$$\beta = \pm 1.3917452002707348 9.$$

Pokažimo ponašanje Newtonove metode ako za startnu točku uzmemo, redom,

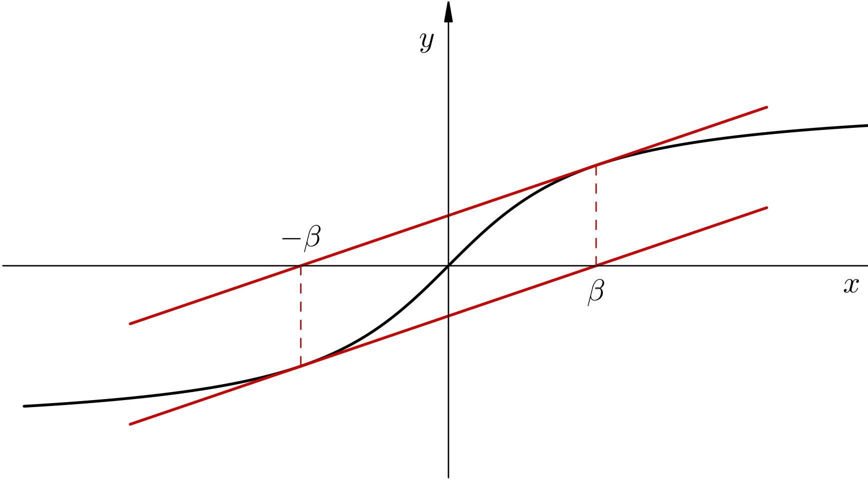
$$x_0 = \beta, \quad x_0 = 1, \quad x_0 = 1.5,$$

a zaustavljamo se

- ako postignemo točnost 10^{-18} za nađenu nultočku,
- ili nakon najviše 10 iteracija.

Primjer cikliranja Newtonove metode

Za $x_0 = \beta$ graf je



Primjer cikliranja Newtonove metode

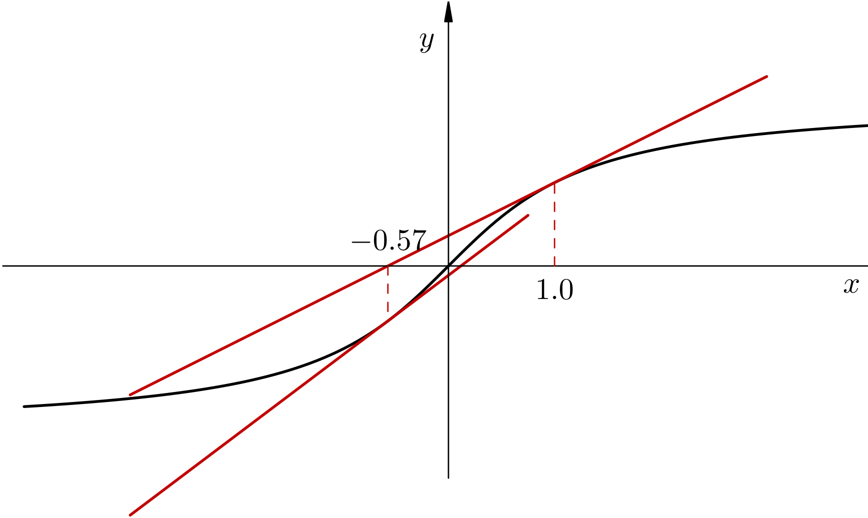
Točnost nije postignuta nakon 10 iteracija.

n	x_n	$f(x_n)$	korekcija
0	1.391745200270735	0.947747133516990	2.783490400541470
1	−1.391745200270735	−0.947747133516990	−2.783490400541470
2	1.391745200270735	0.947747133516990	2.783490400541469
3	−1.391745200270734	−0.947747133516990	−2.783490400541468
4	1.391745200270734	0.947747133516990	2.783490400541466
5	−1.391745200270732	−0.947747133516989	−2.783490400541459
6	1.391745200270727	0.947747133516988	2.783490400541440
7	−1.391745200270713	−0.947747133516983	−2.783490400541391
8	1.391745200270678	0.947747133516971	2.783490400541262
9	−1.391745200270584	−0.947747133516939	−2.783490400540922
10	1.391745200270337	0.947747133516939	

Zbog malih grešaka zaokruživanja, metoda bi konvergirala.

Primjer konvergencije Newtonove metode

Za $x_0 = 1$ graf je



Primjer konvergencije Newtonove metode

Zadana točnost 10^{-18} se postiže nakon 6 iteracija.

n	x_n	$f(x_n)$	korekcija
0	1.000000000000000	0.785398163397448	1.570796326794897
1	−0.570796326794897	−0.518669369255017	−0.687656230793810
2	0.116859903998913	0.116332265113896	0.117920926115958
3	−0.001061022117045	−0.001061021718890	−0.001061022913354
4	0.000000000796310	0.000000000796310	0.000000000796310
5	−0.000000000000000	−0.000000000000000	−0.000000000000000
6	0.000000000000000	−0.000000000000000	

Metoda kubno konvergira (zbog $f''(0) = 0$), ali ne konvergira monotono prema nultočki $\alpha = 0$.

Newton kvg., numerički red konvergencije

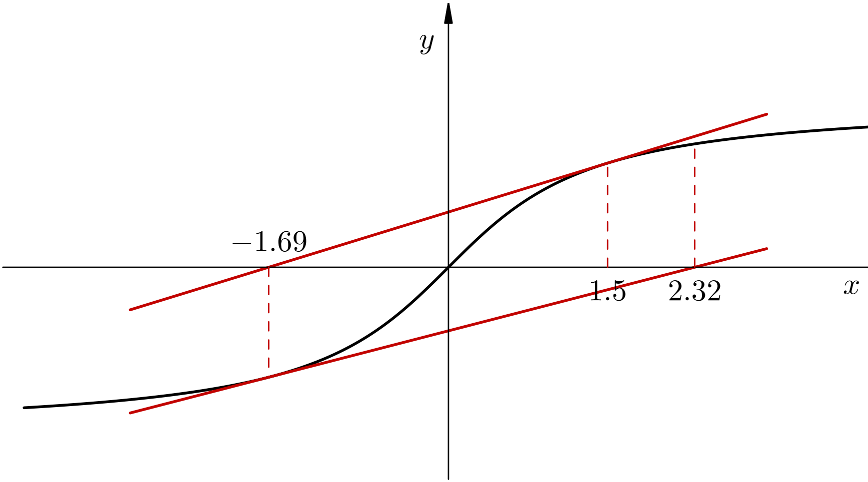
Numerički red konvergencije pokazuje da metoda zaista kubno konvergira prema $\alpha = 0$.

n	x_n	p_n	c_n
0	1.00000000000000000	—	—
1	-0.57079632679489662	—	—
2	0.11685990399891305	—	—
3	-0.00106102211704472	4.02288	1.13367E+00
4	0.00000000079630960	2.97361	6.28237E$-$01
5	-0.00000000000000000	2.99942	6.64032E$-$01
6	0.00000000000000000	2.99655	6.51104E$-$01

Na kraju iteracija dolazi do kraćenja, zbog nemonotonosti. Zato je p_6 malo manje točan od p_5.

Primjer divergencije Newtonove metode

Za $x_0 = 1.5$ graf je



Primjer divergencije Newtonove metode

Metoda kvadratno divergira, ali $f(x)$ “konvergira” prema $\pm\frac{\pi}{2}$.

n	x_n	$f(x_n)$	korekcija
0	1.5000000000E+000	0.9827937232473	3.1940796006E+000
1	−1.6940796006E+000	−1.0375463591379	−4.0152065620E+000
2	2.3211269614E+000	1.1640020424220	7.4352147982E+000
3	−5.1140878368E+000	−1.3776945287028	−3.7409771751E+001
4	3.2295683914E+001	1.5398423269080	1.6076126347E+003
5	−1.5753169508E+003	−1.5701615339901	−3.8965513247E+006
6	3.8949760078E+006	1.5707960700539	2.3830292869E+013
7	−2.3830288974E+013	−1.5707963267949	−8.9202801611E+026
8	8.9202801611E+026	1.5707963267949	1.2499045994E+054
9	−1.2499045994E+054	−1.5707963267949	−2.4539946375E+108
10	2.4539946375E+108	1.5707963267949	

Rješavanje nelinearnih jednadžbi

Derivabilne iteracijske funkcije i metode višeg reda

Newtonova metoda za višestruke nultočke

Primjeri za jednostruke nultočke

Primjeri za višestruke nultočke

Rješavanje sustava nelinearnih jednadžbi

Newtonova metoda

Metoda sekante — Broydenova metoda

Primjer rješavanja sustava nelinearnih jednadžbi

Newtonova metoda i višestruke nultočke

Primjer. Funkcija (polinom)

$$f(x) = x^3 - 5.56x^2 + 9.1389x - 4.68999$$

ima dvostruku nultočku u $x = 1.23$ (treća nultočka je 3.1).

Pokažimo, redom, ponašanje

- obične Newtonove metode,
- Newtonove metode za dvostruku nultočku — stavljen je faktor $m = 2$ za korekciju,
- obične Newtonove metode, ali za funkciju $u = f/f'$.

Startna točka je $x_0 = 1.5$, a tražena točnost je $\varepsilon = 10^{-15}$.

- Točnost za nultočku je namjerno "slabija" nego prije.

Newtonova metoda, $\varepsilon = 10^{-15}$ (početak)

Pažljivo promatrajte kako se ponaša $f(x_n)$ i korekcija.

n	x_n	$f(x_n)$	korekcija
0	1.500000000000000	−0.116640000000000	0.147440273037543
1	1.352559726962457	−0.026248102309329	0.063506947954921
2	1.289052779007536	−0.006315190760541	0.030015778457640
3	1.259037000549896	−0.001552203168203	0.014633908572506
4	1.244403091977390	−0.000384941831542	0.007229603981965
5	1.237173487995426	−0.000095859059122	0.003593663348690
6	1.233579824646736	−0.000023918444247	0.001791630511788
7	1.231788194134948	−0.000005973865556	0.000894525173287
8	1.230893668961661	−0.000001492750955	0.000446941328040
9	1.230446727633621	−0.000000373098481	0.000223390506264
10	1.230223337127356	−0.000000093263474	0.000111675233252
11	1.230111661894104	−0.000000023314476	0.000055832614098

Newtonova metoda, $\varepsilon = 10^{-15}$ (nastavak)

n	x_n	$f(x_n)$	korekcija
12	1.230055829280006	−0.000000005828445	0.000027915056720
13	1.230027914223287	−0.000000001457089	0.000013957215816
14	1.230013957007471	−0.000000000364270	0.000006978529777
15	1.230006978477694	−0.000000000091067	0.000003489245354
16	1.230003489232340	−0.000000000022767	0.000001744617791
17	1.230001744614549	−0.000000000005692	0.000000872307704
18	1.230000872306845	−0.000000000001423	0.000000436153629
19	1.230000436153216	−0.000000000000356	0.000000218076376
20	1.230000218076841	−0.000000000000089	0.000000109038186
21	1.230000109038655	−0.000000000000022	0.000000054517907
22	1.230000054520748	−0.000000000000006	0.000000027259838
23	1.230000027260910	−0.000000000000001	0.000000013624333
24	1.230000013636577	−0.000000000000000	0.000000006828241
25	1.230000006808336	−0.000000000000000	0.000000003389306

Newtonova metoda, $\varepsilon = 10^{-15}$ (kraj)

n	x_n	$f(x_n)$	korekcija
26	1.230000003419030	-0.000000000000000	0.000000001729681
27	1.230000001689349	-0.000000000000000	0.000000000823684
28	1.230000000865665	-0.000000000000000	0.000000000401855
29	1.230000000463810	0.000000000000000	-0.000000000000000
30	1.230000000463810	0.000000000000000	

Korekcija linearno teži u 0, puno sporije nego $f(x)$. Razlog:

- Oko višestruke nultočke, funkcijska vrijednost je relativno mala, čak i kad smo relativno daleko od nultočke, tj.
- u okolini višestruke nultočke α, graf funkcije f se bolje "priljubi" uz os x, nego kad je nultočka jednostruka.

Modificirana Newtonova metoda, $m = 2$

Modificirana metoda pokazuje očitu kvadratnu konvergenciju.

n	x_n	$f(x_n)$	korekcija
0	1.500000000000000	−0.116640000000000	0.294880546075085
1	1.205119453924915	−0.001173009833879	−0.024718265674538
2	1.229837719599453	−0.000000049250590	−0.000162273360038
3	1.229999992959491	−0.000000000000000	−0.000000007049176
4	1.230000000008667	0.000000000000000	−0.000000026758865
5	1.230000026767532	−0.000000000000001	0.000000026771877
6	1.229999999995655	0.000000000000000	0.000000000000000
7	1.229999999995655	0.000000000000000	

Da smo pogriješili m — dobili bismo linearnu konvergenciju!

Newtonova metoda za funkciju $u = f/f'$

Obična Newtonova metoda za funkciju $u = f/f'$, također, pokazuje očitu kvadratnu konvergenciju prema jednostrukoj nultočki funkcije u.

n	x_n	$f(x_n)$	korekcija
0	1.500000000000000	0.147440273037543	0.243748194650388
1	1.256251805349612	0.013220017840880	0.026062273270891
2	1.230189532078721	0.000094770842551	0.000189522471849
3	1.230000009606872	0.000000004803964	0.000000009608985
4	1.229999999997887	0.000000000000000	0.000000000000000
5	1.229999999997887	0.000000000000000	

Cijena:

- računanje vrijednosti druge derivacije funkcije f za u'.

Rješavanje nelinearnih jednadžbi

- Derivabilne iteracijske funkcije i metode višeg reda
- Newtonova metoda za višestruke nultočke
- Primjeri za jednostruke nultočke
- Primjeri za višestruke nultočke

Rješavanje sustava nelinearnih jednadžbi

- Newtonova metoda
- Metoda sekante — Broydenova metoda
- Primjer rješavanja sustava nelinearnih jednadžbi

Formulacija problema

Zadana je funkcija
$$f : \mathbb{R}^n \to \mathbb{R}^m.$$

Radi jednostavnosti, pretpostavljamo da je f definirana na cijelom prostoru $\mathbb{R}^n$.

Kao i prije, tražimo (jednu ili sve) točke $x \in \mathbb{R}^n$ za koje je

$$f(x) = 0 \in \mathbb{R}^m.$$

Najjednostavniji primjer je tzv. linearna (ili afina) funkcija f

$$f(x) = Ax - b, \quad x \in \mathbb{R}^n$$

gdje je $A \in \mathbb{R}^{m \times n}$ zadana pravokutna matrica s m redaka i n stupaca, a $b \in \mathbb{R}^m$ je zadani vektor.
Nultočke ove funkcije su rješenja linearnog sustava $Ax = b$.

Pojednostavljenja — dodatne pretpostavke

Raspisom vektorske jednadžbe $f(x) = 0$ po komponentama u prostoru $\mathbb{R}^m$, dobivamo sustav s m jednadžbi i n nepoznanica

$$f_i(x_1, x_2, \ldots, x_n) = 0, \quad i = 1, \ldots, m.$$

Dodatne pretpostavke u nastavku:

- funkcija f je, općenito, nelinearna,
- broj jednadžbi m jednak je broju nepoznanica n,

tj. funkcija je $f : \mathbb{R}^n \to \mathbb{R}^n$.

Osim toga, pretpostavljamo

- da f ima samo izolirane nultočke $\alpha \in \mathbb{R}^n$, i
- da je f dovoljno glatka — na cijelom $\mathbb{R}^n$, ili barem u nekoj okolini nultočke.

O generalizaciji metoda iz jedne dimenzije

Očita ideja za rješavanje sustava nelinearnih jednadžbi je

- generalizacija metoda za rješavanje jedne jednadžbe.

Problem: U više dimenzija nema uspoređivanja funkcijskih vrijednosti (vektora f) — osim po komponentama f_i.

Zato se "jednostavne" metode — poput bisekcije,

- teško generaliziraju, a i složenost postaje problem (imamo 2^n vrhova "kocke" i n komponentnih funkcija u svakom vrhu). Probajte zamisliti u $\mathbb{R}^2$!

S druge strane, većina "iteracijskih" funkcija se

- relativno jednostavno generalizira — iz "skalarnih", u "vektorske" ili "matrično–vektorske".

Ilustracija toga $=$ Newtonova metoda i metoda sekante.

Newtonova metoda u $\mathbb{R}^n$

Newtonovu metodu dobivamo linearizacijom funkcije f oko trenutne aproksimacije $x^{(k)} \in \mathbb{R}^n$ — slično kao za $n = 1$, samo su oznake malo drugačije (indeksi služe za komponente).

Izaberemo neku početnu točku $x^{(0)} \in \mathbb{R}^n$. Onda, induktivno, generiramo niz aproksimacija $x^{(k)}$, za $k \geq 0$, na sljedeći način.

Oko trenutne točke $x^{(k)}$, svaku komponentnu funkciju f_i

- aproksimiramo linearnim dijelom Taylorovog razvoja,

tj. odbacujemo sve članove nakon linearnih. Dobivamo

$$f_i(x) \approx f_i(x^{(k)}) + \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x^{(k)})(x_j - x_j^{(k)}), \quad i = 1, \ldots, n.$$

Newtonova metoda u $\mathbb{R}^n$

Iskoristimo Jacobijevu matricu $J_f(x)$, koja sadrži parcijalne derivacije svih komponentnih funkcija po svim varijablama

$$[J_f(x)]_{ij} = \frac{\partial f_i}{\partial x_j}, \quad i = 1, \ldots, n, \; j = 1, \ldots, n,$$

ili

$$J_f(x) = \frac{\partial f}{\partial x}(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Newtonova metoda u $\mathbb{R}^n$

Linearizirani aproksimaciju za $f(x)$ oko točke $x^{(k)}$ možemo zapisati u matrično–vektorskom obliku

$$f(x) \approx f(x^{(k)}) + J_f(x^{(k)})(x - x^{(k)}).$$

Novu aproksimaciju $x^{(k+1)}$ za nultočku dobivamo iz zahtjeva da je

$$f(x^{(k+1)}) \approx f(x^{(k)}) + J_f(x^{(k)})(x^{(k+1)} - x^{(k)}) = 0.$$

Ako je $J_f(x^{(k)})$ regularna matrica, onda je pripadna korekcija

$$d^{(k)} := x^{(k+1)} - x^{(k)}$$

jedinstveno rješenje linearnog sustava jednadžbi

$$J_f(x^{(k)})d^{(k)} = -f(x^{(k)}).$$

Newtonova metoda u $\mathbb{R}^n$

Za korekciju $d^{(k)}$ onda vrijedi (ali se ne računa ovako)

$$d^{(k)} = -[J_f(x^{(k)})]^{-1} f(x^{(k)}),$$

pa izlazi standardni zapis za iteracije u Newtonovoj metodi

$$x^{(k+1)} = x^{(k)} - [J_f(x^{(k)})]^{-1} f(x^{(k)}), \quad k = 0, 1, \ldots$$

U usporedbi sa "skalarnim" oblikom Newtonove metode,

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})},$$

- množenje inverzom Jacobijeve matrice $J_f(x)$, i to slijeva, je "zamjena" za dijeljenje derivacijom funkcije $f'(x)$.

Lokalna konvergencija Newtonove metode u $\mathbb{R}^n$

Uočite da se rješavanje jednog nelinearnog sustava jednadžbi svodi na niz rješavanja linearnih sustava.

Što se konvergencije tiče, vrijedi potpuni analogon rezultata o lokalnoj konvergenciji i brzini konvergencije u jednoj dimenziji.

Teorem. Neka je α nultočka funkcije $f : \mathbb{R}^n \to \mathbb{R}^n$ i neka je f klase C^2 u okolini od α.

Ako je $J_f(\alpha)$ regularna matrica (to je analogon jednostruke nultočke), onda postoji okolina nultočke α, takva da

- za bilo koju početnu točku $x^{(0)}$ iz te okoline,
- niz iteracija generiran Newtonovom metodom, konvergira prema α, i konvergencija je (barem) kvadratna,

tj. vrijedi

$$\|\alpha - x^{(k+1)}\| \in O\left(\|\alpha - x^{(k)}\|^2\right).$$

■

Metoda sekante u $\mathbb{R}^n$

Metodu sekante možemo dobiti iz Newtonove metode, slično kao u jednoj dimenziji, ali ne na posve analogan način.

Ideja je, ponovo, da parcijalne derivacije u Jacobijevoj matrici $J_f(x)$ zamijenimo sa izrazima koji uključuju samo funkcijske vrijednosti.

Najjednostavniji način da ih aproksimiramo podijeljenim razlikama, i definiramo matricu

$$\Delta f(x) = [\Delta_1 f, \ldots, \Delta_n f], \quad \Delta_j f(x) = \frac{f(x + h_j e_j) - f(x)}{h_j},$$

za dovoljno male h_j. Dobar izbor koraka h_j može biti problem, jer

- ako je prevelik, konvergencija može biti jako spora ili potpuno izgubljena,
- ako je premalen, može doći do katastrofalnog kraćenja.

Metoda sekante u $\mathbb{R}^n$ — Broydenova metoda

U slučaju da se funkcija f komplicirano računa, dodatna n izvrednjavanja funkcije u svakoj iteraciji mogu biti preskupa.

Tada pokušavamo zamijeniti $Jf(x^{(k)})$ sa nekom matricom B_k koja je još jednostavnija od $\Delta f(x^{(k)})$. Pogodne matrice mogu se dobiti koristeći rezultat koji je dao Broyden.

Teorem. Neka su A i B proizvoljne $n \times n$ matrice, neka je $b \in \mathbb{R}^n$, i neka je $F : \mathbb{R}^n \to \mathbb{R}^n$ afino preslikavanje $F(u) = Au + b$. Pretpostavimo da su $x, x' \in \mathbb{R}^n$ različiti vektori, i definirajmo

$$p = x' - x, \quad q = F(x') - F(x) = Ap.$$

Tada $n \times n$ matrica B' dana sa

$$B' = B + \frac{1}{p^T p}(q - Bp)p^T,$$

zadovoljava

$$\|B' - A\|_2 \leq \|B - A\|_2, \quad \text{i} \quad B'p = Ap = q.$$

Metoda sekante u $\mathbb{R}^n$ — Broydenova metoda

Dokaz. Iz definicije matrice B' slijedi

$$(B'-A)p = Bp+\frac{p^Tp}{p^Tp}(q-Bp)-Ap = Bp+q-Bp-Ap = q-Ap = 0.$$

Svaki vektor $u \in \mathbb{R}^n$ koji ima jediničnu normu $\|u\|_2 = 1$ ima ortogonalnu dekompoziciju oblika

$$u = \alpha p + v, \quad v^T p = 0, \quad \|v\|_2 \leq 1, \quad \alpha \in \mathbb{R}.$$

Tada vrijedi

$$\begin{aligned} \|(B' - A)u\|_2 &= \|\alpha(B' - A)p + (B' - A)v\|_2 = \|(B' - A)v\|_2 \\ &= \left\|Bv + \frac{p^T v}{p^T p}(q - Bp) - Av\right\|_2 = \|(B - A)v\|_2 \\ &\leq \|B - A\|_2 \|v\|_2 \leq \|B - A\|_2. \end{aligned}$$

Prema tome je

$$\|B' - A\|_2 = \max_{\|u\|_2=1} \|(B' - A)u\|_2 \leq \|B - A\|_2.$$

Metoda sekante u $\mathbb{R}^n$ — Broydenova metoda

Rezultat ovog teorema pokazuje

- da se Jacobijeva matrica $JF(x) = A$ afine funkcije F može aproksimirati matricom B' barem tako dobro kao što je aproksimirana matricom B, i
- B' i $JF(x)$ oboje preslikavaju p u isti vektor.

Budući da znamo da se diferencijabilna nelinearna funkcija $f : \mathbb{R}^n \to \mathbb{R}^n$ u okolini nultočke α može aproksimirati afinom funkcijom (preko Jacobijeve matrice), ovaj pristup možemo primijeniti i na nelinearne funkcije.

Metoda sekante u $\mathbb{R}^n$ — Broydenova metoda

Time dobivamo iteracije oblika

$$p^{(k-1)} = x^{(k)} - x^{(k-1)}, \qquad q^{(k-1)} = f(x^{(k)}) - f(x^{(k-1)}),$$

$$B^{(k)} = B^{(k-1)} + \frac{1}{(p^{(k-1)})^T p^{(k-1)}} (q^{(k)} - B^{(k-1)} p^{(k-1)})(p^{(k-1)})^T,$$

$$d^{(k)} = (B^{(k)})^{-1} f(x_k), \qquad x^{(k+1)} = x^{(k)} - d^{(k)}.$$

Pogodna početna matrica $B^{(1)}$ može se uzeti kao

$$B^{(1)} = \Delta f(x^{(1)}),$$

za pogodno odabrane h_j. Primijetimo da na vektor

$$x^{(0)} = [x_1^{(1)} - h_1, \ldots, x_n^{(1)} - h_n]^T,$$

možemo gledati kao na drugi incijalni vektor, kao i kod jednodimenzionalne metode sekante, s time da u $\Delta f(x^{(1)})$ gledamo podijeljene razlike u natrag.

Primjer rješavanja sustava nelinearnih jednadžbi

Pretpostavimo da želimo riješiti sljedeći sustav nelinearnih jednadžbi

$$\begin{aligned} x_1^6 - 5x_1^2 x_2^2 + 136 &= 0 \\ x_2^4 - 3x_1^4 x_2 + 80 &= 0 \end{aligned}$$

Rješavat ćemo ga

- Newtonovom metodom, sa $x^{(0)} = [1, 2]^T$
- Broydenovom metodom, sa $x^{(0)} = [1, 2]^T$, $x^{(1)} = [1.1, 2.1]^T$

Za obje metode kriterij zaustavljanja je kad niz aproksimacija postane dovoljno stacionaran, tj.

$$\max_{i=1,\ldots,n} \frac{|x_i^{(k+1)} - x_i^{(k)}|}{|x_i^{(k)}|} \leq \epsilon_M = 2.22 \cdot 10^{-16}.$$

Rezultat Newtonove metode

Newtonova metoda je iskonvergirala u 12 koraka.

k	$x_1^{(k)}$	$x_2^{(k)}$	$\|f(x^{(k)})\|_2$
0	1.000000000000000	2.000000000000000	147.6109752017105
1	4.542291950886767	1.828103683492497	8863.030301106330
2	3.775877835532193	1.294999404576783	2999.012537206996
3	3.123284551508469	1.027257121946039	1034.769150951805
4	2.545619794676695	1.044378193958253	376.1688218972496
5	1.995630009701346	1.565902932840357	150.7780703407369
6	1.697787817019694	3.303311264075423	116.7608671714409
7	2.073054677755745	2.934926464099915	31.42852345303920
8	2.081716791664326	3.178510792324494	3.361565521966347
9	2.088346798860385	3.168604854289832	0.016918243023167
10	2.088378995687454	3.168732957946968	0.000000716918630
11	2.088378995520735	3.168732953136702	0.000000000000115
12	2.088378995520735	3.168732953136702	0.000000000000043

Na početku iteracija aproksimacija se čak i pogoršala, ali nakon 8-e iteracije dolazi do brze konvergencije.

Rezultat Broydenove metode

Broydenova metoda je iskonvergirala u 37 koraka.

k	$x_1^{(k)}$	$x_2^{(k)}$	$\|f(x^{(k)})\|_2$
0	1.000000000000000	2.000000000000000	147.6109752017105
1	1.100000000000000	2.100000000000000	143.1140899112964
2	4.058394377890478	1.976171198459055	4541.941394741537
3	0.590746087877206	−4.995248773233302	710.5002298559424
4	1.258907674767459	5.728619994002871	1120.247923357512
5	−1.364017538440051	−29.697199542259938	778219.5036503847
6	1.340011467997714	5.771613411370410	1144.686893149320
7	1.447670659500288	5.812926821629026	1164.070771252191
8	0.944811623564473	5.680023928505576	1107.325461214122
9	0.848889970472194	5.713415252184174	1136.826991687819
10	0.785032286474865	5.812482026339942	1215.224818692192
11	0.701339367515614	6.148624262352509	1505.417318914932
12	1.369442293190839	3.091800232523259	148.5203856425290
13	1.053846060301846	4.890109616367509	633.7617724848633
14	1.207406249061289	4.225116637858867	371.8498191008254
15	1.310794120657283	3.858735223866794	267.8558409110436
16	1.335834378875876	3.848719442385333	262.8206785599257
17	1.316116103188205	3.896413226420203	275.5930679349895
18	1.603195996529515	3.252774497727929	128.6125920985108
19	1.507934096215163	3.557932768195186	185.0986140044751

Rezultat Broydenove metode

k	$x_1^{(k)}$	$x_2^{(k)}$	$\|f(x^{(k)})\|_2$
20	1.586921467048508	3.404739081823769	149.7228381683659
21	1.686361813496761	3.260766923784112	114.2074519101122
22	1.714895705179966	3.258111373079208	108.2810351193008
23	1.673194507840681	3.300446647434505	121.1763464004828
24	1.919099364070587	3.079236723061190	46.02392581336729
25	1.813336078922426	3.220737063277526	83.13874384370180
26	1.886081091790109	3.167159860787779	60.43947807269011
27	1.964207001636458	3.130882351624753	36.53586362834241
28	1.964458432114316	3.150142713055224	37.78484835985317
29	1.996856018490091	3.145064958712827	27.91052752166356
30	2.085053324051122	3.152093890708849	2.193495040485576
31	2.082820147524059	3.168952729473998	1.941679259243338
32	2.088545800542688	3.168125410800618	0.134000955914142
33	2.088382888561172	3.168721428461160	0.002749293025919
34	2.088378990156818	3.168732963904817	0.000003086552909
35	2.088378995528641	3.168732953107722	0.000000006378405
36	2.088378995520763	3.168732953136599	0.000000000022529
37	2.088378995520735	3.168732953136702	0.000000000000043
38	2.088378995520735	3.168732953136702	0.000000000000043

Rezultat Broydenove metode

Na početku iteracija aproksimacija jako sporo konvergira i u nekoliko navrata se čak i pogoršala, ali nakon 29-e iteracije dolazi do brze konvergencije.

Na kraju iteracija obaju metoda dobili smo jednake aproksimacije nultočke, do na sve znamenke.