Numerička matematika
12. predavanje

Saša Singer

singer@math.hr

web.math.pmf.unizg.hr/~singer

PMF – Matematički odsjek, Zagreb

Sadržaj predavanja

- Numerička integracija (nastavak):
  - Računanje čvorova i težina Gaussovih formula.
- Rješavanje nelinearnih jednadžbi:
  - Općenito o iterativnim metodama.
  - Brzina konvergencije i pojam reda konvergencije.
  - Metoda raspolavljanja — bisekcije.
  - Regula falsi — metoda pogrešnog položaja.
  - Konstrukcija iterativnih metoda za nultočke.
  - Metoda tangente — Newtonova metoda.
  - Metoda sekante.

Informacije

Trenutno nema bitnih informacija.

Računanje čvorova i težina u Gaussovim formulama

Problem nalaženja Gaussovih formula

Neka je zadana težinska funkcija $w \geq 0$ na intervalu $[a, b]$.

Problem: Za zadani $n \in \mathbb{N}$, treba naći sve "parametre" odgovarajuće Gaussove integracijske formule reda n

$$\int_a^b w(x) f(x)\, dx \approx I_n(f) = \sum_{k=1}^{n} w_k f(x_k).$$

To znači da treba izračunati

- sve čvorove x_k i težine w_k, za $k = 1, \ldots, n$.

Usput, ove parametre treba izračunati maksimalno točno, da osiguramo što točniju numeričku integraciju raznih funkcija f.

Idealno: Izračunati čvorove i težine na punu relativnu točnost aritmetike računala u kojoj radimo — recimo, u tipu `double`.

Sustav jednadžbi iz uvjeta egzaktne integracije

Znamo da Gaussove integracijske formule egzaktno integriraju sve polinome iz $\mathcal{P}_{2n-1}$.

- Možemo izabrati bilo koju bazu u tom prostoru $\mathcal{P}_{2n-1}$
- i napisati sustav od $2n$ jednadžbi s $2n$ nepoznanica,

iz uvjeta egzaktne integracije na toj bazi.

Na primjer, u standardnoj bazi $\{1, x, x^2, \ldots, x^{2n-1}\}$ dobivamo sustav oblika

$$\mu_j = \int\limits_a^b w(x) x^j \, dx = \sum_{k=1}^n w_k x_k^j, \quad j = 0, 1, \ldots, 2n-1.$$

Međutim, to je loš pristup!

Sustav jednadžbi iz uvjeta egzaktne integracije

Što ne valja? Ključni problem je nelinearnost ovog sustava.

- Ovisnost o nepoznanicama x_k je nelinearna.

Već i dokaz da ovaj nelinearni sustav ima jedinstveno rješenje nije jednostavan.

Drugi problem je moguća

- loša uvjetovanost izabrane baze prostora polinoma.

Potencijalni popravak:

- uzeti bazu pripadnih ortogonalnih polinoma p_n.

Nažalost, to pomaže tek kad jednom izračunamo čvorove x_k, pa ostaje linearni sustav (reda n) za težine w_k.

Dakle, nema puno smisla!

Parametri Gaussovih formula

Napomena. Za neke “klasične” izbore težinskih funkcija w i intervala $[a, b]$, postoje

- tablice čvorova i težina pripadnih Gaussovih formula,
- za neke (male) vrijednosti n — tipično je $n \leq 20$,
- na vrlo visoku točnost — 20, pa i više decimala.

Međutim, čak i tad imamo “problem”:

- treba korektno “prekucati” tabelirane vrijednosti u naš program!

Probajte jednom — i provjerite jesu li sve vrijednosti korektne! (Test je egzaktna integracija polinoma.)

Dakle, korisno je znati kako izgleda algoritam za računanje parametara Gaussovih formula.

Ortogonalni polinomi i tročlana rekurzija

Algoritam se bazira na pripadnim ortogonalnim polinomima i

- tročlanoj rekurziji za te polinome.

Neka je $\{p_k \mid k \geq 0\}$ familija ortogonalnih polinoma na intervalu $[a, b]$ s težinskom funkcijom w.

Već smo pokazali da ovi polinomi zadovoljavaju tročlanu homogenu rekurziju oblika

$$p_{k+1}(x) = (a_k x + b_k) p_k(x) - c_k p_{k-1}(x), \quad k \geq 1.$$

Izveli smo i formule za koeficijente a_k, b_k i c_k u ovoj rekurziji.

Proširenje. Uz dogovor $p_{-1}(x) = 0$, rekurzija vrijedi i za $k = 0$, s proizvoljnim c_0, a koeficijente a_0 i b_0 izračunamo iz p_0 i p_1.

Monični ortogonalni polinomi

Izvod algoritma za nalaženje parametara Gaussovih formula obično kreće od ortogonalnih polinoma p_k

- s vodećim koeficijentom $A_k = 1$.

Ovi polinomi zovu se monični ortogonalni polinomi.

Monični ortogonalni polinomi zadovoljavaju

- još jednostavniju tročlanu rekurziju (jer je $a_k = 1$),

koja se standardno piše u sljedećem obliku:

$$p_{k+1}(x) = (x - \alpha_k)p_k(x) - \beta_k p_{k-1}(x), \quad k = 0, 1, \ldots .$$

Uočite "pomak" u rekurziji — rekurzija starta od nule!
Veza ovih koeficijenata s ranijim: $a_k = 1$, $\alpha_k = -b_k$, $\beta_k = c_k$.

Rekurzija za monične ortogonalne polinome

Rekurzija je

$$p_{k+1}(x) = (x - \alpha_k)p_k(x) - \beta_k p_{k-1}(x), \quad k = 0, 1, \ldots.$$

Po definiciji, prva dva polinoma su

$$p_{-1}(x) := 0, \quad p_0(x) = 1.$$

Uz skraćeni zapis integralnog skalarnog produkta $\langle\ ,\ \rangle$, koeficijenti u ovoj rekurziji dani su formulama (v. ranije)

$$\alpha_k = \frac{\langle xp_k, p_k\rangle}{\langle p_k, p_k\rangle}, \quad k = 0, 1, 2, \ldots,$$

$$\beta_k = \frac{\langle xp_k, p_{k-1}\rangle}{\langle p_{k-1}, p_{k-1}\rangle} = \frac{\langle p_k, p_k\rangle}{\langle p_{k-1}, p_{k-1}\rangle} > 0, \quad k = 1, 2, \ldots.$$

Rekurzija za monične ortogonalne polinome

Standardno se još definira da je

$$\beta_0 := \mu_0 = \int\limits_a^b w(x)\, dx.$$

Zbog $p_{-1}(x) = 0$, ovaj koeficijent β_0 nije bitan u rekurziji, već ima drugu svrhu. I za njega vrijedi $\beta_0 > 0$.

Pretpostavimo sad da su

- svi potrebni koeficijenti α_k i β_k poznati.

Ako nisu, postoje numerički postupci za njihovo računanje.

Za zadani n, čvorovi $x_1, \ldots, x_n$ su nultočke polinoma p_n. Zato u rekurziji trebamo koeficijente za $k \le n - 1$.

Matrični zapis rekurzije za ortogonalne polinome

Za početak, rekurziju za monične ortogonalne polinome

$$p_{k+1}(x) = (x - \alpha_k)p_k(x) - \beta_k p_{k-1}(x), \quad k = 0, 1, \ldots,$$

napišemo tako da član $xp_k(x)$ ostane sam na desnoj strani

$$p_{k+1}(x) + \alpha_k p_k(x) + \beta_k p_{k-1}(x) = xp_k(x), \quad k = 0, 1, \ldots.$$

Prvih n relacija iz rekurzije, za $k = 0, \ldots, n-1$, možemo zapisati u matričnom zapisu,

- tako da lijevu stranu svake relacije gledamo kao linearnu kombinaciju vrijednosti

$$p_0(x),\ p_1(x),\ \ldots\ ,\ p_{n-1}(x).$$

U zadnjoj relaciji, $p_n(x)$ pišemo posebno.

Matrični zapis rekurzije za ortogonalne polinome

Dobivamo

$$\begin{bmatrix} \alpha_0 & 1 & & & \\ \beta_1 & \alpha_1 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{n-2} & \alpha_{n-2} & 1 \\ & & & \beta_{n-1} & \alpha_{n-1} \end{bmatrix} \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{n-2}(x) \\ p_{n-1}(x) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ p_n(x) \end{bmatrix} = x \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{n-2}(x) \\ p_{n-1}(x) \end{bmatrix}.$$

Uvedimo oznake

$$T_n = \begin{bmatrix} \alpha_0 & 1 & & & \\ \beta_1 & \alpha_1 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{n-2} & \alpha_{n-2} & 1 \\ & & & \beta_{n-1} & \alpha_{n-1} \end{bmatrix}, \quad z(x) = \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{n-2}(x) \\ p_{n-1}(x) \end{bmatrix}.$$

Matrični zapis rekurzije za ortogonalne polinome

Onda dobivamo "skraćeni" matrični zapis

$$T_n z(x) + p_n(x) e_n = x\, z(x),$$

gdje je $e_n = (0, \ldots, 0, 1)^T$ zadnji vektor standardne baze u $\mathbb{R}^n$. Dodatno još, zbog $p_0(x) = 1$, uvijek vrijedi $z(x) \neq 0$.

Sad ide ključna primjedba:

- ako je x_k nultočka polinoma p_n, onda je x_k svojstvena vrijednost matrice T_n, a $z(x_k)$ je pripadni svojstveni vektor.

Vrijedi i obrat:

- ako je x svojstvena vrijednost matrice T_n, onda je x nultočka polinoma p_n.

Čvorovi kao svojstvene vrijednosti

Dakle, sve svojstvene vrijednosti matrice T_n su, upravo, sve nultočke polinoma p_n, tj. svi čvorovi integracije $x_1, \ldots, x_n$.

Zaključak: za računanje čvorova možemo koristiti algoritme

- za računanje svojstvenih vrijednosti tridijagonalne (općenito, nesimetrične) matrice T_n.

Međutim, to se u praksi nikad ne radi tako,

- preko nesimetrične matrice T_n.

Razlog: Postoji i puno bolji pristup!

- Matrica T_n se uvijek može simetrizirati u tzv. Jacobijevu matricu J_n.

Simetrizacija matrice — Jacobijeva matrica J_n

Tvrdnja. Matrica T_n je dijagonalno slična simetričnoj matrici

$$J_n = \begin{bmatrix} \alpha_0 & \sqrt{\beta_1} & & & \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & \\ & \ddots & \ddots & \ddots & \\ & & \sqrt{\beta_{n-2}} & \alpha_{n-2} & \sqrt{\beta_{n-1}} \\ & & & \sqrt{\beta_{n-1}} & \alpha_{n-1} \end{bmatrix}$$

koju zovemo Jacobijeva matrica. Ovdje je bitno da je $\beta_k > 0$.

Preciznije, vrijedi $D_n^{-1} T_n D_n = J_n$, pri čemu je

$$D_n = d_0 D_n' = d_0 \cdot \operatorname{diag}\big(1, \sqrt{\beta_1}, \sqrt{\beta_1 \beta_2}, \ldots, \sqrt{\beta_1 \cdots \beta_{n-1}}\,\big),$$

a $d_0 \neq 0$ je proizvoljan skalar (d_0 se skrati u izrazu za J_n). ■

Čvorovi kao svojstvene vrijednosti matrice J_n

Slične matrice T_n i J_n imaju iste svojstvene vrijednosti.

Zaključak: Čvorove integracije možemo izračunati kao

- svojstvene vrijednosti simetrične tridijagonalne matrice J_n.

Prednosti ovog pristupa:

- Simetrična matrica J_n ima realne svojstvene vrijednosti,
- pripadni svojstveni vektori su ortogonalni,
- iz njih se lako računaju težine (Golub–Welsch algoritam).

Dodatno, za simetrične tridijagonalne matrice postoje

- vrlo efikasni i točni algoritmi za svojstveni problem.

Simetrizacija matrice i rekurzija

Simetrizaciji matrice T_n u Jacobijevu matricu J_n odgovara

- simetrizacija rekurzije za pripadne ortogonalne polinome.

Iz moničnih polinoma p_k, supstitucijom

$$\widetilde{p}_k(x) = \frac{1}{\sqrt{\beta_0\,\beta_1 \cdots \beta_k}}\, p_k(x), \qquad k = 0, 1, 2, \ldots,$$

prelazimo na ortogonalne polinome $\widetilde{p}_k$ koji nisu monični, nego normirani ($\|\widetilde{p}_k\| = 1$), i zadovoljavaju simetriziranu rekurziju

$$\sqrt{\beta_{k+1}}\,\widetilde{p}_{k+1}(x) = (x - \alpha_k)\widetilde{p}_k(x) - \sqrt{\beta_k}\,\widetilde{p}_{k-1}(x),$$

za $k = 0, 1, \ldots$. Start je, ovdje, $\widetilde{p}_{-1}(x) = 0$ i $\widetilde{p}_0(x) = 1/\sqrt{\beta_0}$.

Ovoj rekurziji odgovara Jacobijeva matrica J_n.

Svojstveni vektori Jacobijeve matrice J_n

Ortogonalni polinomi p_n i $\widetilde{p}_n$, naravno, imaju iste nultočke, a to su, ujedno, i svojstvene vrijednosti matrice J_n.

Za bilo koju nultočku x_k polinoma $\widetilde{p}_n$, iz matričnog zapisa rekurzije slijedi

$$J_n \widetilde{z}_k = x_k \widetilde{z}_k,$$

gdje je

$$\widetilde{z}_k := \widetilde{z}(x_k) = \begin{bmatrix} \widetilde{p}_0(x_k) \\ \widetilde{p}_1(x_k) \\ \vdots \\ \widetilde{p}_{n-1}(x_k) \end{bmatrix}$$

svojstveni vektor matrice J_n, koji pripada svojstvenoj vrijednosti x_k, za $k = 1, \ldots, n$. Vektore $\widetilde{z}_k$ smo već spominjali!

Ortogonalnost svojstvenih vektora (još jednom)

Znamo da su sve svojstvene vrijednosti x_k međusobno različite (to su nultočke ortogonalnog polinoma $\widetilde{p}_n$). Onda su

- pripadni svojstveni potprostori jednodimenzionalni,
- i još moraju biti ortogonalni, jer je J_n simetrična matrica!

To znači da su svojstveni vektori $\widetilde{z}_k$ međusobno ortogonalni. Uz oznaku $\langle\ ,\ \rangle_n$ za "obični" skalarni produkt u $\mathbb{R}^n$, vrijedi

$$\langle \widetilde{z}_j, \widetilde{z}_k \rangle_n = 0, \quad \text{za} \quad j \neq k.$$

Napomena. Ovu ortogonalnost vektora $\widetilde{z}_k$ dokazali smo ranije,

- kod Christoffel–Darbouxovog identiteta.

Tamo su vektori $\widetilde{z}_k$ bili stupci (neke) matrice Z_n, a sad vidimo da je $Z_n =$ matrica svojstvenih vektora Jacobijeve matrice J_n.

Ortonormirana baza svojstvenih vektora

Svojstveni vektori $\widetilde{z}_k$ matrice J_n, općenito,

- nisu normirani, tj. vrijedi $\|\widetilde{z}_k\| \neq 1$,

već su skalirani tako da im je prva komponenta jednaka

$$\widetilde{z}_{k,1} = \widetilde{p}_0(x_k) = \frac{1}{\sqrt{\mu_0}} = \frac{1}{\sqrt{\beta_0}}, \qquad k = 1, \ldots, n.$$

Ako želimo ortonormiranu bazu svojstvenih vektora, možemo ih normirati,

$$v_k := \frac{\widetilde{z}_k}{\|\widetilde{z}_k\|}, \qquad k = 1, \ldots, n,$$

i onda vrijedi

$$\langle v_j, v_k \rangle_n = \delta_{j,k}.$$

Dakle, $v_1, \ldots, v_n$ je ortonormirana baza svojstvenih vektora matrice J_n u prostoru $\mathbb{R}^n$ (v. matricu V_n kod diskretne ortog.).

Ortonormirana baza svojstvenih vektora

Za prve komponente vektora v_k ortonormirane baze onda vrijedi

$$v_{k,1} = \frac{\widetilde{z}_{k,1}}{\|\widetilde{z}_k\|} = \frac{1}{\sqrt{\beta_0}\,\|\widetilde{z}_k\|}, \qquad k = 1, \ldots, n.$$

Ako znamo v_k, odavde dobivamo norme

$$\|\widetilde{z}_k\| = \frac{1}{\sqrt{\beta_0}\, v_{k,1}}, \qquad k = 1, \ldots, n.$$

Ova veza je korisna u praksi. Naime,

- ako numerički računamo svojstvene vektore matrice J_n,
- kao rezultat, dobivamo ortonormiranu bazu $v_1, \ldots, v_n$.

Razlog: Dijagonalizacija simetrične matrice J_n uvijek se radi ortogonalnim transformacijama (sličnosti = kongruencije)!

Računanje težina Gaussovih formula

Iz ranijeg teorema o težinama, znamo da je

$$w_k = \frac{1}{\|\widetilde{z}_k\|^2}, \quad k = 1, \ldots, n.$$

Kad uvrstimo izraz za $\|\widetilde{z}_k\|$ u izračunatoj ortonormiranoj bazi, $\|\widetilde{z}_k\| = 1/(\sqrt{\beta_0}\, v_{k,1})$, dobivamo da za težine vrijedi

$$w_k = \beta_0\, v_{k,1}^2, \quad k = 1, \ldots, n.$$

Težine su kvadrati prvih komponenti normiranih svojstvenih vektora matrice J_n, pomnoženi s β_0.

Ovo je tzv. Golub–Welsch algoritam za računanje parametara Gaussovih integracijskih formula, a objavljen je 1969. g.

Složenost cijelog postupka

Složenost:

- $O(n^3)$ — ako za matricu J_n računamo svojstvene vrijednosti $x_1, \ldots, x_n$ i (ortonormirane) svojstvene vektore $v_1, \ldots, v_n$,
- $O(n^2)$ — ako računamo samo svojstvene vrijednosti x_k, a elemente $\widetilde{p}_j(x_k)$ svojstvenih vektora $\widetilde{z}_1, \ldots, \widetilde{z}_n$ računamo na kraju, po rekurziji.

Za jednu klasu posebnih Gaussovih formula, postoji još brži algoritam. Složenost je linearna, tj. $O(n)$, što je optimalno.

- Polinomi p_n zadovoljavaju posebni oblik diferencijalne jednadžbe drugog reda. Vrijedi za sve klasične formule.
- Autori su Glaser, Liu i Rokhlin, a članak je iz 2007. g.

Rješavanje nelinearnih jednadžbi

Općenito o iterativnim metodama

Neka je zadana nelinearna funkcija

$$f : I \to \mathbb{R},$$

gdje je I neki interval. Tražimo sve one točke $x \in I$ za koje je

$$f(x) = 0.$$

Takve točke x zovu se

- rješenja ili korijeni pripadne jednadžbe,
- ili nultočke funkcije f.

U pravilu, pretpostavljamo da je

- f neprekidna na I i
- da su joj nultočke izolirane.

Neprekidnost funkcije f

Neprekidnost funkcije f obično se koristi pri određivanju intervala gdje se nalazi nultočka. Naime, ako je

$$f(a) \cdot f(b) < 0$$

na nekom intervalu $[a, b]$, to znači da funkcija je promijenila znak na $[a, b]$. To se može dogoditi na dva načina:

- ili f ima nultočku na $[a, b]$,
- ili f ima prekid na $[a, b]$.

Ako je f neprekidna funkcija na $[a, b]$

- i u rubovima vrijedi $f(a) \cdot f(b) < 0$,

onda f sigurno ima nultočku na $[a, b]$ — čak unutar (a, b).

Izoliranost nultočaka

Definicija (Izolirana nultočka). Za nultočku α reći ćemo da je izolirana, ako postoji krug nekog pozitivnog radijusa oko α,

- takav da je α jedina nultočka od f unutar tog kruga.

U protivnom, kažemo da je nultočka neizolirana. ■

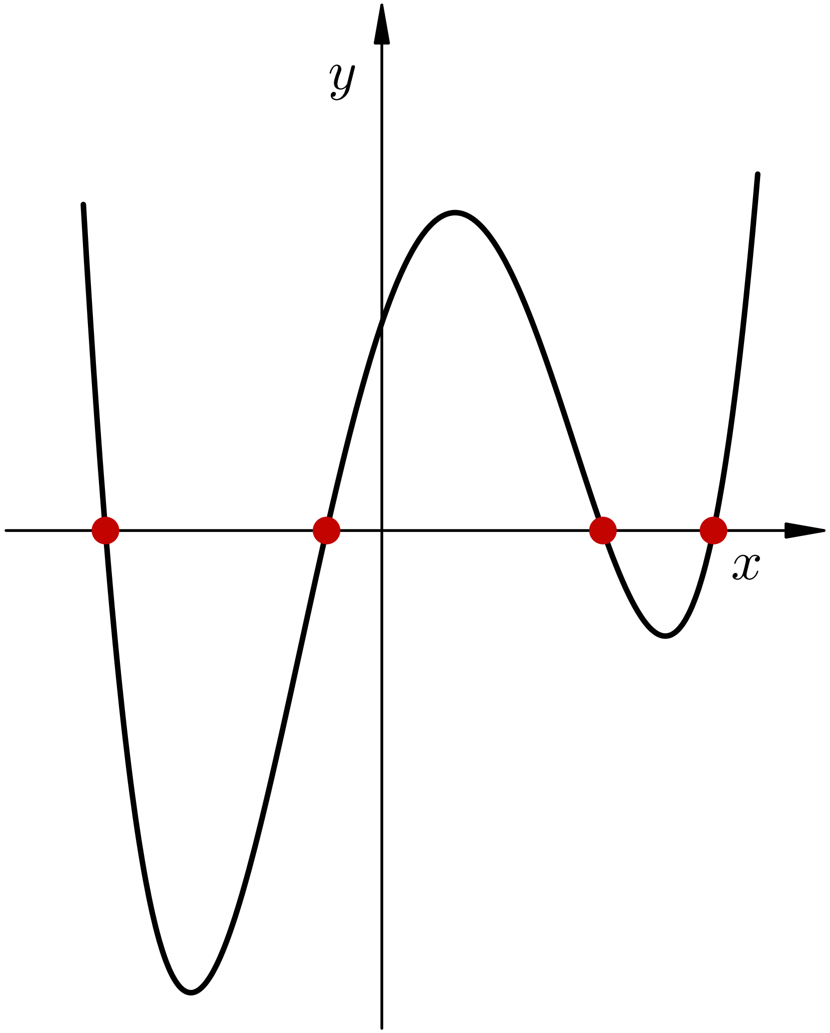
Kod neizoliranih nultočaka postoji problem konvergencije metoda za nalaženje nultočaka (lokalna nejedinstvenost).

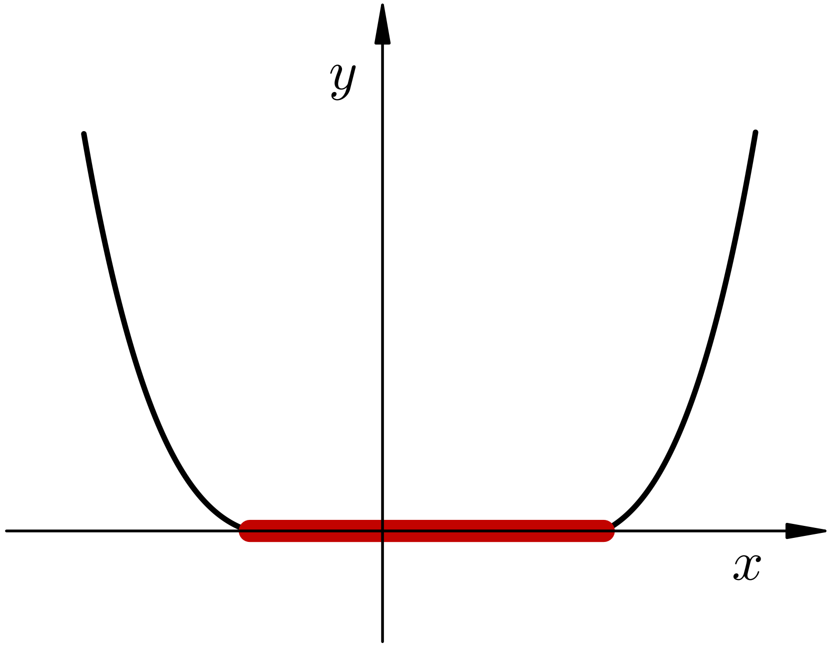
Odsad nadalje, pretpostavljamo da f ima samo izolirane nultočke.

Na sljedećoj stranici su primjeri funkcije s

- izoliranim nultočkama (lijevo),
- neizoliranim nultočkama (desno).

Izoliranost nultočaka





Računanje nultočke na zadanu točnost

Traženje nultočki na zadanu točnost sastoji se od dvije faze:

1. Izolacija jedne ili više nultočki, tj. nalaženje intervala I unutar kojeg se nalazi barem jedna nultočka. Ovo je teži dio posla i obavlja se na temelju analize toka funkcije.
2. Iterativno nalaženje nultočke na traženu točnost.

Postoji mnogo metoda za nalaženje nultočaka nelinearnih funkcija na zadanu točnost. One se bitno razlikuju po tome

- imamo li sigurnu konvergenciju ili ne,
- i po brzini konvergencije (ako/kad konvergiraju).

Uobičajeno:

- brze metode nemaju sigurnu konvergenciju,
- dok je sporije metode imaju.

Brzina ili red konvergencije

Definirajmo sada brzinu konvergencije nekog niza "iteracija". Te iteracije mogu, ali ne moraju biti iteracije za računanje nultočke funkcije.

Definicija. Niz iteracija $(x_n, n \in \mathbb{N}_0)$ konvergira prema točki α

- s redom konvergencije p, gdje je $p \geq 1$,

ako je p najveći realan broj, za kojeg postoji konstanta $c > 0$, tako da vrijedi

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

Ako je $p = 1$, kažemo da niz konvergira linearno prema α. U tom slučaju, mora biti $c < 1$ (za konvergenciju niza), i obično se c naziva faktor linearne konvergencije. ■

Linearna konvergencija

U nekim slučajevima, prethodna definicija nije zgodna za linearne iterativne algoritme.

Ako u prethodnoj formuli upotrijebimo indukciju za $p = 1$, uz $c < 1$, onda dobivamo da je

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

Katkad će biti mnogo lakše pokazati ovu relaciju, nego onu iz definicije. I u ovom slučaju, kažemo da niz iteracija konvergira linearno s faktorom c. Zbog $c^n \to 0$, niz zaista konvergira.

Iz očitih razloga, ovakvu linearnu konvergenciju još zovemo i

- geometrijska konvergencija s faktorom c.

Metoda bisekcije (raspolavljanja)

Uvodno o metodi raspolavljanja

Najjednostavnija metoda nalaženja nultočaka funkcije je metoda bisekcije ili raspolavljanja.

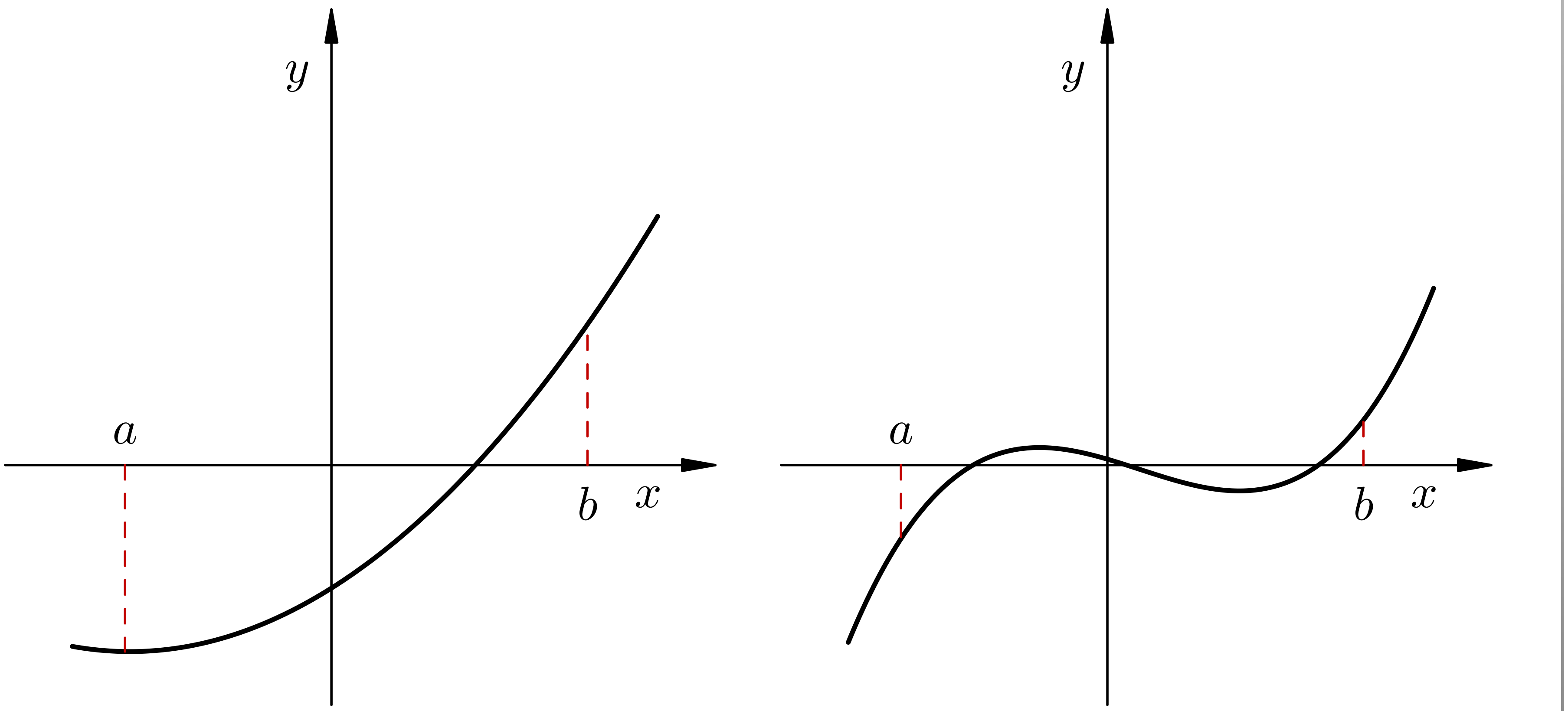
- Osnovna ili startna pretpostavka za početak algoritma raspolavljanja je neprekidnost funkcije f na intervalu $[a, b]$, s tim da u rubovima intervala vrijedi

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

To znači da f ima barem jednu nultočku u intervalu $[a, b]$. Međutim, f može imati i više nultočaka u intervalu $[a, b]$. Na sljedećoj stranici su primjeri kad funkcija f ima

- točno jednu nultočku unutar $[a, b]$ (lijevo),
- više nultočaka unutar $[a, b]$ — točnije, neparan broj njih, brojeći kratnost (desno).

Uvodno o metodi raspolavljanja



Uvodno o metodi raspolavljanja

Naravno, ako je $f(a) \cdot f(b) = 0$, onda f sigurno ima nultočku u (barem) jednom rubu intervala — a ili b. Provjerom $f(a) = 0$, odnosno, $f(b) = 0$, otkrivamo nultočke u rubovima.

Na kraju, ako je

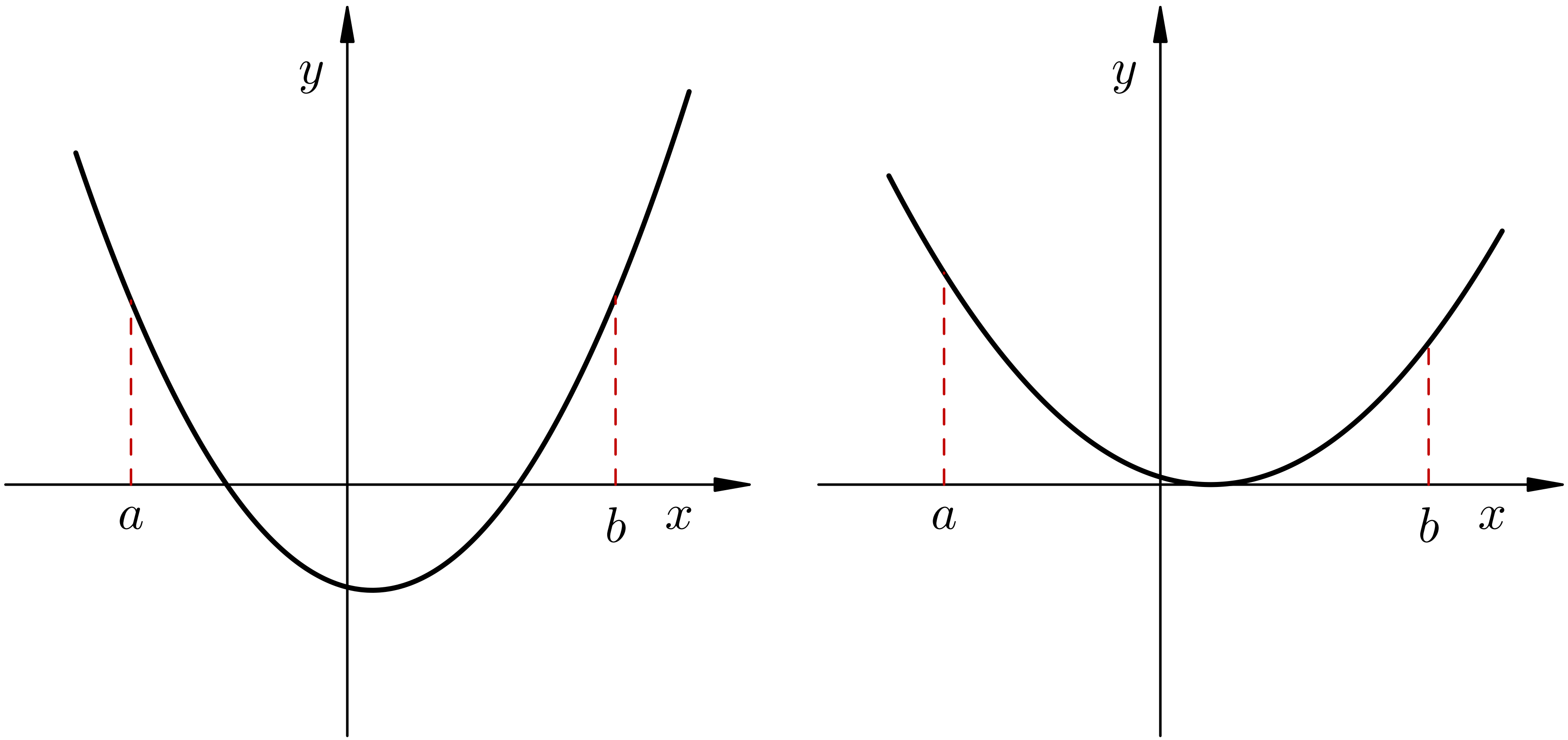
$$f(a) \cdot f(b) > 0,$$

to ne mora značiti da f nema nultočku unutar $[a, b]$.

Na primjer, moglo se dogoditi da smo loše separirali (locirali) nultočke i da f, unutar intervala $[a, b]$, ima

- paran broj nultočaka (slika lijevo),
- ili nultočku parnog reda (slika desno).

Uvodno o metodi raspolavljanja



Uvodno o metodi raspolavljanja

Zaključak.

- Boljom separacijom nultočaka na lijevoj slici, lako ćemo postići da je $f(a) \cdot f(b) < 0$.
- Nultočke parnog reda nemoguće je direktno naći metodom bisekcije (nema promjene predznaka).

Kad ćemo govoriti o nultočkama višeg reda, onda ćemo pokazati kako treba modificirati funkciju, tako da i metodom bisekcije možemo naći višestruku nultočku.

- Umjesto f, treba raditi s funkcijom f/f'.

Algoritam

Označimo s α pravu nultočku funkcije (nju tražimo), a zatim s

- $a_0 := a$,
- $b_0 := b$ i
- $x_0 :=$ polovište intervala $[a_0, b_0]$, tj.

$$x_0 = \frac{a_0 + b_0}{2}.$$

Ideja metode: U n-tom koraku algoritma, počev od intervala $[a_{n-1}, b_{n-1}]$ koji sigurno sadrži neku nultočku α,

- konstruiramo interval $[a_n, b_n]$ kojemu je
- duljina = polovina duljine prethodnog intervala,
- ali tako da je nultočka α ostala unutar intervala $[a_n, b_n]$.

Algoritam

Konstrukcija intervala $[a_n, b_n]$ sastoji se u raspolavljanju intervala $[a_{n-1}, b_{n-1}]$ točkom x_{n-1}, na sljedeći način:

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

Uočimo da je dovoljno ispitivati koji predznak imamo za $f(a_{n-1}) \cdot f(x_{n-1})$. Imamo tri mogućnosti:

- $f(a_{n-1}) \cdot f(x_{n-1}) = 0$ znači da je nultočka upravo x_{n-1},
- $f(a_{n-1}) \cdot f(x_{n-1}) < 0$ znači da je barem jedna nultočka unutar $[a_{n-1}, x_{n-1}]$ — lijeva polovina,
- $f(a_{n-1}) \cdot f(x_{n-1}) > 0$ znači da je barem jedna nultočka unutar $[x_{n-1}, b_{n-1}]$ — desna polovina.

Algoritam

Objasnimo posljednju činjenicu. Množenjem lijevih strana nejednakosti

$$f(a_{n-1}) \cdot f(b_{n-1}) < 0$$

$$f(a_{n-1}) \cdot f(x_{n-1}) > 0$$

dobivamo

$$\big(f(a_{n-1})\big)^2 \cdot f(x_{n-1}) \cdot f(b_{n-1}) < 0,$$

pa mora biti

$$f(x_{n-1}) \cdot f(b_{n-1}) < 0.$$

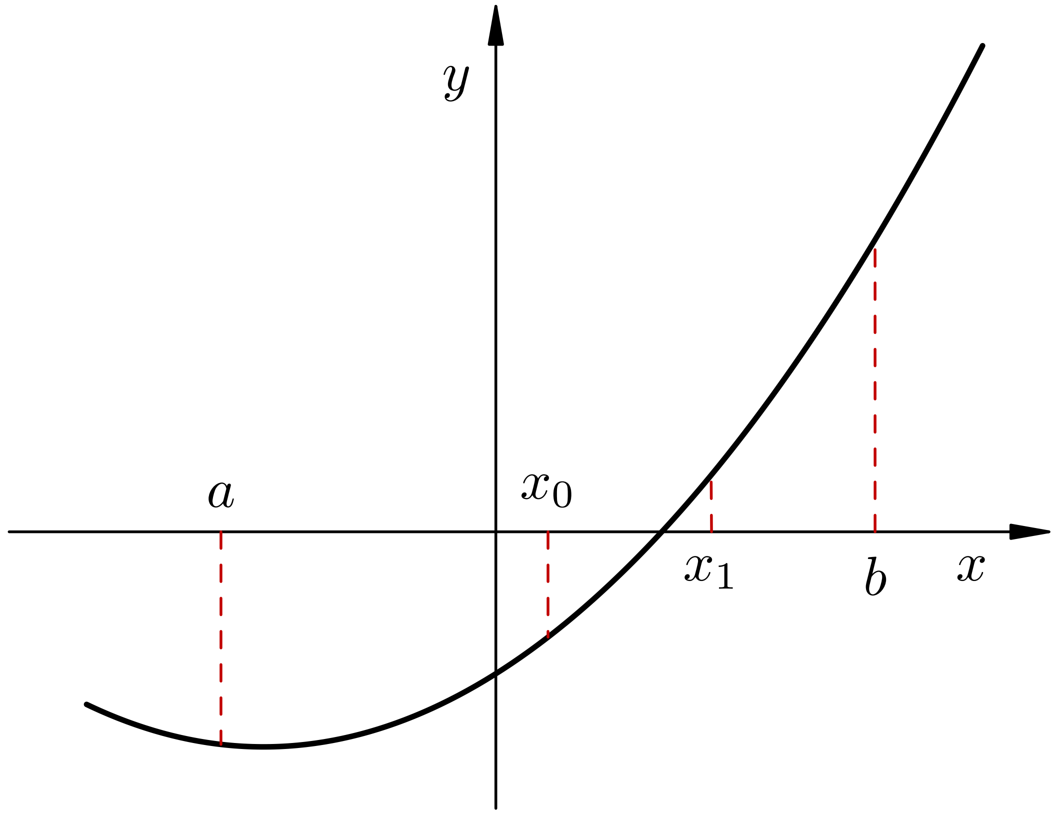
Dakle, nultočka se onda mora nalaziti u intervalu $[x_{n-1}, b_{n-1}]$.

Za svaki korak metode raspolavljanja, očito, vrijedi zaključak

$$\alpha \in [a_{n-1}, b_{n-1}] \Longrightarrow \alpha \in [a_n, b_n].$$

Metoda raspolavljanja — grafički

Grafički, metoda raspolavljanja izgleda ovako



Algoritam

Metoda raspolavljanja — za zadanu (apsolutnu) točnost ε:

```
x = (a + b) / 2;
dok je b - x > epsilon radi {  // ili x - a > ...
  ako je f(a) * f(x) <= 0.0 onda {
    b = x;
  }
  inače {
    a = x;
  }
  x = (a + b) / 2;
}
/* Na kraju je x ≈ alpha. */
```

Pedantni algoritam "čuva" i stare vrijednosti funkcije.

Konvergencija i zaustavljanje algoritma

Tvrdnja. Ako vrijede startne pretpostavke na funkciju f za metodu raspolavljanja, dobiveni niz x_n konvergira prema nekoj nultočki α iz intervala $[a, b]$.

Dokaz. Nultočku α smo našli sa zadanom točnošću ε ako je

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

Kako ćemo znati da je to ispunjeno, ako ne znamo α?

- Budući da je x_n polovište intervala $[a_n, b_n]$ i $\alpha \in [a_n, b_n]$, onda je

$$|\alpha - x_n| \leq b_n - x_n = x_n - a_n = \frac{1}{2}\,(b_n - a_n),$$

- pa je dovoljno zahtijevati

$$b_n - x_n \leq \varepsilon \quad \text{ili} \quad x_n - a_n \leq \varepsilon.$$

Ocjena greške

Iz konstrukcije metode raspolavljanja, lako se izvodi ocjena pogreške n-te aproksimacije x_n nultočke α. Vrijedi:

$$|\alpha - x_n| \le b_n - x_n = \frac{1}{2}\,(b_n - a_n) = \frac{1}{2^2}\,(b_{n-1} - a_{n-1})$$

$$= \cdots = \frac{1}{2^{n+1}}\,(b - a).$$

Nadalje, vrijedi $(b-a)/2 = b - x_0 = x_0 - a$, pa slijedi da je

$$|\alpha - x_n| \le \frac{1}{2^n}\,(b - x_0) = \frac{1}{2^n}\,(x_0 - a).$$

■

Ova relacija podsjeća na linearnu konvergenciju s $c = 1/2$, ali se zdesna ne pojavljuje $|\alpha - x_0|$. Ipak, desna strana sugerira da će konvergencija biti dosta spora (jedan bit po iteraciji).

Ocjena greške i broj koraka

Relacija

$$|\alpha - x_n| \le \frac{1}{2^{n+1}}\,(b-a)$$

omogućava da se unaprijed odredi koliko je koraka (iteracija) raspolavljanja potrebno za postizanje zadane točnosti ε.

Da osiguramo $|\alpha - x_n| \le \varepsilon$, dovoljno je zahtijevati da je

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

Množenjem prethodne jednadžbe s 2^{n+1} i dijeljenjem s ε, dobivamo

$$\frac{b-a}{\varepsilon} \le 2^{n+1}.$$

Ocjena greške i broj koraka

Zatim, logaritmiranje (u bilo kojoj bazi) daje

$$\log(b-a) - \log\varepsilon \le (n+1)\log 2,$$

odnosno,

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

Ako je funkcija f još i klase $C^1[a,b]$, tj. ako f ima neprekidnu prvu derivaciju, može se dobiti i tzv. dinamička ocjena greške.

Po Teoremu srednje vrijednosti za funkciju f oko α, imamo

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

pri čemu je ξ između x_n i α.

Dinamička ocjena greške

Prvo iskoristimo da je α nultočka, tj. da je $f(\alpha) = 0$, a zatim uzmemo apsolutne vrijednosti obje strane. Dobivamo

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

Znamo da su α i x_n iz $[a, b]$, odakle slijedi i $\xi \in [a, b]$. Onda $|f'(\xi)|$ ocijenimo odozdo, preko cijelog intervala $[a, b]$,

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

Ako je $m_1 > 0$, odnosno, ako f' ima fiksni predznak na $[a, b]$, onda izlazi ocjena

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

Dinamička ocjena greške

Drugim riječima, ako želimo da je $|\alpha - x_n| \leq \varepsilon$, dovoljno je zahtijevati da je

$$\frac{|f(x_n)|}{m_1} \leq \varepsilon,$$

odnosno, da vrijedi

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

Ovaj uvjet možemo provjeriti u svakoj iteraciji.

- Iteracije smijemo "prekinuti" čim je ovaj uvjet ispunjen,
- neovisno o unaprijed izračunatom potrebnom broju iteracija.

Napomena. Pretpostavka $m_1 > 0$ znači da f' nema nultočku na $[a, b]$, tj. da je f monotona na $[a, b]$ $(\Rightarrow \alpha$ je jedinstvena).

Regula falsi
(metoda pogrešnog položaja)

Uvodno o metodi pogrešnog položaja

Znamo da metoda raspolavljanja ima

- sigurnu konvergenciju, ali je vrlo spora.

Regula falsi ili metoda pogrešnog položaja je prirodan pokušaj ubrzavanja metode raspolavljanja. I ova metoda ima

- sigurnu konvergenciju,

uz iste pretpostavke kao u metodi raspolavljanja.

Pretpostavimo da je funkcija $f : [a, b] \to \mathbb{R}$

- neprekidna na intervalu $[a, b]$
- i da u rubovima intervala vrijedi "promjena znaka"

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

Ideja i skica algoritma

Ideja metode: Aproksimirajmo funkciju f pravcem koji prolazi točkama $(a, f(a))$ i $(b, f(b))$.

Traženu nultočku α tada možemo aproksimirati

- nultočkom tog pravca — označimo ju s x_0.

Uočite da pravac sigurno siječe os x, zbog $f(a) \cdot f(b) < 0$.

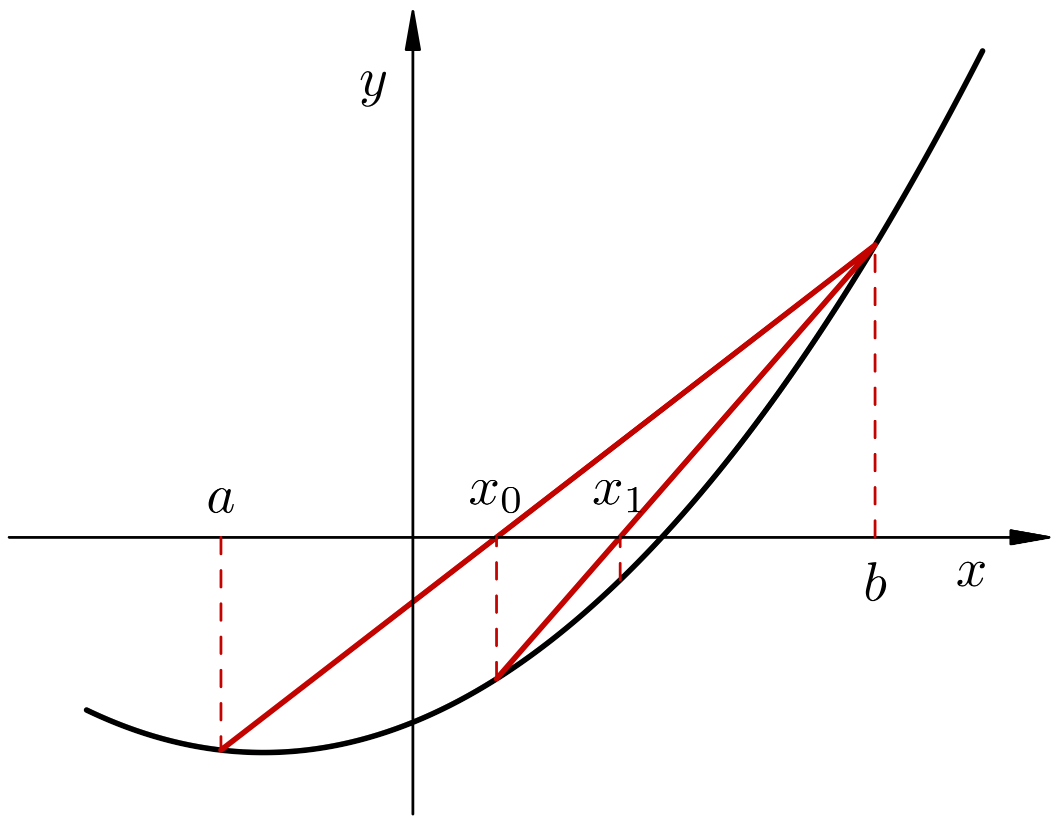
Nakon toga,

- pomaknemo ili točku a, ili točku b — u točku x_0,
- ali tako da je nultočka ostala unutar novodobivenog intervala (test predznaka, kao kod raspolavljanja).

Postupak ponavljamo sve dok nismo postigli željenu točnost.

Metoda pogrešnog položaja — grafički

Grafički, regula falsi ili metoda pogrešnog položaja izgleda ovako



Regula falsi — osnovne ideje

Točka x_0 dobiva se jednostavno iz jednadžbe pravca, pa je

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

Dakle, osnovne ideje metode su:

- aproksimacija pravcem
- i “zatvaranje” nultočke u određeni — sve manji interval.

Iz slike zaključujemo da je to sasvim dobra ideja,

- za monotone i konveksne (ili konkavne) funkcije.

Nažalost, postoje ozbiljni problemi i s ovom metodom.

- Konvergencija je i dalje linearna, kao kod raspolavljanja.
- Može biti vrlo spora — sporija nego kod raspolavljanja.

Regula falsi — red konvergencije

Izvedimo red konvergencije metode pogrešnog položaja.

Uz oznaku za prvu podijeljenu razliku

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

formula za prvu aproksimaciju x_0 glasi

$$x_0 = b - \frac{f(b)}{f[a, b]}.$$

Treba naći izraz za grešku $\alpha - x_0$.

Prethodnu relaciju pomnožimo s -1 i dodamo α na obje strane, tako da lijeva strana postane upravo $\alpha - x_0$.

Regula falsi — red konvergencije

Dobivamo redom

$$\begin{aligned}
\alpha - x_0 &= \alpha - b + \frac{f(b)}{f[a,b]} = (\text{izlučimo } \alpha - b) \\
&= (\alpha - b)\left(1 + \frac{f(b)}{(\alpha - b) f[a,b]}\right) = (\text{uvalimo } f(\alpha) = 0) \\
&= (\alpha - b)\left(1 + \frac{f(b) - f(\alpha)}{(\alpha - b) f[a,b]}\right) = (\text{sredimo u } f[b,\alpha]) \\
&= (\alpha - b)\left(1 - \frac{f[b,\alpha]}{f[a,b]}\right) = (\alpha - b)\,\frac{f[a,b] - f[b,\alpha]}{f[a,b]} \\
&= -(\alpha - b)\,(\alpha - a)\,\frac{f[a,b,\alpha]}{f[a,b]}.
\end{aligned}$$

Regula falsi — red konvergencije

Ako je funkcija f dovoljno glatka, onda

- podijeljene razlike $f[a, b, \alpha]$ i $f[a, b]$ možemo napisati preko derivacija funkcije f.

Ako je f klase $C^1[a, b]$, onda po Teoremu srednje vrijednosti imamo

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

Na sličan način, ako je f klase $C^2[a, b]$, onda vrijedi

$$f[a, b, \alpha] = \frac{1}{2} f''(\zeta),$$

gdje se ζ nalazi između minimuma i maksimuma vrijednosti a, b i α. Zbog $\alpha \in [a, b]$, to opet daje $\zeta \in [a, b]$.

Sad ove dvije relacije uvrstimo u izraz za grešku.

Regula falsi — red konvergencije

Za funkciju $f \in C^2[a, b]$, dobivamo sljedeći izraz za grešku

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

Uočimo još da, zbog $\alpha \in [a, b]$, vrijedi

$$-(\alpha - b)\,(\alpha - a) = (b - \alpha)\,(\alpha - a) > 0.$$

Da bismo pojednostavnili analizu, pretpostavimo da

- prva derivacija f' i druga derivacija f'' imaju konstantan predznak na $[a, b]$ (pozitivne ili negativne).

Onda je α jedina nultočka funkcije f u intervalu $[a, b]$, jer je f monotona.

Regula falsi — red konvergencije

U nastavku gledamo samo jedan od četiri moguća slučaja za predznake f' i f''.

Pretpostavimo da je $f' > 0$ i $f'' > 0$ na $[a, b]$, tj. da je

- f monotono rastuća i konveksna na $[a, b]$.

U tom slučaju,

- spojnica točaka $(a, f(a))$ i $(b, f(b))$ uvijek se nalazi iznad grafa funkcije f, kao na prethodnoj slici.

Uvrštavanjem podataka o predznaku prve i druge derivacije u

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

dobivamo da je desna strana veća od 0, pa je i $\alpha > x_0$.

Regula falsi — red konvergencije

Što sve slijedi iz $\alpha > x_0$?

Po pretpostavci, funkcija f monotono raste na $[a, b]$, pa je

$$f(a) < f(x_0) < f(\alpha) = 0 < f(b).$$

To znači da treba "pomaknuti" a za sljedeći korak metode, jer $f(a)$ i $f(x_0)$ imaju isti predznak, tj. $\alpha \in [x_0, b]$. Dakle, imamo

- $a_1 := x_0$, a b ostaje fiksan, $b_1 := b$.

Potpuno isto će se dogoditi i u svim narednim koracima. Drugim riječima,

- aproksimacije x_n neprestano ostaju lijevo od nultočke α,
- tj. pomiče se lijevi rub intervala, $a_n := x_{n-1}$,
- a desni rub b ostaje fiksan, $b_n := b_{n-1} = \cdots = b$.

Regula falsi — red konvergencije

U proizvoljnoj iteraciji — kad računamo x_n, uz $\alpha \in [a_n, b_n]$, relacija za grešku ima oblik

$$\alpha - x_n = (b_n - \alpha)\,(\alpha - a_n)\,\frac{f''(\zeta_n)}{2f'(\xi_n)}.$$

Kad uvrstimo prethodne zaključke $a_n = x_{n-1}$ i $b_n = b$, izlazi

$$\alpha - x_n = \left((b - \alpha)\,\frac{f''(\zeta_n)}{2f'(\xi_n)}\right)(\alpha - x_{n-1}).$$

Uzimanjem apsolutnih vrijednosti slijeva i zdesna, slijedi da

- u ovom slučaju, regula falsi konvergira linearno.

Zadatak. Dokažite da je apsolutna vrijednost faktora u prvoj (velikoj) zagradi strogo manja od 1 ($\Rightarrow$ konvergencija metode).

Regula falsi — red konvergencije

Za metodu bisekcije dobili smo sličnu relaciju za grešku, samo je faktor bio $1/2$. Usporedbom izraza za greške, vidimo da

- nije teško konstruirati primjere kad je metoda bisekcije brža no regula falsi.

Probajte naći takav primjer!

Napomena. Sasvim analogno se analiziraju i ostala tri slučaja za predznake f' i f'' na $[a, b]$. Na primjer, ako je f konveksna, ali monotono pada, tj. $f'' > 0$ i $f' < 0$,

- aproksimacija x_n je uvijek desno od α,
- a uvijek se pomiče desni rub b.

U svim slučajevima izlazi da regula falsi konvergira linearno.

Konstrukcija iterativnih metoda za nalaženje nultočaka

Konstrukcija iterativnih metoda

Kod konstrukcije metode bisekcije i regule falsi, koristili smo dvije startne pretpostavke o funkciji f i intervalu $[a, b]$ na kojem radimo. Te pretpostavke su:

- "glatkoća" — funkcija f je neprekidna na intervalu $[a, b]$,
- "lokacija nultočke" — u rubovima intervala vrijedi

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

  tj. funkcijske vrijednosti u rubovima intervala imaju različit predznak.

Uz prvu, druga pretpostavka osigurava da f ima bar jednu nultočku u $[a, b]$

- i ove rubne točke a, b su bitne za start iteracija.

Obje metode sigurno konvergiraju, ali sporo — linearno.

Konstrukcija iterativnih metoda

U nastavku, krećemo od drugačijih pretpostavki za konstrukciju iterativnih metoda.

Ako funkcija f ima veću glatkoću na nekoj domeni,

- smijemo koristiti vrijednosti prve derivacije f' u točkama,
- može i vrijednosti viših derivacija, ako postoje.

Do sada smo veću glatkoću koristili samo kod ocjene greške.

Što se lokacije tiče, krećemo od slabijih pretpostavki.

- Za konstrukciju iterativne metode — ne pretpostavljamo ništa, tj. lokacija nultočke nije bitna.

Kod analize konvergencije metode — lokacija nultočke, naravno, ima ključnu ulogu.

Konstrukcija iterativnih metoda

Opće ideje za konstrukciju iterativnih metoda za nalaženje nultočaka funkcije su sljedeće.

- Jednu ili više startnih točaka možemo, bar u principu, izabrati bilo gdje.

Zatim, počevši od tih startnih točaka,

- iterativno generiramo neki niz aproksimacija (= metoda).

U svakoj iteraciji, novu aproksimaciju generiramo tako da je

$$\text{aproksimacija nultočke} = \text{nultočka aproksimacije},$$

s tim da se aproksimacija određuje na osnovu jedne ili više prethodnih točaka — tzv. iteracijskom funkcijom.

Ideja iteracije je slična onoj kod integracijskih formula.

Konstrukcija iterativnih metoda

Na primjer, kod Newtonove metode, funkciju aproksimiramo

- tangentom u jednoj (zadnjoj) točki,

a kod metode sekante,

- sekantom kroz dvije (zadnje) točke.

Kod ovakvog pristupa, gubimo sigurnu konvergenciju, tj.

- može se dogoditi da metode ne konvergiraju, ovisno o startu iteracija.

Međutim, ako konvergiraju,

- konvergencija je brža — kad smo dovoljno blizu nultočke.

Za pojedine metode, navest ćemo i neke uvjete koji osiguravaju konvergenciju (slično kao kod regule falsi).

Metoda tangente (Newtonova metoda)

Ideja metode tangente (Newtonove metode)

Pretpostavimo da je funkcija f barem

- neprekidno derivabilna na nekoj domeni,
- idealno — na cijelom $\mathbb{R}$.

Nadalje, neka je zadana, ili nekako izabrana,

- početna točka x_0.

Ideja metode tangente je

- povući tangentu na graf funkcije f u točki $(x_0, f(x_0))$,

i definirati novu aproksimaciju x_1

- u točki gdje ta tangenta siječe os x — ako takva točka x_1 postoji. U protivnom, treba uzeti neku drugu točku x_0.

Ideja metode tangente (Newtonove metode)

Drugim riječima, funkciju f

- aproksimiramo pravcem — tangentom u točki $(x_0, f(x_0))$ (to je najbolja linearna aproksimacija funkcije f u okolini točke x_0),

a nepoznatu nultočku funkcije f

- aproksimiramo nultočkom x_1 tog pravca — te tangente na graf funkcije f u zadanoj točki (ako x_1 postoji).

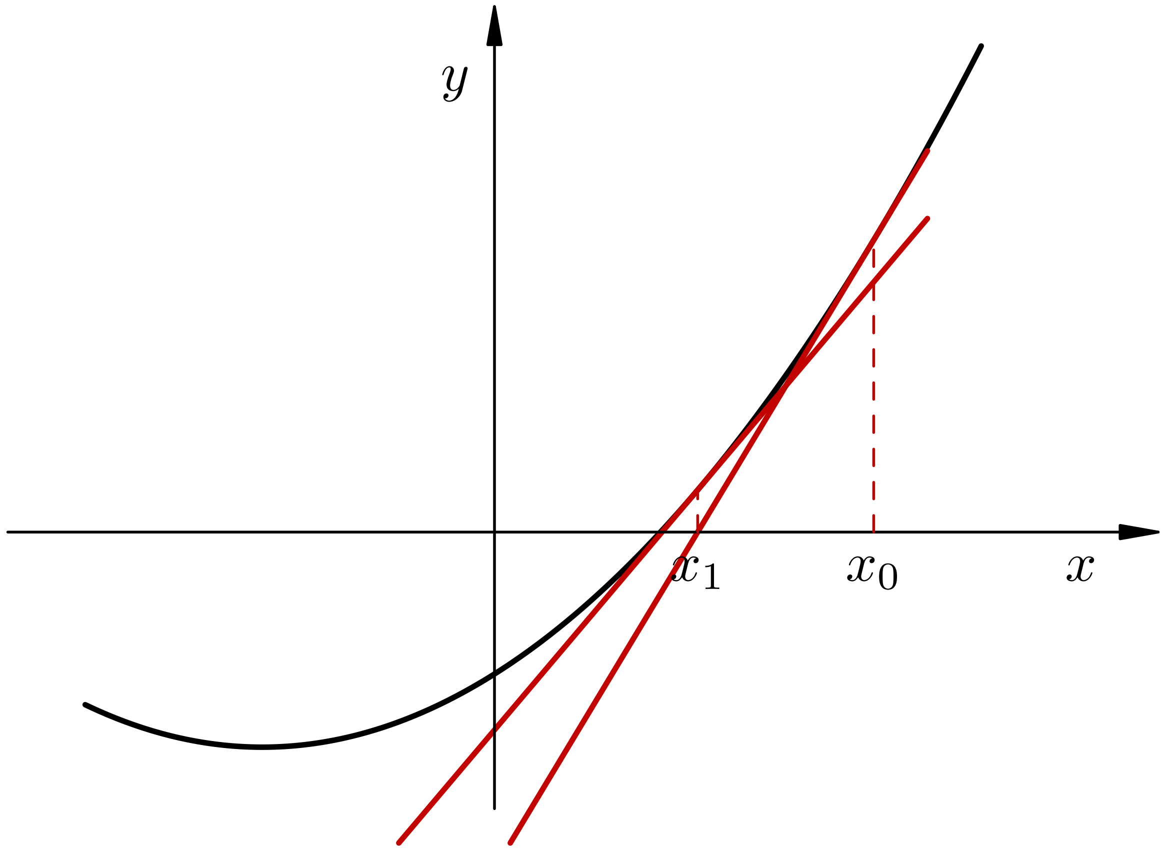
Isti postupak možemo ponoviti u točki x_1 i dobiti sljedeću točku x_2. Kad ovaj postupak iteriramo — induktivno ponovimo u svakoj sljedećoj točki x_n, za $n \geq 0$, dobivamo

- metodu tangente ili Newtonovu metodu

za nalaženje nultočke funkcije f.

Newtonova metoda — grafički

Grafički, Newtonova metoda za nalaženje nultočke izgleda ovako



Geometrijski izvod Newtonove metode

Geometrijski izvod metode je jednostavan.

- U točki x_n napišemo jednadžbu tangente i pogledamo gdje tangenta siječe os x.

Jednadžba tangente je

$$y - f(x_n) = f'(x_n)(x - x_n).$$

Iz zahtjeva $y = 0$ za $x = x_{n+1}$, izlazi da je nova aproksimacija x_{n+1} dana izrazom

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n \geq 0.$$

Za računanje je dovoljno pretpostaviti da $f'(x_n)$ postoji (neprekidnost nije bitna) i da je $f'(x_n) \neq 0$ u svim točkama x_n.

Analitički izvod Newtonove metode

Do Newtonove metode može se doći i analitički,

- ali uz malo jače pretpostavke,
- iz kojih onda slijedi i izraz za grešku.

Neka je α neka nultočka funkcije f i pretpostavimo da je

- f dva puta neprekidno derivabilna na nekom intervalu oko α.

Tj. pretpostavljamo da je $f \in C^2(I)$, gdje je I segment takav da je $\alpha \in I$.

Neka je $x_n \in I$ bilo koja točka — neka aproksimacija za α.

Onda funkciju f možemo razviti u Taylorov red oko x_n, do uključivo prvog člana, a kvadratni član je greška.

Analitički izvod Newtonove metode

Za $x \in I$, dobivamo

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

pri čemu je ξ_n između x i x_n, tj. $\xi_n \in I$.

Uvrštavanjem nultočke $x = \alpha \in I$, dobivamo

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

Uz pretpostavku $f'(x_n) \neq 0$, dijeljenjem i prebacivanjem izlazi

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

Analitički izvod Newtonove metode

Ako pretpostavimo da je aproksimacija x_n dovoljno blizu α,

- tj. da je $|\alpha - x_n|$ mali,
- onda očekujemo da je $(\alpha - x_n)^2$ još puno manji.

Zato možemo "zanemariti" zadnji član u relaciji

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

i definirati novu aproksimaciju x_{n+1} kao početak desne strane

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

s idejom da je $x_{n+1} \approx \alpha$ još bolja aproksimacija od x_n.

Relacija za grešku Newtonove metode

Oduzimanjem odmah dobivamo i izraz koji veže greške dviju susjednih iteracija

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

Iz ovog "izvoda" samo očekujemo da se greška "smanjuje", ali to tek treba dokazati, uz odgovarajuće pretpostavke!

Sasvim općenito, to ne mora vrijediti!

Čak i kad startamo u nekom intervalu $I = [a, b]$ koji sadrži nultočku α, bez dodatnih pretpostavki — nema garancije

- da aproksimacije ostaju u tom intervalu I,
- a kamo li da konvergiraju (v. primjere kasnije).

Napomene o konvergenciji Newtonove metode

Dakle, Newtonova metoda ne mora konvergirati prema nultočki!

S druge strane, dokazat ćemo da ovaj zaključak vrijedi

- kad je x_n, odnosno, startna točka x_0 — dovoljno blizu α.

Takva konvergencija se obično naziva

- lokalna konvergencija metode.

Općenito, zaključke o konvergenciji metode (uz dovoljno jake pretpostavke) možemo podijeliti u tri grupe:

- brzina (lokalne) konvergencije — ako niz konvergira,
- lokalna konvergencija metode — uz start dovoljno blizu,
- globalna konvergencija metode na nekom intervalu I.

Brzina konvergencije Newtonove metode

Brzina lokalne konvergencije izlazi direktno iz izraza za grešku u susjednim iteracijama.

Teorem. Neka je α jednostruka nultočka funkcije f i pretpostavimo da je $f \in C^2(I)$, na nekom segmentu I koji sadrži tu nultočku α.

Ako niz aproksimacija x_n, generiran Newtonovom metodom, konvergira prema α,

- onda je brzina konvergencije (barem) kvadratna,

i na limesu vrijedi

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

Brzina konvergencije Newtonove metode

Dokaz. Za greške u susjednim iteracijama vrijedi

$$\alpha - x_{n+1} = -(\alpha - x_n)^2 \, \frac{f''(\xi_n)}{2f'(x_n)},$$

s tim da je ξ_n između x_n i α.

Po pretpostavci, niz x_n konvergira prema α. Onda mora vrijediti i $\xi_n \to \alpha$.

Iz $f \in C^2(I)$ slijedi da su f' i f'' neprekidne na I, pa dobivamo

$$f'(x_n) \to f'(\alpha), \quad f''(\xi_n) \to f''(\alpha).$$

Na kraju, iz pretpostavke da je α jednostruka nultočka, slijedi $f'(\alpha) \neq 0$.

Zato smijemo prijeći na limes $x_n \to \alpha$ u relaciji za grešku.

Brzina konvergencije Newtonove metode

Prijelazom na limes $x_n \to \alpha$ dobivamo

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

Odavde čitamo da je Newtonova metoda, kad konvergira,

- (barem) kvadratno konvergentna.

Ako je $f''(\alpha) = 0$, konvergencija može biti i brža od kvadratne!

Ipak, treba biti oprezan, jer prethodni zaključci vrijede

- samo ako je $f'(\alpha) \neq 0$, tj. ako je α jednostruka nultočka.

Za višestruke nultočke ovo ne vrijedi, jer

- konvergencija može biti i samo linearna (v. kasnije).

Lokalna konvergencija Newtonove metode

Sljedeći rezultat o konvergenciji Newtonove metode je

- lokalna konvergencija za jednostruke nultočke α,

uz pretpostavku da je funkcija $f \in C^2(I)$, na nekom segmentu I koji sadrži tu nultočku α.

Neformalno rečeno, tvrdnja kaže sljedeće:

- ako je početna točka x_0 dovoljno blizu nultočke α,
- onda je Newtonova metoda

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}, \quad n \geq 0,$$

dobro definirana — tj. vrijedi $f'(x_n) \neq 0$, za sve $n \geq 0$,

- i ovaj niz konvergira prema α (i to barem kvadratno).

Lokalna konvergencija Newtonove metode

Prava formulacija je malo složenija, jer precizno opisuje što znači “dovoljno blizu”.

Teorem. Neka je α jednostruka nultočka funkcije f i neka je

$$I_\varepsilon := \{x \in \mathbb{R} \mid |x - \alpha| \leq \varepsilon\}$$

segment radijusa ε oko α (duljina tog segmenta je 2ε).

Pretpostavimo da je $f \in C^2(I_\varepsilon)$, za sve dovoljno male $\varepsilon > 0$. Definiramo brojeve

$$m_1(\varepsilon) := \min_{x \in I_\varepsilon} |f'(x)|, \qquad M_2(\varepsilon) := \max_{x \in I_\varepsilon} |f''(x)|,$$

i, na kraju,

$$M(\varepsilon) := \frac{M_2(\varepsilon)}{2m_1(\varepsilon)}.$$

Lokalna konvergencija Newtonove metode

Onda postoji $\varepsilon > 0$, toliko mali da vrijedi

$$\varepsilon M(\varepsilon) < 1.$$

Nadalje, za svaki takav ε, za bilo koju startnu točku $x_0 \in I_\varepsilon$,

- Newtonova metoda je dobro definirana,
- i konvergira prema jedinoj nultočki $\alpha \in I_\varepsilon$.
  (Već znamo da je tada konvergencija barem kvadratna.)

Dokaz. Ide u 4 koraka i počiva na sljedeće dvije pretpostavke:

- α je jednostruka nultočka, tj. vrijedi $f'(\alpha) \neq 0$,
- po definiciji I_ε, za svaku točku $x \in I_\varepsilon$ vrijedi $|x - \alpha| \leq \varepsilon$.

To ćemo iskoristiti nekoliko puta!

Lokalna konvergencija Newtonove metode

1\. korak. Pokažimo da postoji $\varepsilon > 0$, takav da je

$$\varepsilon M(\varepsilon) < 1.$$

Iz pretpostavke da je $f \in C^2(I_\varepsilon)$, za sve dovoljno male $\varepsilon > 0$, slijedi da je

$$M(\varepsilon) := \frac{M_2(\varepsilon)}{2m_1(\varepsilon)}$$

rastuća funkcija od ε (za dovoljno male $\varepsilon > 0$), jer se

- $\max |f''(x)|$ u brojniku i $\min |f'(x)|$ u nazivniku,
- uzimaju po sve većim segmentima I_ε oko α,
- tako da brojnik raste, a nazivnik pada (kad ε raste).

Naravno, ako f' ima nultočku u nekom segmentu I_ε, onda je $m_1(\varepsilon) = 0$, odnosno, $M(\varepsilon) = \infty$, i to treba izbjeći!

Lokalna konvergencija Newtonove metode

Međutim, zbog $f'(\alpha) \neq 0$ i neprekidnosti f' oko α (za sve dovoljno male $\varepsilon > 0$), sigurno postoji (dovoljno mali) $\varepsilon_0 > 0$,

- takav da je $f'(x) \neq 0$, za svaki $x \in I_{\varepsilon_0}$,
- pa onda vrijedi $m_1(\varepsilon_0) > 0$, odnosno, $M(\varepsilon_0) < \infty$.

No, čim je $M(\varepsilon_0)$ "konačan", onda postoji i $\varepsilon > 0$, takav da je $\varepsilon \leq \varepsilon_0$, i vrijedi

$$\varepsilon M(\varepsilon_0) < 1.$$

Zato što $M(\varepsilon)$ raste po ε, iz $\varepsilon \leq \varepsilon_0$ izlazi da je i $\varepsilon M(\varepsilon) < 1$.

Osim toga, iz $m_1(\varepsilon) \geq m_1(\varepsilon_0) > 0$, slijedi da je $f'(x) \neq 0$, za svaki $x \in I_\varepsilon$. Dodatno, iz neprekidnosti f', vidimo da

- $f'(x)$ ima isti predznak kao i $f'(\alpha)$, za svaki $x \in I_\varepsilon$.

Lokalna konvergencija Newtonove metode

2. korak. Pokažimo da je α jedina nultočka funkcije f u I_ε.

Upravo smo pokazali da je $f'(x) \neq 0$ na I_ε, pa je funkcija f monotona na I_ε, tj. može imati najviše jednu nultočku.

Dakle, α je jedina nultočka od f u I_ε.

Digresija: Isti zaključak se može dobiti i direktno, iz $f'(\alpha) \neq 0$.

Iz Taylorovog razvoja f oko α, za bilo koji $x \in I_\varepsilon$, imamo

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

za neki ξ između α i x, pa je i $\xi \in I_\varepsilon$.

Znamo da je $f(\alpha) = 0$ i $f'(\alpha) \neq 0$, pa izlučimo $f'(\alpha)$.

Lokalna konvergencija Newtonove metode

Dobivamo

$$f(x) = f'(\alpha)(x-\alpha) \cdot \left(1 + \frac{f''(\xi)}{2f'(\alpha)}\,(x-\alpha)\right).$$

Prvi faktor je, očito, različit od nule. Za drugi član u trećem faktoru imamo ocjenu

$$\left|\frac{f''(\xi)}{2f'(\alpha)}\,(x-\alpha)\right| \leq \frac{M_2(\varepsilon)}{2m_1(\varepsilon)}\,\varepsilon = \varepsilon M(\varepsilon) < 1,$$

pa za treći faktor vrijedi

$$1 + \frac{f''(\xi)}{2f'(\alpha)}\,(x-\alpha) > 0.$$

Dakle, $f(x) = 0$, za neki $x \in I_\varepsilon$, ako i samo ako je drugi faktor jednak nuli, tj. za $x = \alpha$.

Lokalna konvergencija Newtonove metode

3. korak. Neka je $x_0 \in I_\varepsilon$ bilo koja startna točka. Onda je $f'(x_0) \neq 0$, pa je prvi korak Newtonove metode dobro definiran

$$x_1 := x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Kad je $n = 0$, relacija za greške dviju susjednih iteracija glasi

$$\alpha - x_1 = -(\alpha - x_0)^2 \, \frac{f''(\xi_0)}{2f'(x_0)},$$

gdje je ξ_0 između x_0 i α, pa vrijedi i $\xi_0 \in I_\varepsilon$.

Za drugi faktor na desnoj strani onda vrijedi ocjena

$$\left| \frac{f''(\xi_0)}{2f'(x_0)} \right| \leq \frac{M_2(\varepsilon)}{2m_1(\varepsilon)} = M(\varepsilon).$$

Lokalna konvergencija Newtonove metode

Zatim, imamo redom,

$$\begin{aligned} |\alpha - x_1| &\le |\alpha - x_0|^2 \, M(\varepsilon) \quad (\text{iskoristimo } |\alpha - x_0| \le \varepsilon) \\ &\le |\alpha - x_0| \, \varepsilon M(\varepsilon) \quad (\text{zbog } \varepsilon M(\varepsilon) < 1) \\ &< |\alpha - x_0| \le \varepsilon, \end{aligned}$$

pa je $|\alpha - x_1| \le \varepsilon$, što dokazuje da je $x_1 \in I_\varepsilon$.

Dakle, ako startamo u bilo kojoj točki $x_0 \in I_\varepsilon$, onda

- sljedeća aproksimacija x_1 po Newtonovoj metodi, ostaje unutar segmenta I_ε.

Induktivnom primjenom ovog argumenta (x_0 je proizvoljan), dobivamo da isto vrijedi i za svaku sljedeću aproksimaciju x_n.

Lokalna konvergencija Newtonove metode

Dakle, za svaki $n \geq 0$, u točki x_n vrijedi $f'(x_n) \neq 0$, pa je sljedeća aproksimacija dobro definirana relacijom

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}, \qquad n \geq 0,$$

i ostaje unutar polaznog segmenta, tj. vrijedi $x_{n+1} \in I_\varepsilon$.

4. korak. Preostaje još samo dokazati konvergenciju niza (x_n) prema nultočki α.

Relacija koja veže greške dviju susjednih iteracija x_n i x_{n+1} je

$$\alpha - x_{n+1} = -(\alpha - x_n)^2 \, \frac{f''(\xi_n)}{2f'(x_n)},$$

gdje je ξ_n između x_n i α, pa je $\xi_n \in I_\varepsilon$.

Lokalna konvergencija Newtonove metode

Za drugi faktor na desnoj strani opet vrijedi ocjena

$$\left|\frac{f''(\xi_n)}{2f'(x_n)}\right| \leq \frac{M_2(\varepsilon)}{2m_1(\varepsilon)} = M(\varepsilon).$$

Slično kao u prethodnom koraku, dobivamo redom

$$\begin{aligned}
|\alpha - x_{n+1}| &\leq |\alpha - x_n|^2 \, M(\varepsilon) \qquad (\text{iskoristimo } |\alpha - x_n| \leq \varepsilon) \\
&\leq \varepsilon M(\varepsilon) \, |\alpha - x_n| \\
&\leq \ldots \ (\text{induktivno spuštamo } n \text{ do nule}) \\
&\leq \big(\varepsilon M(\varepsilon)\big)^{n+1} \, |\alpha - x_0|.
\end{aligned}$$

Zbog $\varepsilon M(\varepsilon) < 1$, odavde odmah slijedi da $x_n \to \alpha$ kad $n \to \infty$ ($=$ "geometrijska" konvergencija s faktorom $c = \varepsilon M(\varepsilon)$). ■

Lokacija i osiguranje konvergencije — problem

Ovaj rezultat o lokalnoj konvergenciji teško možemo iskoristiti u praksi — za osiguranje konvergencije Newtonove metode

- bar iz neke startne točke x_0, koju znamo naći (što bi bilo sasvim dovoljno za nalaženje nultočke).

Problem = traženi ε iz tvrdnje ne znamo i nije ga lako naći!

Pretpostavimo da smo locirali nultočku funkcije f u segmentu $[a, b]$ i znamo da je $f \in C^2[a, b]$. Neka je "globalno" na $[a, b]$

$$M_2 = \max_{x \in [a,b]} |f''(x)|, \quad m_1 = \min_{x \in [a,b]} |f'(x)|.$$

Pretpostavimo još da je

- f strogo monotona na $[a, b]$, što je ekvivalentno s $m_1 > 0$.

Tada f ima jedinstvenu jednostruku nultočku α u $[a, b]$.

Lokacija i osiguranje konvergencije — problem

To znači da imamo sve osnovne pretpostavke prethodnog teorema o lokalnoj konvergenciji Newtonove metode

- i još znamo da je $f' \neq 0$ na $[a, b]$.

Umjesto "lokalnog" $M(\varepsilon)$, izračunamo "globalnu" veličinu

$$M := \frac{M_2}{2m_1}.$$

Sad možemo pokušati izabrati ε tako da vrijedi

- $I_\varepsilon \subseteq [a, b]$ — pa mora biti $M \geq M(\varepsilon)$, jer $M(\varepsilon)$ raste s ε,
- i da je $\varepsilon M < 1$ — to onda osigurava i $\varepsilon M(\varepsilon) < 1$.

Našalost, prvi uvjet je problem!

Lokacija i osiguranje konvergencije — problem

Primjer. Uzmemo ε tako "globalno" vrijedi

$$\varepsilon M < 1.$$

Nažalost, bez dodatnih uvjeta, još uvijek se može dogoditi da

- pripadni interval I_ε nije sadržan u $[a, b]$, već "viri" preko ruba intervala $[a, b]$.

To se sigurno događa ako je $\varepsilon > (b - a)/2$. Inače, dovoljno je da nultočka α leži blizu jednog ruba intervala — bliže od ε.

- Tada ne mora vrijediti $M \geq M(\varepsilon)$, pa ni $\varepsilon M(\varepsilon) < 1$.

Onda više nemamo garanciju lokalne konvergencije, tj.

- čak i da uzmemo početnu iteraciju $x_0 \in I_\varepsilon \cap [a, b]$,
- već prva sljedeća iteracija x_1 može izaći izvan $[a, b]$, čak izvan I_ε (taj interval ne znamo, jer ne znamo α).

Lokacija i osiguranje konvergencije — problem

Naime, 3. korak je ključni dio teorema o lokalnoj konvergenciji!

Poanta. Treba uzeti još manji ε tako da bude $I_\varepsilon \subseteq [a, b]$, tj. tako da vrijedi još i

$$a + \varepsilon \le \alpha \le b - \varepsilon.$$

No, baš to je problem — jer ne znamo gdje je α.

Kad ima neke koristi? Ako nađemo ε takav da je

$$\varepsilon M < 1 \quad \text{i} \quad \varepsilon < \frac{b - a}{2},$$

i još pokažemo da je $\alpha \in [a + \varepsilon, b - \varepsilon]$ — na primjer,

- provjerom da $f(a)$ i $f(a + \varepsilon)$ imaju isti znak, te $f(b)$ i $f(b - \varepsilon)$ imaju isti znak,

onda znamo da je $I_\varepsilon \subseteq [a, b]$ i da vrijedi $\varepsilon M(\varepsilon) < 1$.

Problem — “lokalnost” lokalne konvergencije

Tada imamo sigurnu lokalnu konvergenciju Newtonove metode

- za bilo koju startnu točku $x_0 \in I_\varepsilon$.

Međutim, još uvijek ne znamo gdje treba startati, jer ne znamo gdje se nalazi “pravi” I_ε unutar $[a, b]$. Što sad?

Možemo napraviti pretragu intervala $[a, b]$ s korakom $h \leq 2\varepsilon$, tj. “testirati” startne točke oblika

$$x_0 = a + kh, \quad k = 0, 1, \ldots,$$

Bar jedna takva startna točka x_0 daje sigurnu konvergenciju Newtonove metode. Oprez: za svaku takvu točku x_0 treba

- pažljivo provjeravati sve sljedeće iteracije x_n — ostaju li unutar $[x_0 - \varepsilon, x_0 + \varepsilon]$, ili, barem, unutar $[a, b]$.

Finale: ova pretraga je spora — bisekcija je, vjerojatno, brža!

Ocjena greške iteracija u Newtonovoj metodi

Pretpostavimo da sve iteracije x_n leže unutar intervala $[a, b]$. Onda možemo dobiti i ocjenu

- lokalne greške susjednih iteracija u Newtonovoj metodi,

u terminima veličina M_2 i m_1 (na tom intervalu $[a, b]$).

Iz ranije relacije za grešku

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

gdje je ξ_{n-1} između nultočke α i x_{n-1}, odmah slijedi

$$|\alpha - x_n| \leq \frac{M_2}{2m_1} \, (\alpha - x_{n-1})^2.$$

Ova ocjena nije naročito korisna za praksu, jer nultočku α ne znamo. Tražimo ocjenu preko veličina koje znamo izračunati.

Ocjena greške iteracija u Newtonovoj metodi

Za dvije susjedne iteracije x_{n-1} i x_n u Newtonovoj metodi, također, vrijedi veza preko Taylorove formule

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

pri čemu je ξ_{n-1} između x_{n-1} i x_n.

Po definiciji iteracija u Newtonovoj metodi, vrijedi

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

pa je

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

Ocjena greške iteracija u Newtonovoj metodi

Sad iskoristimo pretpostavku da je x_{n-1}, $x_n \in [a, b]$, pa onda mora biti i $\xi_{n-1} \in [a, b]$. Dobivamo

$$|f(x_n)| \leq \frac{M_2}{2} (x_n - x_{n-1})^2.$$

Kao i kod metode bisekcije, ako je $m_1 > 0$, onda vrijedi ocjena

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

Kombinacijom ovih ocjena dobivamo ocjenu greške za svaku iteraciju x_n u Newtonovoj metodi — preko razlike $x_n - x_{n-1}$,

$$|\alpha - x_n| \leq \frac{M_2}{2m_1} (x_n - x_{n-1})^2.$$

Ocjena greške iteracija u Newtonovoj metodi

Ova ocjena se može iskoristiti. Ako je ε tražena točnost, onda zahtjev

$$\frac{M_2}{2m_1}(x_n - x_{n-1})^2 \le \varepsilon$$

garantira da je $|\alpha - x_n| \le \varepsilon$, do na greške zaokruživanja.

Pripadni test zaustavljanja iteracija u Newtonovoj metodi je

$$|x_n - x_{n-1}| \le \sqrt{\frac{2m_1\varepsilon}{M_2}}.$$

Naravno, uz ovaj, možemo koristiti i raniji test zaustavljanja

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

Veznik između ova dva testa je ili, tj. pitamo je li ispunjen jedan ili drugi.

Globalna konvergencija Newtonove metode

U analizi konvergencije i ocjenama greške koristili smo pretpostavku da je

- f strogo monotona na $[a, b]$,
- tj. da prva derivacija f' ima fiksni predznak na $[a, b]$.

Ako i druga derivacija ima fiksni predznak na tom intervalu, onda možemo dobiti i

- globalnu konvergenciju Newtonove metode,
- uz odgovarajući izbor startne točke x_0,

slično kao kod regule falsi.

Globalna konvergencija Newtonove metode

Teorem. Neka je $f \in C^2[a,b]$ i neka je $f(a) \cdot f(b) < 0$.

Ako prva i druga derivacija f' i f'' nemaju nultočku u $[a,b]$, tj.

- ako f' i f'' imaju konstantni predznak na $[a,b]$,

onda Newtonova metoda konvergira prema

- jedinstvenoj jednostrukoj nultočki α, funkcije f u $[a,b]$,

i to za svaku startnu aproksimaciju $x_0 \in [a,b]$, za koju vrijedi

$$f(x_0) \cdot f''(x_0) > 0.$$

Dokaz. Gledamo samo jedan od četiri moguća slučaja za predznake f' i f''.

U ostalim slučajevima, dokaz ide potpuno analogno.

Globalna konvergencija Newtonove metode

Uzmimo, na primjer, da je $f' > 0$ i $f'' > 0$ na $[a, b]$, tj. da je

- f monotono rastuća i konveksna na $[a, b]$.

U tom slučaju, jer f raste, mora biti $f(a) < 0$ i $f(b) > 0$.

Zbog $f'' > 0$, startna aproksimacija x_0

- mora zadovoljavati $f(x_0) > 0$, tj. $\alpha < x_0$, jer f raste.

U praksi možemo uzeti $x_0 = b$, jer je to jedina točka za koju sigurno znamo da vrijedi $f(x_0) > 0$.

Neka je $(x_n, n \in \mathbb{N}_0)$ niz iteracija generiran Newtonovom metodom iz bilo koje startne točke x_0, za koju je $f(x_0) > 0$,

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

Globalna konvergencija Newtonove metode

Za početak, znamo da je $x_0 > \alpha$. Tvrdimo da je

- $\alpha < x_n \leq x_0$ za svaki $n \in \mathbb{N}_0$.

Dokaz ide indukcijom, a bazu već imamo.

Za korak indukcije, pretpostavimo da je $\alpha < x_n \leq x_0$. Onda je

$$f(x_n) > f(\alpha) = 0.$$

Osim toga, jer f raste, znamo da je $f'(x_n) > 0$, pa dobivamo

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} < x_n \leq x_0,$$

što pokazuje da niz (x_n) monotono pada.

Globalna konvergencija Newtonove metode

Iz Taylorove formule oko x_n je

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

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

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

odakle slijedi

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

Dakle, niz (x_n) je odozdo ograničen s α i monotono pada, pa postoji limes

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

Odmah znamo i da je $\alpha \leq \alpha' \leq x_0$, tj. vrijedi $\alpha' \in [a, b]$.

Globalna konvergencija Newtonove metode

Prijelazom na limes u formuli za Newtonove iteracije dobivamo

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

Zbog $f'(\alpha') \neq 0$ (jer $\alpha' \in [a, b]$), odavde slijedi $f(\alpha') = 0$.

No, znamo da f ima

- jedinstvenu nultočku α u intervalu $[a, b]$,

pa mora biti $\alpha = \alpha'$. Dakle, niz (x_n) konvergira baš prema α.

Preostala tri slučaja za predznake prve i druge derivacije dokazuju se potpuno analogno. ∎

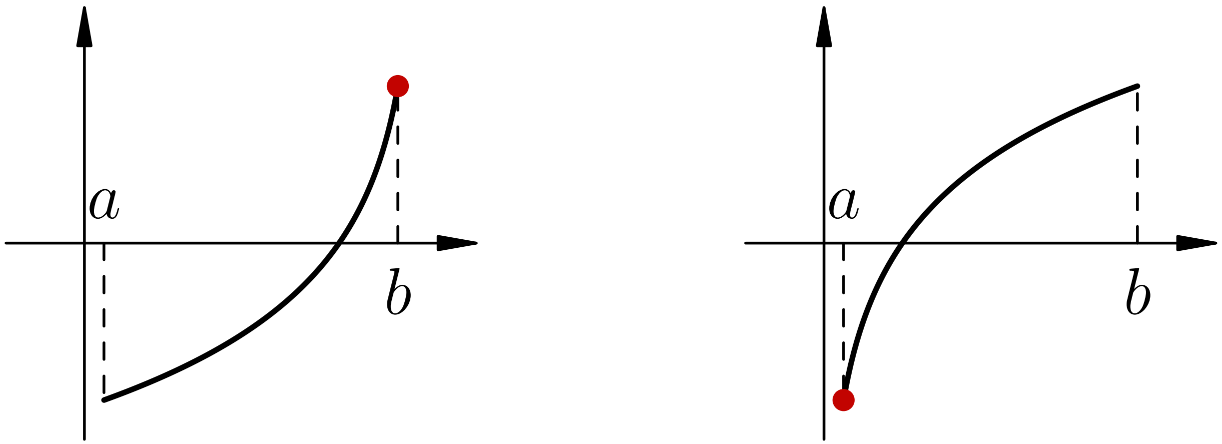
Uvjet $f(x_0) \cdot f''(x_0) > 0$, na izbor startne točke u prethodnom teoremu, ima vrlo jednostavnu geometrijsku interpretaciju.

Izbor startne točke za Newtonovu metodu

Ako pogledamo graf funkcije f na $[a, b]$, startnu točku x_0

- treba odabrati na "strmijoj" strani grafa funkcije.

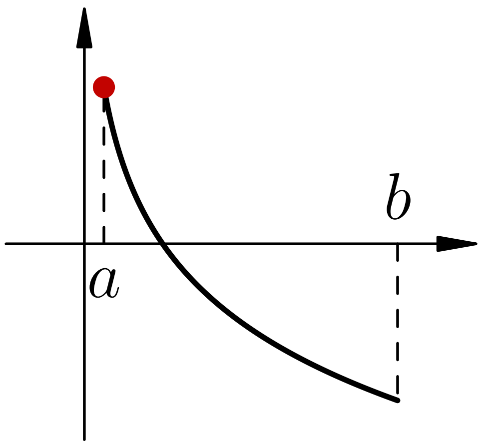
Izbor startne točke x_0 — ako je $f' > 0$, tj. f raste.



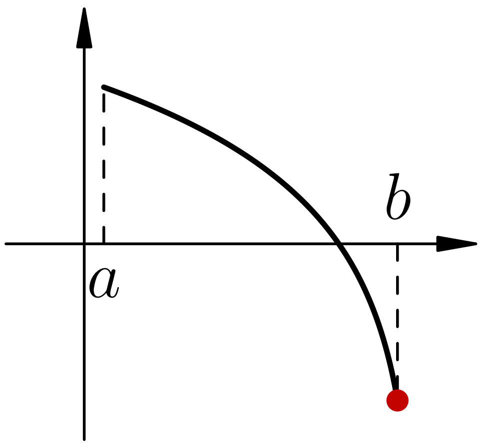
$f'' > 0$ $\qquad$ $f'' < 0$

Izbor startne točke za Newtonovu metodu

Izbor startne točke x_0 — ako je $f' < 0$, tj. f pada



$f'' > 0$



$f'' < 0$

Newtonova metoda — komentari

Prednost:

- brza $=$ kvadratna konvergencija, za jednostruke nultočke.

Potencijalna mana: osim vrijednosti funkcije,

- trebamo i vrijednost prve derivacije u svakoj iteraciji.

Ako se f' komplicirano računa,

- Newtonova metoda može biti sporija od metode sekante,
- iako ima veći red konvergencije (v. malo dalje).

Metoda sekante

Uvodno o metodi sekante

U Newtonovoj metodi koristimo

- tangentu u točki x_0 kao aproksimaciju funkcije f.

Ako ne znamo derivaciju f' funkcije f, ili se ona teško računa, onda možemo

- tangentu u točki x_0 aproksimirati sekantom kroz dvije startne točke x_0 i x_1,

što odgovara aproksimaciji derivacije $f'(x_0)$ podijeljenom razlikom

$$f'(x_0) \approx \frac{f(x_1) - f(x_0)}{x_1 - x_0} = f[x_0, x_1].$$

Tako dobivamo metodu sekante.

Ideja metode sekante

Počinjemo s dvije početne točke x_0 i x_1. Poredak je bitan!

Ideja metode sekante je

- povući sekantu grafa funkcije f kroz točke $(x_0, f(x_0))$ i $(x_1, f(x_1))$,

i definirati novu aproksimaciju x_2

- u točki gdje ta sekanta siječe os x (ako x_2 postoji).

Postupak nastavljamo povlačenjem sekante

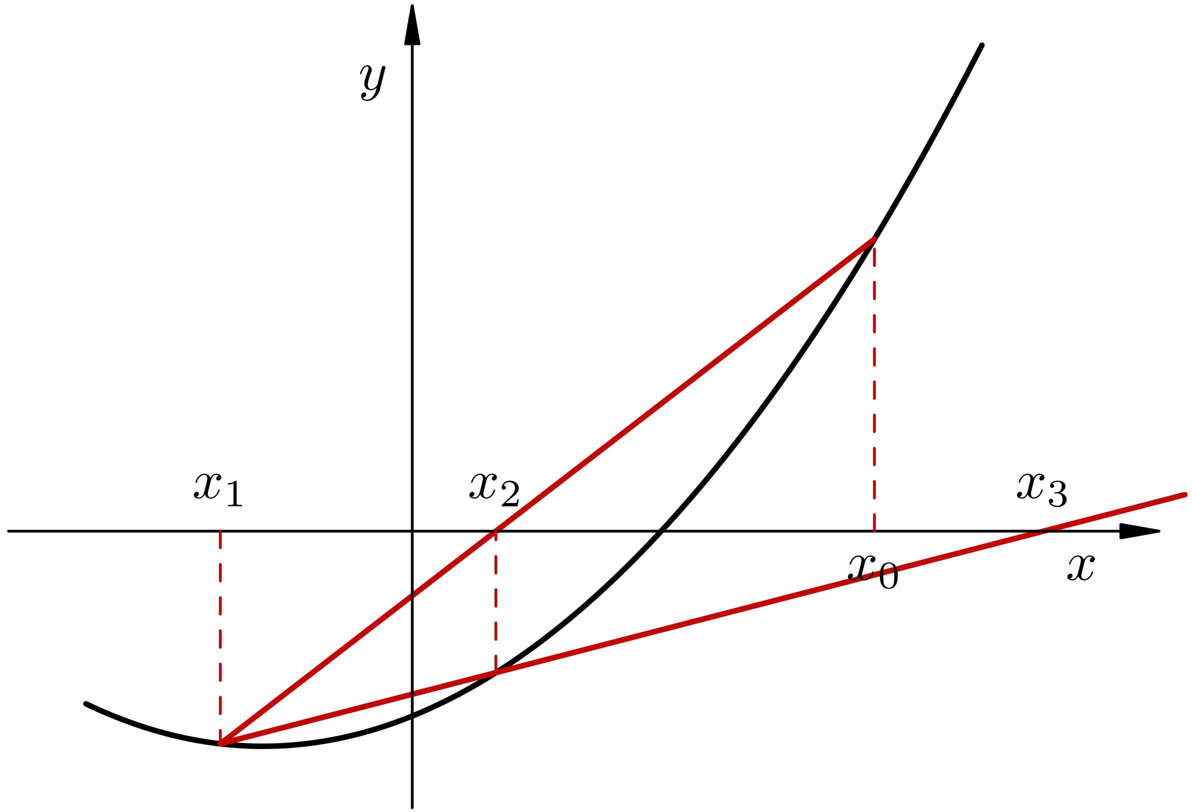
- kroz posljednje dvije točke $(x_1, f(x_1))$ i $(x_2, f(x_2))$,

i tako redom.

Napomena. Tu je ključna razlika od regule falsi — tamo se jedna početna točka drži fiksnom.

Metoda sekante — grafički

Grafički, metoda sekante za nalaženje nultočke izgleda ovako



Primijetite da je treća iteracija izašla izvan početnog intervala, pa metoda sekante ne mora konvergirati. Za obrnute x_0 i x_1?

Geometrijski izvod metode sekante

Geometrijski izvod metode je jednostavan.

- Napišemo jednadžbu sekante u točkama x_{n-1} i x_n i pogledamo gdje taj pravac siječe os x.

Jednadžba sekante je (linearna interpolacija za f u x_{n-1} i x_n)

$$y - f(x_n) = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} (x - x_n).$$

Iz zahtjeva $y = 0$ za $x = x_{n+1}$, izlazi da je nova aproksimacija x_{n+1} dana izrazom

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

Za računanje je dovoljno pretpostaviti da je $f(x_n) \neq f(x_{n-1})$ u svim "susjednim" točkama (iteracijama) x_{n-1} i x_n.

Formula za iteracije u metodi sekanti

Formulu za metodu sekante možemo dobiti i iteriranjem početne formule za regulu falsi.

Izraz za novu aproksimaciju možemo napisati i ovako

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

za $n \geq 1$.

Relacija koja "veže" greške susjednih aproksimacija, izvodi se na isti način kao kod regule falsi, i ima oblik

$$\alpha - x_{n+1} = -(\alpha - x_n)\,(\alpha - x_{n-1})\,\frac{f''(\zeta_n)}{2f'(\xi_n)},$$

gdje je ζ_n između x_{n-1}, x_n i α, a ξ_n između x_{n-1} i x_n.

Red konvergencije metode sekante

Iz ove relacije može se izračunati red konvergencije metode sekante, uz odgovarajuće pretpostavke. Dobivamo da je

$$p = \frac{1 + \sqrt{5}}{2} \approx 1.61803.$$

Dokaz je dosta kompliciran (i ima veze s Fibonaccijevim brojevima, a p je veće rješenje jednadžbe $p^2 = p + 1$).

Napomena. Metoda sekante se još naziva i

- metoda (inverzne) linearne interpolacije,

jer sljedeću aproksimaciju x_{n+1} dobivamo kao

- nultočku linearne interpolacije za f, u prethodne dvije iteracije x_{n-1} i x_n.