Matematički praktikum
Predavanje 7 – [*]

prof. dr. sc. Rudolf Scitovski
prof. dr. sc. Kristian Sabo
doc. dr. sc. Danijel Grahovac
dr. sc. Matea Ugrica[†]

1. prosinca 2020.

Sadržaj

1 Jednodimenzionalna minimizacija strogo kvazikonveksnih funkcija

U Poglavlju **??** uveli smo pojam kvazikonveksne funkcije, a u Lemi **??**, str.**??** naveli smo operativni kriterij za ispitivanje kvazikonveksnosti neke funkcije. Dalje ćemo se baviti

---

[*]Matematički praktikum obavezni je predmet u zimskom semestru druge godine sveučilišnog Diplomskog studija matematike na smjerovima Financijska matematika i statistika i Računarstvo, te na petoj godini sveučilišnog Nastavničkog studija matematike i informatike (30 sati predavanja i 30 sati vježbi, 6 ECTS bodova)

[†]scitowsk@mathos.hr, ksabo@mathos.hr, dgrahova@mathos.hr, mugrica@mathos.hr

strogo kvazikonveksnim funkcijama. Kažemo da je funkcija $f\colon [a,b] \to \mathbb{R}$ strogo kvazikonveksna ako za sve $x, y \in [a,b]$, $(x \neq y)$ vrijedi

$$f(\alpha x + (1-\alpha)y) < \max\{f(x), f(y)\}, \quad \forall \alpha \in (0,1),$$

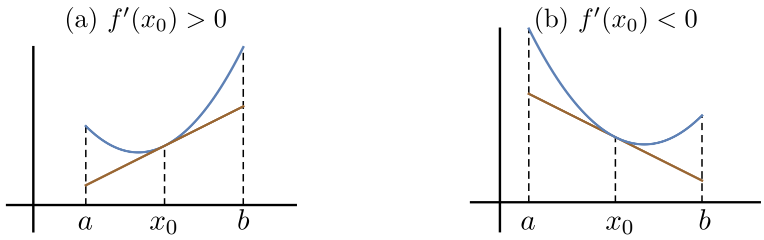
a za svaki $\lambda \in \mathbb{R}$ nivo-skup $\mathcal{D}_\lambda$ je konveksan. Točka $x^\star \in \operatorname*{argmin}_{x\in[a,b]} f(x)$ strogo kvazikonveksne funkcije točka je globalnog minimuma te funkcije.

2 Metoda polovljenja za kvazikonveksnu derivabilnu funkciju

Za kvazikonveksnu derivabilnu funkciju $f\colon [a,b] \to \mathbb{R}$ uz točnost $\epsilon > 0$ treba riješiti optimizacijski problem, tj. odrediti

$$x^\star \in \operatorname*{argmin}_{x\in[a,b]} f(x), \quad \text{takav da je} \quad f(x^\star) = \min_{x\in[a,b]} f(x).$$

Najjednostavnija metoda je tzv. metoda polovljenja (bisekcije), koju opisujemo u nastavku. Najprije odredimo polovište x_0 intervala $[a,b]$. Ako je $f'(x_0) = 0$, točka x_0 rješenje je problema. Ako je $f'(x_0) > 0$, dalje razmatramo podinterval $[a, x_0]$ (vidi Sliku 1a). U protivnom, dalje razmatramo podinterval $[x_0, b]$ (vidi Sliku 1b). Proces se nastavlja tako dugo dok duljina podintervala ne postane manja od ϵ, ili dok derivacija funkcije f u nekoj točki ne postane jednaka nuli.



Slika 1: *Metoda polovljenja* za kvazikonveksnu derivabilnu funkciju

Primjedba 1. Primijetite da se ova metoda može interpretirati kao metoda polovljenja za traženje nultočke derivacije funkcije f (vidi [7]).

Neka je $[a_0, b_0] := [a,b]$ početni interval. Početnu aproksimaciju odredimo kao polovište tog intervala $x_0 = \frac{1}{2}(a_0 + b_0)$, a apsolutna pogreška je $\Delta x_0 = |x_0 - x^\star| \leq \frac{1}{2}(b_0 - a_0)$. Zatim izaberemo novi podinterval

$$[a_1, b_1] := \begin{cases} [a_0, x_0], & f'(x_0) > 0 \\ [x_0, b_0], & f'(x_0) < 0 \end{cases}$$

i odredimo sljedeću aproksimaciju $x_1 = \frac{1}{2}(a_1 + b_1)$. Primijetite da za apsolutnu pogrešku prve aproksimacije x_1 vrijedi $\Delta x_1 := |x_1 - x^\star| \le \frac{1}{2}(b_1 - a_1)$.

Neka je općenito
$$x_k = \tfrac{1}{2}(a_k + b_k).$$
Vrijedi
$$|x^\star - x_k| \le \tfrac{1}{2}(b_k - a_k) = (\tfrac{1}{2})^2(b_{k-1} - a_{k-1}) = \cdots = \left(\tfrac{1}{2}\right)^{k+1}(b_0 - a_0). \tag{1}$$

Na taj način dokazali smo sljedeći teorem o konvergenciji niza (x_k), brzini konvergencije iterativnog postupka te odgovarajuće ocjene pogreške.

Teorem 1. *Neka je $f\colon [a,b] \to \mathbb{R}$ kvazikonveksna derivabilna funkcija. Niz $\left(x_k\right)$ definiran na prethodno opisan način konvergira prema globalnom minimumu funkcije f linearnom brzinom i vrijedi ocjena pogreške*
$$|x^\star - x_k| \le \left(\tfrac{1}{2}\right)^{k+1}|b - a|.$$

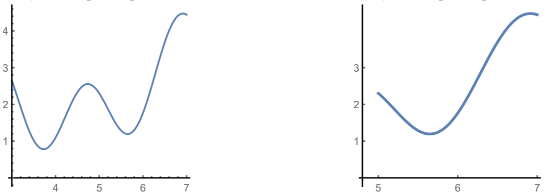
Korištenjem ocjene (1) može se odrediti potreban broj iteracija za traženu točnost $\epsilon > 0$. Iz
$$\left(\tfrac{1}{2}\right)^{(k+1)}|b_0 - a_0| < \epsilon,$$
dobivamo
$$k > \frac{\ln \epsilon - \ln |b_0 - a_0|}{\ln 0.5} - 1.$$

Algorithm 1 **(Metoda polovljenja za kvazikonveksnu derivabilnu funkciju)**

Require: $f\colon [a,b] \to \mathbb{R}$;
Require: $a_0, b_0 \in [a,b]$; $\epsilon > 0$; k=0;
{Učitati funkciju f, početni interval $[a_0, b_0]$, točnost $\epsilon > 0$ i staviti $k = 0$;}
while $b_k - a_k \ge \epsilon$, **do**
  $x_p := \frac{1}{2}(a_k + b_k)$
  if $f'(x_p) > 0$, **then**
    $a_{k+1} = a_k$; $b_{k+1} = x_p$;
  else
    $a_{k+1} = x_p$; $b_{k+1} = b_k$;
  end if
  $k = k + 1$;
end while
Ensure: x_p

Primjer 1. *Metodom polovljenja potražimo lokalni minimum funkcije $f\colon [3,7] \to \mathbb{R}$, $f(x) = \sin x + \sin 3x + \ln x + 1$ na intervalu $[5,7]$ (vidi Sliku 2) s jednom točnom decimalom.*

(a) $f\colon [3,7] \to \mathbb{R}$ (b) $f\colon [5,7] \to \mathbb{R}$

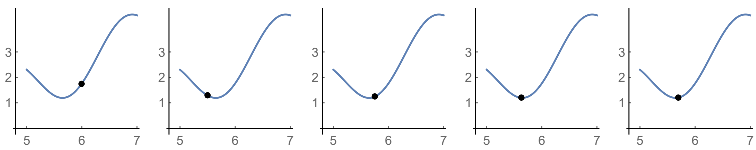


Slika 2: $f(x) = \sin x + \sin 3x + \ln x + 1$

Za $\epsilon = 0.05$ dobivamo $k > \frac{\ln(.05) - \ln|7-3|}{\ln 0.5} - 1 \approx 5.32$, što znači da ćemo nakon $k = 6$ iteracija dobiti aproksimaciju točnu na barem jednu decimalu (vidi Tablicu 1 i Sliku 3). Primijetite da je tražena točnost postignuta već u petoj iteraciji. Dobiva li se isti rezultat primjenom *Mathematica*-modula `NMinimize[]`?

k	a_k	b_k	x_k	$\frac{1}{2}(b_k - a_k)$	$f'(x_k)$
0	5	7	6	1	3.108
1	5	6	5.5	.5	-1.217
2	5.5	6	5.75	.25	.949
3	5.5	5.75	5.625	.125	-.210
4	5.625	5.75	5.6875	.0625	.360
5	5.625	5.6875	5.65625	.03125	0.071
6	5.625	5.65625	5.65625	.015625	0.071

Tablica 1: *Metoda polovljenja*: $f\colon [5,7] \to \mathbb{R}$, $f(x) = \sin x + \sin 3x + \ln x + 1$



Slika 3: *Metoda polovljenja* za kvazikonveksnu derivabilnu funkciju

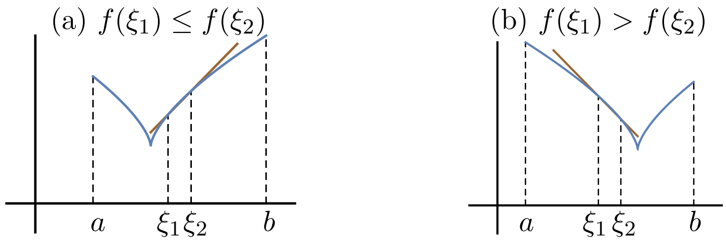
Zadatak 1. *Za funkciju* $f\colon [1,4] \to \mathbb{R}$, $f(x) = \frac{x}{2} + \frac{1}{x}$ *treba odrediti aproksimaciju globalnog minimizatora.*

(a) Primjenom Metode polovljenja *odredite prve tri aproksimacije i ocijenite pogrešku dobivene aproksimacije.*

(b) Koliko iteracija Metodom polovljenja *treba provesti da bi se dobila točnost globalnog minimuma na 4 decimalna mjesta?*

3 Metoda polovljenja za nederivabilnu strogo kvazikonveksnu funkciju

Neka je $f\colon [a,b] \to \mathbb{R}$ strogo kvazikonveksna funkcija. Analogno prethodnoj metodi, konstruirat ćemo *Metodu polovljenja* za ovakvu funkciju. Uz to što funkcija f općenito nije derivabilna, ona na intervalu $[a,b]$ može imati i prekide prve vrste. Neka je $\delta > 0$ realan broj takav da je $b - a > \delta$.



Slika 4: *Metoda polovljenja* za strogo kvazikonveksnu nederivabilnu funkciju

Neka je $[a_0, b_0] := [a, b]$. Budući da u ovom slučaju općenito ne možemo koristiti derivaciju funkcije f u polovištu intervala $[a_0, b_0]$, u okolini polovišta simetrično ćemo izabrati dvije točke

$$\begin{aligned}\xi_1 &= \tfrac{1}{2}(a_0 + b_0) - \tfrac{\delta}{2},\\ \xi_2 &= \tfrac{1}{2}(a_0 + b_0) + \tfrac{\delta}{2},\end{aligned}$$

i promatrati koeficijent smjera odgovarajuće sekante

$$\kappa := \frac{f(\xi_2) - f(\xi_1)}{\delta}.$$

Zato će nam kriterij za odabir podintervala u kojemu ostaje ležati točka minimuma $x^\star$ biti sljedeći (vidi Sliku 4):

$$[a_1, b_1] := \left\{ \begin{array}{ll} [a_0, \xi_2], & \kappa \ge 0 \\ [\xi_1, b_0], & \kappa < 0 \end{array} \right\} = \left\{ \begin{array}{ll} [a_0, \xi_2], & f(\xi_2) \ge f(\xi_1) \\ [\xi_1, b_0], & f(\xi_2) < f(\xi_1) \end{array} \right. .$$

Širina novog podintervala je

$$b_1 - a_1 (= \xi_2 - a_0 = b_0 - \xi_1) = \tfrac{1}{2}(b_0 - a_0) + \tfrac{1}{2}\delta.$$

Prvu aproksimaciju x_1 možemo definirati u polovištu intervala $[a_1, b_1]$

$$x_1 = \tfrac{1}{2}(a_1 + b_1).$$

Ocjena apsolutne pogreške prve aproksimacije je

$$\Delta x_1 := |x_1^\star - x^\star| \le \tfrac{1}{2}(b_1 - a_1) = \tfrac{1}{2^2}(b_0 - a_0) + \tfrac{1}{2^2}\delta,$$

što ćemo zapisati u obliku

$$\Delta x_1 := |x_1^\star - x^\star| \leq \tfrac{1}{2^2}(b_0 - a_0) + \tfrac{1}{2}(1 - \tfrac{1}{2})\delta,$$

Ponavljajući postupak nadalje definiramo dvije simetrične točke u okolini polovišta intervala $[a_1, b_1]$

$$\begin{aligned} \xi_3 &= \tfrac{1}{2}(a_1 + b_1) - \tfrac{1}{2}\delta, \\ \xi_4 &= \tfrac{1}{2}(a_1 + b_1) + \tfrac{1}{2}\delta, \end{aligned}$$

novi koeficijent smjera $\kappa = \frac{1}{\delta}(f(\xi_4) - f(\xi_3))$ i novi podinterval

$$[a_2, b_2] := \begin{cases} [a_1, \xi_4], & \kappa \geq 0 \\ [\xi_3, b_1], & \kappa < 0 \end{cases},$$

u kojemu ostaje točka minimuma $x^\star$ funkcije f. Širina novog podintervala je

$$b_2 - a_2 = \tfrac{1}{2}(b_1 - a_1) + \tfrac{1}{2}\delta\Big[= \tfrac{1}{2^2}(b_0 - a_0) + \tfrac{1}{2^2}\delta + \tfrac{1}{2}\delta\Big] = \tfrac{1}{2^2}(b_0 - a_0) + \tfrac{1}{2}\delta(1 + \tfrac{1}{2}).$$

Novu aproksimaciju x_2 odredit ćemo u polovištu intervala $[a_2, b_2]$

$$x_2 = \tfrac{1}{2}(a_2 + b_2),$$

čiju apsolutnu pogrešku možemo ocijeniti s

$$\Delta x_2 := |x_2 - x^\star| \leq \tfrac{1}{2}(b_2 - a_2) = \tfrac{1}{2^3}(b_0 - a_0) + \tfrac{1}{2}\delta \cdot \tfrac{3}{4},$$

što ćemo zapisati kao

$$\Delta x_2 := |x_2 - x^\star| \leq \tfrac{1}{2^3}(b_0 - a_0) + \tfrac{1}{2}\Big(1 - \tfrac{1}{2^2}\Big)\delta.$$

Općenito, neka je poznat interval $[a_{k-1}, b_{k-1}]$ širine

$$b_{k-1} - a_{k-1} = \tfrac{1}{2}(b_{k-2} - a_{k-2}) + \tfrac{1}{2}\delta, \quad k \geq 2,$$

u kojemu leži točka minimuma $x^\star$. Kao i ranije, definiramo dvije simetrične točke u okolini polovišta

$$\begin{aligned} \xi_{2k-1} &= \tfrac{1}{2}(a_{k-1} + b_{k-1}) - \tfrac{1}{2}\delta, \\ \xi_{2k} &= \tfrac{1}{2}(a_{k-1} + b_{k-1}) + \tfrac{1}{2}\delta, \end{aligned}$$

koeficijent smjera $\kappa = \frac{1}{\delta}(f(\xi_{2k}) - f(\xi_{2k-1})$ i sljedeći podinterval

$$[a_k, b_k] := \begin{cases} [a_{k-1}, \xi_{2k}], & \kappa \geq 0 \\ [\xi_{2k-1}, b_{k-1}], & \kappa < 0 \end{cases}$$

širine

$$b_k - a_k = \tfrac{1}{2}(b_{k-1} - a_{k-1}) + \tfrac{\delta}{2} = \tfrac{1}{2^k}(b_0 - a_0) + \tfrac{1}{2}\left(1 + \cdots + \tfrac{1}{2^{k-1}}\right)\delta.$$

Sljedeću aproksimaciju x_k definirat ćemo s

$$x_k := \tfrac{1}{2}(a_k + b_k).$$

Apsolutnu pogrešku možemo ocijeniti s

$$\begin{aligned}\Delta x_k := |x_k - x^\star| \le \tfrac{1}{2}(b_k - a_k) &= \tfrac{1}{2^{k+1}}(b_0 - a_0) + \tfrac{1}{2^2}\left(1 + \cdots + \tfrac{1}{2^{k-1}}\right)\delta \\ &= \tfrac{1}{2^{k+1}}(b_0 - a_0) + \tfrac{1}{2}\left(1 - \tfrac{1}{2^k}\right)\delta.\end{aligned}$$

Za $k \to +\infty$ ocjena apsolutne pogreške konvergira prema $\frac{\delta}{2}$. Dakle, nakon dovoljnog broja iteracija dobivena aproksimacija može odstupati od točke minimuma za najviše $\frac{\delta}{2}$.

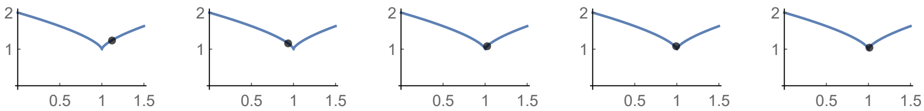
Metodu nazivamo *Metoda polovljenja* jer u svakom koraku interval dijelimo gotovo na dvije polovine – čim je δ manji, time je dijeljenje intervala bliže raspolavljanju. Naravno, ova metoda može se primijeniti i na proizvoljnu strogo konveksnu funkciju (vidi Primjedbu **??**, str.**??**).

Algorithm 2 **(Metoda polovljenja za kvazikonveksnu funkciju)**

Require: $f\colon [a,b] \to \mathbb{R}$;
Require: a, b; $0 < \delta < (b-a)/2$; $\epsilon > \delta$; $k = 0$;
  {Učitati funkciju f, početni interval $[a,b]$, točnost $\epsilon > 0$ i staviti $k = 0$;}
  $x_p = (a+b)/2$;
  while $b - a \ge \epsilon$, **do**
    $k = k + 1$;
    $x_1 = x_p - \frac{\delta}{2}$; $x_2 = x_p + \frac{\delta}{2}$;
    if $f(x_1) \le f(x_2)$, **then**
      $b = x_2$;
    else
      $a = x_1$;
    end if
    $x_p = (a+b)/2$;
  end while
Ensure: x_p

Primjedba 2. Primijetite da se ova metoda može interpretirati kao metoda polovljenja gdje je derivacija zamijenjena s podijeljenom razlikom (vidi [7]).

Primjer 2. *Metodu polovljenja ilustrirat ćemo na primjeru minimizacije funkcije* $f : [-1, 2] \to \mathbb{R}$, $f(x) = 1 + \sqrt[3]{(x-1)^2}$. *Na Slici 5 prikazano je nekoliko iteracija.*



Slika 5: Ilustracije *Metode polovljenja* na funkciji iz Primjera 2

Zadatak 2. *Je li* $f\colon [-2,2] \to \mathbb{R}$, $f(x) = 1 - e^{-|x-1|}$ *kvazikonveksna funkcija? Ako je,* Metodom polovljenja *odredite aproksimaciju njezine točke globalnog minimuma.*

Zadatak 3. *Lokalni minimumi funkcije iz Primjera* **??**, *str.***??** *postižu se u točkama*

$$\begin{array}{lll} x_1^\star \approx -3.217, & x_2^\star \approx 0.243, & x_3^\star \approx 3.617, \\ f(x_1^\star) \approx 6.251, & f(x_2^\star) \approx 3.289, & f(x_3^\star) \approx 0.618. \end{array}$$

Definirajte intervale u kojima ćete Metodom polovljenja *odrediti pojedine točke lokalnih minimuma.*

Zadatak 4. Metodom polovljenja *odredite prve četiri aproksimacije minimuma funkcije* $f\colon [0,3] \to \mathbb{R}$, $f(x) = |x-1| + 2$ *uz* $\delta = 0.2$. *Ocijenite pogrešku dobivene aproksimacije. Kolika je stvarna apsolutna pogreška?*

Zadatak 5. *Za dani* $\epsilon > 0$ *i parametar* $\delta > 0$ *odredite broj iteracija* k *koji je potreban da* Metoda polovljenja *daje točku globalnog minimuma* $x^\star$ *kvazikonveksne funkcije* $f\colon [a,b] \to \mathbb{R}$ *s točnošću* $\frac{\delta}{2} + \epsilon$. *Kako treba odabrati* $\delta > 0$ *i* k *da bismo postigli točnost na dvije decimale? U tom smislu analizirajte Primjer 2.*

Zadatak 6. *Neka je* $f\colon [a,b] \to \mathbb{R}$ *strogo kvazikonveksna funkcija. Definirajte modifikaciju* Metode polovljenja *za minimizaciju funkcije* f *u kojoj je vrijednost parametra* δ *moguće mijenjati u svakoj iteraciji* $\delta = \delta_k$. *Izvedite ocjenu apsolutne pogreške i dokažite konvergenciju metode uz dodatne pretpostavke na niz* (δ_k). *Što možete reći o brzini konvergencije metode?*

4 Metoda zlatnog reza

Kao što smo već primijetili, *Metoda polovljenja za kvazikonveksnu derivabilnu funkciju* ne može točno odrediti točku minimuma, već točnost aproksimacije opisi o veličini parametra $\delta > 0$. Zato ćemo točke x_{2k-1}, x_{2k}, pomoćukojih smo u svakom koraku određivali sekantu, pokušati definirati neovisno o ovakvom parametru δ. Takvu mogućnost pružanam *Metoda zlatnog reza.*

Neka je ℓ dužina duljine 1. Točka (odnosno realni broj) $.5 < x < 1$ dijeli dužinu ℓ u omjeru zlatnog reza ako se omjer duljine cijele dužine ℓ u odnosu na veći dio x odnosi kao omjer većeg dijela x u odnosu na manji dio $1-x$, tj.

$$\frac{1}{x} = \frac{x}{1-x}.$$

Pozitivno rješenje ove jednadžbe je $x = \frac{\sqrt{5}-1}{2} \approx 0.618$, pri čemu broj $\frac{1}{x} = \frac{\sqrt{5}+1}{2} \approx 1.61812$ nazivamo omjer zlatnog reza.

Slika 6: Dijeljenje dužine ℓ u omjeru zlatnog reza

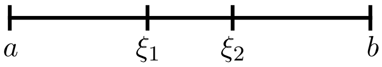
Zadatak 7. *Pokažite da i broj* $1-x=\frac{3-\sqrt{5}}{2}\approx 0.382$ *također dijeli dužinu* ℓ *u omjeru zlatnog reza.*

Neka je $f\colon [a,b]\to\mathbb{R}$ strogo kvazikonveksna funkcija. Niz aproksimacija točke globalnog minimuma $x^\star\in[a,b]$ funkcije f definirat ćemo na sljedeći način. Formalno označimo $[a_0,b_0]:=[a,b]$.

U cilju definiranja prve aproksimacije najprije definirajmo točke $\xi_1,\xi_2\in[a_0,b_0]$, $\xi_1<\xi_2$, na sljedeći način

$$\xi_1=a_0+(1-x)(b_0-a_0)=a_0+\tfrac{3-\sqrt{5}}{2}(b_0-a_0), \tag{2}$$

$$\xi_2=a_0+x(b_0-a_0)=a_0+\tfrac{\sqrt{5}-1}{2}(b_0-a_0). \tag{3}$$



Slika 7: Izbor točaka ξ_1,ξ_2

Točka ξ_1 čini zlatni rez intervala $[a_0,b_0]$. Naime,

$$\frac{b_0-a_0}{b_0-\xi_1}=\frac{b_0-\xi_1}{\xi_1-a_0}.$$

Slično se može pokazati da i točka ξ_2 čini zlatni rez intervala $[a_0,b_0]$ jer vrijedi

$$\frac{b_0-a_0}{\xi_2-a_0}=\frac{\xi_2-a_0}{b_0-\xi_2}.$$

Točke ξ_1 i ξ_2 poslužit će nam slično kao u t.3, str.5 za definiranje intervala $[a_1,b_1]$ i prve aproksimacije x_1. Uz oznaku $\kappa:=\frac{f(\xi_2)-f(\xi_1)}{\xi_2-\xi_1}$ definiramo

$$[a_1,b_1]=\begin{cases}[a_0,\xi_2], & \kappa\geq 0,\\ [\xi_1,b_0], & \kappa<0,\end{cases}\qquad x_1=\begin{cases}\xi_1, & \kappa\geq 0,\\ \xi_2, & \kappa<0.\end{cases}$$

Širina intervala $[a_1,b_1]$ je

$$b_1-a_1=\begin{cases}\xi_2-a_0\overset{(3)}{=}\frac{\sqrt{5}-1}{2}(b_0-a_0),\\ b_0-\xi_1\overset{(2)}{=}b_0-a_0-\frac{\sqrt{5}-1}{2}(b_0-a_0)=\frac{\sqrt{5}-1}{2}(b_0-a_0)\end{cases},$$

a pogrešku prve aproksimacije procijenit ćemo slično kao u t.3 koristeći $\frac{\sqrt{5}-1}{2}>\frac{1}{2}$:

$$\Delta x_1:=|x_1-x^\star|\leq\tfrac{1}{2}(b_1-a_1)\leq\Big(\tfrac{\sqrt{5}-1}{2}\Big)(b_1-a_1)=\Big(\tfrac{\sqrt{5}-1}{2}\Big)^2(b_0-a_0).$$

Također primijetimo da je

$$\xi_2 - \xi_1 \overset{(3)-(2)}{=} (\sqrt{5}-2)(b_0 - a_0).$$

Ponavljajući analogno kao u (3), (2), definiramo točke x_3, x_4 u intervalu $[a_1, b_1]$

$$\xi_3 = a_1 + \tfrac{3-\sqrt{5}}{2}(b_1 - a_1),$$
$$\xi_4 = a_1 + \tfrac{\sqrt{5}-1}{2}(b_1 - a_1).$$

Primjedba 3. Označimo li $\phi = \frac{\sqrt{5}-1}{2}$, možemo uočiti da ako je $f(\xi_1) \leq f(\xi_2)$, onda zbog $\phi^2 = 1 - \phi = \frac{3-\sqrt{5}}{2}$ vrijedi

$$\begin{aligned}\xi_4 &= a_0 + \phi(\xi_2 - a_0) = a_0 + \phi(a_0 + \phi(b_0 - a_0) - a_0)\\ &= a_0 + \phi^2(b_0 - a_0) = \xi_2.\end{aligned}$$

S druge strane, ako je $f(\xi_1) > f(\xi_2)$, onda

$$\begin{aligned}\xi_3 &= \xi_1 + (1-\phi)(b_0 - \xi_1)\\ &= a_0 + (1-\phi)(b_0 - a_0) + (1-\phi)(b_0 - a_0 - (1-\phi)(b_0 - a_0))\\ &= a_0 + (1-\phi^2)(b_0 - a_0) = \xi_1.\end{aligned}$$

Stoga, u svakoj je iteraciji potrebno izračunati samo jednu novu točku.

Pokažite da i točka ξ_3 i točka ξ_4 dijele segment $[a_1, b_1]$ u omjeru zlatnog reza.

Pomoću točaka ξ_3, ξ_4 koristeći oznaku $\kappa := \frac{f(\xi_4)-f(\xi_3)}{\xi_4-\xi_3}$ definirat ćemo novi podinterval $[a_2, b_2]$ i drugu aproksimaciju x_2

$$[a_2, b_2] = \begin{cases}[a_1, \xi_4], & \kappa \geq 0,\\ [\xi_3, b_1], & \kappa < 0,\end{cases} \qquad x_2 = \begin{cases}\xi_3, & \kappa \geq 0,\\ \xi_4, & \kappa < 0.\end{cases}$$

Širina intervala $[a_2, b_2]$ je

$$b_2 - a_2 (= \xi_4 - a_1 = b_1 - \xi_3) = \tfrac{\sqrt{5}-1}{2}(b_1 - a_1) = \left(\tfrac{\sqrt{5}-1}{2}\right)^2 (b_0 - a_0),$$

a pogreška druge aproksimacije x_2 procijenjena je s

$$\Delta x_2 := |x_2 - x^\star| \leq \tfrac{1}{2}(b_2 - a_2) \leq \left(\tfrac{\sqrt{5}-1}{2}\right)(b_2 - a_2) = \left(\tfrac{\sqrt{5}-1}{2}\right)^3 (b_0 - a_0).$$

Također,

$$\xi_4 - \xi_3 = (\sqrt{5}-2)(b_1 - a_1) = \tfrac{\sqrt{5}-1}{2}(\sqrt{5}-2)(b_0 - a_0).$$

Općenito, pretpostavimo da je poznat interval $[a_{k-1}, b_{k-1}]$ u kojemu leži točka minimuma x. Kao i ranije, definiramo dvije točke koje dijele interval $[a_{k-1}, b_{k-1}]$ u omjeru zlatnog reza

$$\xi_{2k-1} = a_{k-1} + \tfrac{3-\sqrt{5}}{2}(b_{k-1} - a_{k-1}),$$
$$\xi_{2k} = a_{k-1} + \tfrac{\sqrt{5}-1}{2}(b_{k-1} - a_{k-1}).$$

Pomoću točaka ξ_{2k-1}, ξ_{2k} koristeći oznaku $\kappa := \frac{f(\xi_{2k})-f(\xi_{2k-1})}{\xi_{2k}-\xi_{2k-1}}$ definirat ćemo novi podinterval $[a_k, b_k]$ i novu aproksimaciju x_k

$$[a_k, b_k] = \begin{cases} [a_{k-1}, \xi_{2k}], & \kappa \geq 0, \\ [\xi_{2k}, b_k], & \kappa < 0, \end{cases} \qquad x_k = \begin{cases} \xi_{2k-1}, & \kappa \geq 0, \\ \xi_{2k}, & \kappa < 0. \end{cases}$$

Širina podintervala $[a_k, b_k]$ je

$$b_k - a_k (= \xi_{2k} - a_k = b_k - \xi_{2k-1}) = \tfrac{\sqrt{5}-1}{2}(b_{k-1} - a_{k-1}) = \left(\tfrac{\sqrt{5}-1}{2}\right)^k (b_0 - a_0),$$

a apsolutna pogreška k-te aproksimacije zadana je s

$$\Delta x_k := |x_k - x^\star| \leq \left(\tfrac{\sqrt{5}-1}{2}\right)(b_k - a_k) = \left(\tfrac{\sqrt{5}-1}{2}\right)^{k+1} (b_0 - a_0),$$

te je

$$\xi_{2k} - \xi_{2k-1} = (\sqrt{5} - 2)(b_{k-1} - a_{k-1}) = \left(\tfrac{\sqrt{5}-1}{2}\right)^{k-1} (\sqrt{5} - 2)(b_0 - a_0).$$

Primijetite da za $k \to +\infty$ niz apsolutnih pogrešaka $\left(\Delta x_k\right)$ konvergira prema nuli te da isto tako niz $(\xi_{2k} - \xi_{2k-1})$ konvergira prema nuli. To znači da primjenom *Metode zlatnog reza* za traženje globalnog minimizatora strogo kvazikonveksne funkcije f (za razliku od *Metode polovljenja*) možemo postići proizvoljnu točnost aproksimacije. Može se također pokazati[1] da je pri tome izbor točaka $\{\xi_{2k-1}, \xi_{2k}\}$ gotovo optimalan.

Više o metodi vidi u [3, 5, 8].

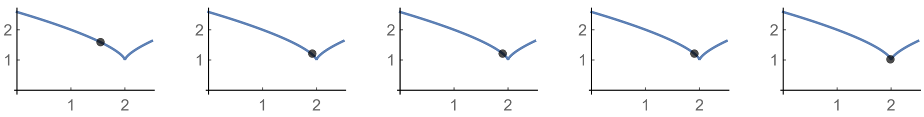
Algorithm 3 **(Metoda zlatnog reza)**

Require: $f\colon [a, b] \to \mathbb{R}$;
Require: a, b; $\epsilon > 0$; $\tau = \frac{\sqrt{5}-1}{2}$; $k = 0$;
  {Učitati funkciju f, početni interval $[a, b]$, točnost $\epsilon > 0$ i staviti $k = 0$;}
while $b - a \geq \epsilon$, **do**
  $k = k + 1$;
  $x_1 = a + \frac{3-\sqrt{5}}{2}(b - a)$; $x_2 = a + \frac{\sqrt{5}-1}{2}(b - a)$;
  if $f(x_1) \leq f(x_2)$, **then**
    $b = x_2$; $x^\star = x_1$
  else
    $a = x_1$; $x^\star = x_2$
  end if
end while
Ensure: $x^\star$

Primjer 3. *Metodu zlatnog reza ilustrirat ćemo na primjeru minimizacije funkcije* $f : [0, 2.5] \to \mathbb{R}$, $f(x) = 1 + \sqrt[3]{(x-2)^2}$. *Na Slici 8 prikazano je nekoliko iteracija.*

---

[1] Iz povijesnih razloga navodimo originalnu referencu:
Ф. П. Василев, Лекции по методам решения екстремалных задач, Издавателство Московского униврзитета, Москва, 1974.



Slika 8: Traženje globalnog minimuma funkcije iz Primjera 3 *Metodom zlatnog reza*

Zadatak 8. *Neka je* $f\colon [a,b] \to \mathbb{R}$ *strogo kvazikonveksna funkcija. Odredite broj iteracija* k *koji je potrebnan da bi se* Metodom zlatnog reza *odredila aproksimacija točke globalnog minimuma s točnošću* $\epsilon > 0$. *Koliko treba iteracija da bi se* Metodom zlatnog reza *pronašla točka globalnog minimuma funkcije iz Primjera 3 s točnošću na dvije decimale?*

Zadatak 9. Metodom zlatnog reza *odredite prve četiri aproksimacije minimuma funkcije* $f\colon [0,3] \to \mathbb{R}$, $f(x) = |x-1| + 2$. *Ocijenite pogrešku dobivene aproksimacije. Kolika je stvarna apsolutna pogreška? Koliko je iteracija potrebno za točnost* $\varepsilon = 0.0005$*?*

Zadatak 10. *Neka je* $f\colon [a,b] \to \mathbb{R}$ *konveksna derivabilna funkcija. Usporedite broj iteracija* k *koji je potreban za određivanje točke globalnog minimuma funkcije* f *s točnošću* $\epsilon > 0$ *pomoću* Metode polovljenja *i pomoću* Metode zlatnog reza*. Usporedbu provedite općenito i također na primjeru funkcije iz Primjera 3.*

Zadatak 11. *Neka je* $f\colon [a,b] \to \mathbb{R}$ *strogo kvazikonveksna funkcija. Izbor intervala koji sadrži točku minimimuma funkcije* f *u* Metodi zlatnog reza *temelji se na izboru točaka*

$$\begin{aligned}\xi_{2k-1} &= a_{k-1} + (1-\tau)(b_{k-1} - a_{k-1}),\\ \xi_{2k} &= a_{k-1} + \tau(b_{k-1} - a_{k-1}),\end{aligned}$$

gdje je $\tau = \frac{\sqrt{5}-1}{2}$. *Za proizvoljni* $\tau \in (1/2, 1)$, *koliko iznosi stopa smanjenja duljine intervala*

$$\frac{b_k - a_k}{b_{k-1} - a_{k-1}}?$$

Izradite Mathematica*-modul za* Metodu zlatnog reza *s proizvoljnim parametrom* $\tau \in \langle \frac{1}{2}, 1\rangle$. *Na primjeru funkcije* $f\colon [0, 2.5] \to \mathbb{R}$, $f(x) = 1 + \sqrt[3]{(x-2)^2}$ *testirajte brzinu konvergencije metode za različite izbore parametra* τ.

5 Jednodimenzionalna globalna optimizacija

5.1 Lipschitz-neprekidne funkcije

U ovom odjeljku razmatrat ćemo jednu metodu traženja globalnog minimuma realne neprekidne funkcije jedne varijable definirane na nekom intervalu $[a,b]$. Od funkcije se neće zahtijevati derivabilnost, već samo ograničeni rast na tom intervalu. Metodu su gotovo istovremeno, ali nezavisno, razradili S. A. Pijavskij[2] i B. O. Shubert[3].

Definicija 1. Kažemo da je funkcija $f\colon \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}$ *Lipschitz-neprekidna* (zadovoljava Lipschitzov uvjet) s konstantom $L > 0$ na $\mathcal{D}$ i pišemo $f \in Lip_L\mathcal{D}$ ako za sve $x, y \in \mathcal{D}$ vrijedi

$$|f(x) - f(y)| \leq L|x - y|. \tag{4}$$

Konstantu $L > 0$ nazivamo Lipschitzova konstanta.

Primjer 4. *Funkcija $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$ je Lipschitz-neprekidna s konstantom $L = 1$ na $\mathbb{R}$ jer je $||x| - |y|| \leq 1 \cdot |x - y|$ za sve $x, y \in \mathbb{R}$.*

Primjer 5. *Odredimo Lipschitzovu konstantu funkcije $f\colon [-3, 2] \to \mathbb{R}$, $f(x) = x^3 - 6x + 2$.*

Primijetite da je ova funkcija derivabilna u svakoj točki segmenta $[-3, 2]$ i da za nju vrijedi Lagrangeov teorem o srednjoj vrijednosti, tj. za proizvoljni $x, y \in [-3, 2]$, $x < y$, postoji $c \in \langle x, y\rangle$ takav da je

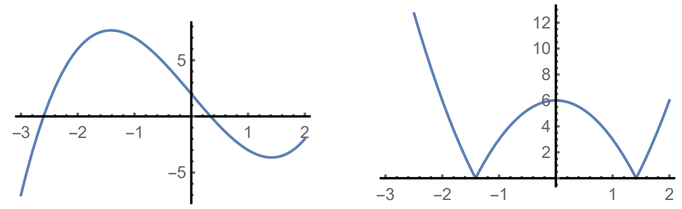
$$f(x) - f(y) = f'(c)(x - y).$$

Zbog neprekidnosti funkcije f' na segmentu $[-3, 2]$ sukladno Weierstrasseovom teoremu (vidi [10]) postoji $L = \max_{x\in[a,b]} |f'(x)|$ iz čega slijedi

$$|f(x) - f(y)| \leq L|x - y|, \quad \forall x, y \in [-3, 2].$$

U ovom slučaju iz Slike 9 vidi se da je najmanja moguća vrijednost za Lipschitzovu konstantu $L = f'(-3) = 21$, iako i svaki drugi broj veći od 21 zadovoljava (4).

(a) $x \mapsto f(x)$ (b) $x \mapsto |f'(x)|$



Slika 9: $f\colon [-3, 2] \to \mathbb{R}$, $f(x) = x^3 - 6x + 2$ je Lipschitz-neprekidna funkcija

U nastavku pretpostavljamo da je područje definicije promatrane funkcije f interval realnih brojeva $[a, b]$. Za $x \neq y$ Lipschitzov uvjet (4) možemo zapisati u obliku

$$-L \leq \frac{f(x) - f(y)}{x - y} \leq L \quad \forall x, y \in [a, b],$$

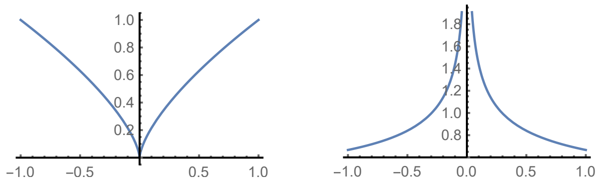
[2] Iz povijesnih razloga navodimo originalne reference:
С. А. Пиявский, Один алгоритм отыскания абсолютного екстрмума функции, Ж. вычисл. матем. и матем. физ. **12**(1972), 888-896. `http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=zvmmf&paperid=6654&option_lang=eng`
Vidi također:
Ф. П. Василев, Лекции по методам решения екстремалных задач, Издавателство Московского универзитета, Москва, 1974.

[3] Bruno O. Shubert, *A sequential method seeking the global maximum of a function*, SIAM Journal on Numerical Analysis, **9**(1972), 379–388

odakle se vidi da je relativna promjena (kvocijent diferencija) funkcije f ograničena između $-L$ i L. Ako x fiksiramo, a y pustimo prema x, onda možemo zaključiti da ako postoji derivacija funkcije f, ona je ograničena između $-L$ i L.

Primjer 6. *Promatrajmo funkciju $f\colon [-1,1] \to \mathbb{R}$, $f(x) = \sqrt[3]{x^2}$. Kao što se može vidjeti na Slici 10, derivacija (brzina pada) ove funkcije u okolini 0 raste u beskonačnost. Zato ne postoji $L > 0$ takav da vrijedi (4), pa ova funkcija nije Lipschitz-neprekidna na $[-1,1]$.*

(a) $x \mapsto f(x)$ (b) $x \mapsto |f'(x)|$



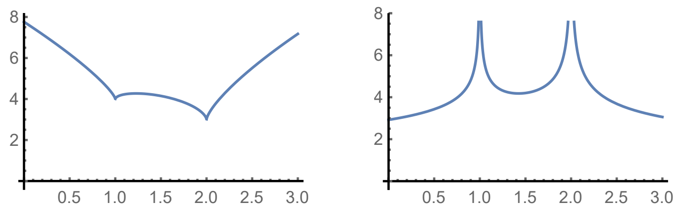
Slika 10: $f\colon [-1,1] \to \mathbb{R}$, $f(x) = \sqrt[3]{x^2}$ nije Lipschitz-neprekidna funkcija

Primjer 7. *Odredimo Lipschitzovu konstantu funkcije $f\colon [0,3] \to \mathbb{R}$, $f(x) = 1+2\sqrt[3]{(x-1)^2}+3\sqrt[3]{(x-2)^2}$ (vidi Sliku 11). Njena derivacija*

$$f'(x) = \frac{4}{3\sqrt[3]{x-1}} + \frac{2}{\sqrt[3]{x-2}}$$

nije definirana u točkama iz $\{1,2\}$, a Lipschitzova konstanta $L > 0$ ne postoji.

(a) $x \mapsto f(x)$ (b) $x \mapsto |f'(x)|$

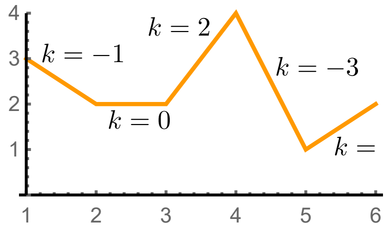


Slika 11: Funkcija $f\colon [0,3] \to \mathbb{R}$, $f(x) = 1+2\sqrt[3]{(x-1)^2}+3\sqrt[3]{(x-2)^2}$ nije Lipschitz-neprekidna funkcija na $[0,3]$

Primjer 8. *Odredimo Lipschitzovu konstantu funkcije $f\colon [1,6] \to \mathbb{R}$,*

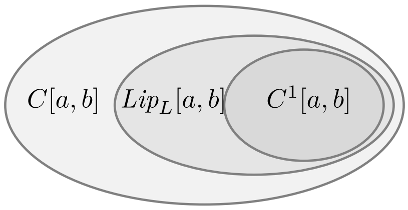
$$f(x) = \begin{cases} -x+4, & 1 \le x \le 2, \\ 2, & 2 \le x \le 3, \\ 2x-4, & 3 \le x \le 4, \\ -3x+16, & 4 \le x \le 5, \\ x-4, & 5 \le x \le 6, \end{cases}$$

čiji je graf prikazan na Slici 12. Očigledno je $L = 3$.



Slika 12: Lipschitz-neprekidna funkcija s konstantom $L = 3$

Iako je određivanje Lipschitzove konstante općenito vrlo složen problem (vidi primjerice [11**?**]), u slučaju da promatrana funkcija f ima konačnu derivaciju u svakoj točki intervala $[a, b]$, problem se može riješiti relativno jednostavno (kao u Primjerima 4, 5 ili 8).



Slika 13: Skup neprekidnih, Lipschitz-neprekidnih i neprekidno derivabilnih funkcija na $[a, b]$

Primjedba 4. Svaka Lipschitz-neprekidna funkcija neprekidna je funkcija, ali neprekidna funkcija ne mora biti Lipschitz-neprekidna (vidi Primjere 6 i Primjere 7). Nadalje, ako je $f \in C^1[a, b]$, onda je $f \in Lip_L[a, b]$ (Lagrangeov teorem o srednjoj vrijednosti), ali obrat ne vrijedi (vidi Primjer 4 i Primjer 8). Dakle, mogli bismo reći da je Lipschitz-neprekidna funkcija neprekidna funkcija, a ako ima derivaciju, ona mora biti ograničena.

Ako je $f : [a, b] \to \mathbb{R}$ neprekidna funkcija, koja ima po dijelovima neprekidnu ograničenu derivaciju, onda Lipschitzovu konstantu $L > 0$ možemo odrediti primjenom Lagrangeovog teorema o srednjoj vrijednosti (vidi Primjer 8). Preciznije, neka je $f\colon [a, b] \to \mathbb{R}$ neprekidna po dijelovima glatka funkcija, u smislu da je f derivabilna, osim eventualno u konačno mnogo točaka $x_1 < x_2 < \cdots < x_n$, te da je f' neprekidna na intervalima (x_j, x_{j+1}), $j = 1, \ldots, n-1$. Ako je $|f'(x)| \le L$ za sve $x \in [a, b] \setminus \{x_1, \ldots, x_n\}$, tada je $f \in Lip_L[a, b]$.

Primjedba 5. Definicija Lipschitz-neprekidne funkcije lako se može proširiti na funkciju više varijabli $f\colon \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$. Takva je funkcija *Lipschitz-neprekidna* (zadovoljava *Lipschitzov uvjet*) ako postoji konstanta $L > 0$ tako da za sve $x, y \in \mathcal{D}$ vrijedi

$$|f(x) - f(y)| \le L\|x - y\|,$$

gdje je $\|\cdot\|$ neka norma. Uočite da je definicija neovisna o izboru norme $\|\cdot\|$ s obzirom na to da su sve norme na $\mathbb{R}^n$ ekvivalentne.

Zadatak 12. *Provjerite jesu li sljedeće funkcije Lipschitz-neprekidne te ako jesu, odredite Lipschitzovu konstantu.*

(a) $f_1 \colon [0,2] \to \mathbb{R}$, $f_1(x) = e^x$,

(b) $f_2 \colon [0,2] \to \mathbb{R}$,

$$f_2(x) = \begin{cases} 5 - x, & 0 \le x \le 3, \\ 4 - \frac{1}{2}(x-5)^2, & 3 < x \le 7, \\ \frac{1}{2}(x-9)^2, & 7 < x \le 9, \end{cases}$$

(c) $f_3 \colon [-1,1] \to \mathbb{R}$, $f_3(x) = 1 + \sqrt[3]{x^2}$.

Zadatak 13. *Neka je* $f_1 \in Lip_{L_1}[a,b]$ *i* $f_2 \in Lip_{L_2}[a,b]$. *Pokažite da su tada i sljedeće funkcije Lipschitz-neprekidne i odredite Lipschitzovu konstantu:*

(a) $f_1 + f_2$,

(b) cf_1, $c \in \mathbb{R}$,

(c) $c_1 f_1 + c_2 f_2$, $c_1, c_2 \in \mathbb{R}$,

(d) $f_1 \cdot f_2$,

(e) $f_1 \circ f_2$, *pod pretpostavkom da je dobro definirana.*

Zadatak 14. *Neka je* $f : [a,b] \to \mathbb{R}$ *po dijelovima Lipschitz-neprekidna funkcija u smislu da postoje* $a = x_0 < x_1 < \cdots < x_n = b$ *takvi da je* f *Lipschitz-neprekidna na intervalima* $\langle x_i, x_{i+1} \rangle$, $i = 0, \ldots, n-1$. *Pokažite da je tada* f *Lipschitz-neprekidna te odredite Lipschitzovu konstantu.*

Zadatak 15. *Dokažite da konveksna funkcija* $f : [a,b] \to \mathbb{R}$ *zadovoljava Lipschitzov uvjet na svakom segmentu* $[\alpha, \beta] \subset [a,b]$. *Korištenjem funkcije* $f(x) = -\sqrt{1-x^2}$, $|x| \le 1$, *pokažite da u općem slučaju nije dobro birati* $\alpha = a$, $\beta = b$.
Uputa. *Korištenjem nejednakosti (**??**) i (**??**) pokažite da je*

$$f'(\alpha+) \le \frac{f(u)-f(v)}{u-v} \le f'(\beta-), \qquad \forall u, v \in [\alpha, \beta].$$

Zadatak 16. *Neka je* f_λ, $\lambda \in \Lambda$ *proizvoljna familija funkcija iz* $Lip_L[a,b]$. *Provjerite jesu li tada i funkcije* $\inf\limits_{\lambda \in \Lambda} f_\lambda$ *i* $\sup\limits_{\lambda \in \Lambda} f_\lambda$ *Lipschitz-neprekidne na* $[a,b]$ *te odredite Lipschitzove konstante.*

Zadatak 17. *Provjerite je li funkcija* $f \colon [0,1] \times [0,2] \to \mathbb{R}$, $f(x_1, x_2) = x_1 x_2^2$ *Lipschitz-neprekidna te ako jest, odredite Lipschitzovu konstantu.*

Zadatak 18. *Neka je* $\mathcal{D} = [a_1, b_1] \times \cdots \times [a_n, b_n]$ *i* $f \colon \mathcal{D} \to \mathbb{R}$ *funkcija. Dokažite:*

a) f *je Lipschitz-neprekidna ako i samo ako postoji konstanta* $L > 0$ *takva da je* f *Lipschitz-neprekidna s konstantom* L *u svakoj varijabli. Preciznije, za svaki* $i = 1, \ldots, n$ *i za proizvoljne* $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ *funkcija*

$$x \mapsto f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)$$

je Lipschitz-neprekidna na $[a_i, b_i]$ *s konstantom* L.

b) *Ako su sve parcijalne derivacije funkcije* f *definirane i ograničene na* $\mathcal{D}$, *tada je* f *Lipschitz-neprekidna.*

6 Pijavskijeva metoda slomljenih pravaca

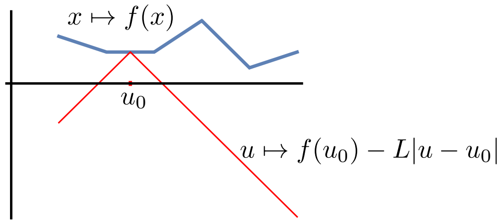
Pretpostavimo da je $f : [a, b] \to \mathbb{R}$ Lipschitz-neprekidna funkcija. Tada prema (4) za sve $x, y \in [a, b]$ vrijedi

$$-L|x - y| \leq f(x) - f(y) \leq L|x - y|,$$

iz čega slijedi

$$f(x) \geq f(y) - L|x - y|, \quad \forall x, y, \in [a, b], \tag{5}$$

(vidi Sliku 14). Iz toga zaključujemo da funkciju f u bilo kojoj točki $y \in [a, b]$ možemo odozdo ograničiti funkcijom $x \mapsto f(y) - L|x - y|$.



Slika 14: Lipschitz-donja ograda funkcije

Izaberimo proizvoljni $u_0 \in [a, b]$ i definirajmo donju ogradu funkcije f kao Lipschitz-neprekidnu po dijelovima linearnu funkciju $K : [a, b] \to \mathbb{R}$ (vidi Sliku 14):

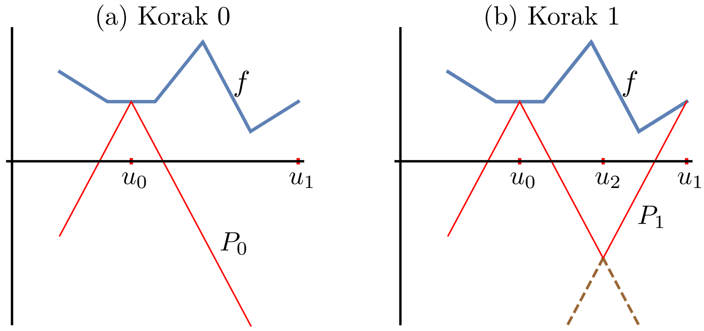
$$K(u; u_0) = f(u_0) - L|u - u_0|, \quad u \in [a, b]. \tag{6}$$

Za slomljeni pravac K lako se vidi da vrijedi:

$$K(u_0; u_0) = f(u_0), \tag{7}$$
$$K(u; u_0) \leq f(u), \quad \forall u \in [a, b]. \tag{8}$$

Svojstvo (8) direktno slijedi iz (5).

Uz pretpostavku da postoji $u^\star \in [a, b]$ na kojemu se postiže globalni minimum Lipschitz-neprekidne funkcije $f \colon [a, b] \to \mathbb{R}$ s konstantom $L > 0$, tj. postoji $u^\star \in \operatorname*{argmin}_{u \in [a,b]} f(u)$, *Pijavskij* je konstruirao sljedeći Algoritam slomljenih pravaca.



Slika 15: *Algoritam Pijavskog*

Najprije, uz poznavanje u_0 i $K(u; u_0)$ odredimo $u_1 \in \underset{u\in[a,b]}{\operatorname{argmin}} K(u; u_0)$ i stavimo $P_0(u) :=$ $K(u; u_0)$ (vidi Sliku 15a). Očigledno će biti $u_1 = a$ ili $u_1 = b$.

U prvom koraku definiramo

$$
\begin{aligned}
K(u; u_1) &:= f(u_1) - L|u - u_1|,\\
P_1(u) &:= \max_{i=0,1} K(u; u_i) = \max\{K(u; u_1), P_0(u)\},
\end{aligned}
$$

i odredimo $u_2 \in \underset{u\in[a,b]}{\operatorname{argmin}} P_1(u)$ (vidi Sliku 15b).

U sljedećem (drugom) koraku definiramo

$$
\begin{aligned}
K(u; u_2) &:= f(u_2) - L|u - u_2|,\\
P_2(u) &:= \max_{i=0,1,2} K(u; u_i) = \max\{K(u; u_2), P_1(u)\},
\end{aligned}
$$

i odredimo $u_3 \in \underset{u\in[a,b]}{\operatorname{argmin}} P_2(u)$;

Općenito, u k-tom koraku, poznavanjem čvorova $u_0, u_1, \ldots, u_k$ i funkcija $K(u; u_0)$, $K(u; u_1)$, $\ldots$, $K(u; u_{k-1})$ definiramo

$$
K(u; u_k) := f(u_k) - L|u - u_k|, \tag{9}
$$

$$
P_k(u) = \max_{i=0,\ldots,k} K(u; u_i) = \max\{K(u; u_k), P_{k-1}(u)\} \tag{10}
$$

i odredimo $u_{k+1} \in \underset{u\in[a,b]}{\operatorname{argmin}} P_n(k)$. Ako se $\min_{u\in[a,b]} P_n(k)$ postiže u više točaka, onda za u_{k+1} biramo bilo koju od njih. Pseudokod Algoritma 4 naveden je niže, a animaciju algoritma i odgovarajući modul možemo vidjeti u t.**??**, str.**??**).

Algorithm 4 (Algoritam slomljenih pravaca)

Require: $f\colon [a,b] \to \mathbb{R}, \quad L > 0$
Require: $u_0 \in [a,b]$; $n \geq 1$; $k = 0$;
{Učitati funkciju i dozvoljeni broj iteracija;}
Definirati $K(u;u_0) := f(u_0) - L|u - u_0|$ i $P_0(u) := K(u;u_0)$;
Odrediti $u_1 \in \operatorname*{argmin}_{u\in[a,b]} P_0(u) \subseteq \{a,b\}$;
while $k = k+1; k < n$, **do**
    Definirati $K(u;u_k) := f(u_k) - L|u - u_k|$;
    $P_k(u) := \max\limits_{i=0,1,\ldots,k} K(u;u_i) = \max\{K(u;u_k), P_{k-1}(u)\}$;
    Odrediti $u_{k+1} \in \operatorname*{argmin}_{u\in[a,b]} P_k(u)$;
end while
Ensure: $\{u_{k+1}, f(u_{k+1})\}$

Sljedeća lema pokazuje da Algoritam 4 za funkciju $f \in Lip_L[a,b]$ daje rastući niz donjih ograda (P_k), $P_k \in Lip_L[a,b]$, (po dijelovima linearne funkcije čiji dijelovi leže na pravcima s koeficijentima smjera L ili $(-L)$), koje se s funkcijom f podudaraju u točkama u_k.

Lema 1. *Neka je $f \in Lip_L[a,b]$, (u_n) niz čvorova i (P_n) niz funkcija definiranih kao u Algoritmu 4. Tada vrijedi:*

(i) *P_n je Lipschitz-neprekidna po dijelovima linearna funkcija, čiji dijelovi leže na pravcima s koeficijentima smjera L ili $(-L)$, tj. $P_n \in Lip_L[a,b]$,*

(ii) $P_n(u) \leq P_{n+1}(u)$, $\forall u \in [a,b]$ *i* $\forall n = 0,1,\ldots$ *(monotoni rast),*

(iii) $P_n(u) \leq f(u)$, $\forall u \in [a,b]$ *i* $\forall n = 0,1,\ldots$ *(ograničenost odozgo),*

(iv) $P_n(u_i) = f(u_i)$, $\forall i = 0,1,\ldots,n$ *(podudaranje u čvorovima).*

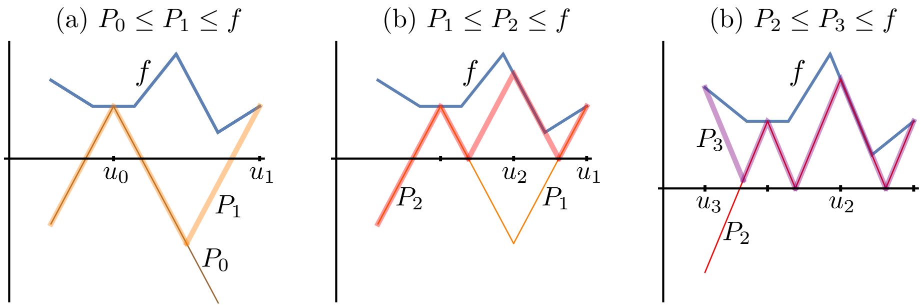
Dokaz. Tvrdnje *(i)* i *(ii)* su očigledne (vidi Sliku 16).

Tvrdnja *(iii)* dokazuje se induktivno. Najprije primijetimo da prema (8) vrijedi

$$P_0(u) = K(u;u_0) \leq f(u),$$

i pretpostavimo da je

$$P_{n-1}(u) \leq f(u).$$



Slika 16: Monotonost i ograničenost niza funkcija (P_n)

Kako je sukladno (8), $K(u; u_n) \leq f(u)$, vrijedi

$$P_n(u) = \max\{K(u; u_n), P_{n-1}(u)\} \leq \max\{f(u), f(u)\} = f(u),$$

tvrdnja *(iii)* vrijedi za svaki $n \in \mathbb{N}$.

U svrhu dokaza tvrdnje *(iv)* primijetimo najprije da vrijedi

$$f(u_i) = K(u_i; u_i), \quad \forall i \in \{0, 1, \ldots, n\}. \tag{11}$$

Nadalje, na osnovi (10) vrijedi $[P_n(u) := \max\limits_{i=0,1,\ldots,n} K(u; u_i) \geq K(u; u_i)]$

$$K(u_i; u_i) \leq P_n(u_i), \quad \forall i \in \{0, 1, \ldots, n\},$$

iz čega zaključujemo

$$f(u_i) = K(u_i; u_i) \leq P_n(u_i), \quad \forall i \in \{0, 1, \ldots, n\}.$$

Zato korištenjem *(iii)* dobivamo

$$f(u_i) \leq P_n(u_i) \overset{(iii)}{\leq} f(u_i),$$

iz čega slijedi tražena tvrdnja *(iv)*. □

Sljedeći teorem pokazuje da Algoritam 4 pronalazi sve točke globalnog minimuma funkcije $f \in Lip_L[a, b]$.

Teorem 2. *Pretpostavimo:*

(a) $f \in Lip_L[a, b]$;

(b) Postoji $u^\star \in \operatorname*{argmin}\limits_{u \in [a,b]} f(u)$;

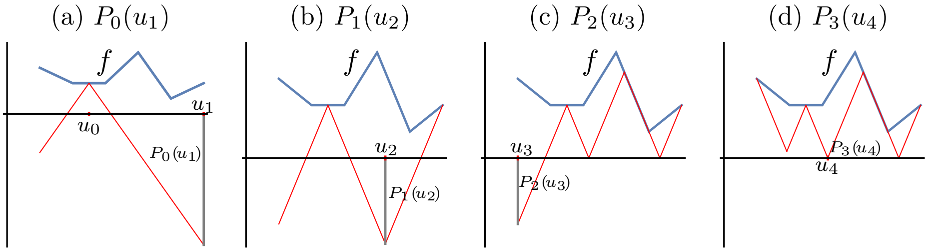
(c) Niz realnih brojeva (u_n) *i niz funkcija* (P_n) *definiran je kao u Algoritmu 4.*

Tada vrijedi:

(i) $\lim\limits_{n\to\infty} P_n(u_{n+1}) = \min\limits_{u\in[a,b]} f(u) = f(u^\star) =: f^\star$;

(ii) *Ako je* $\hat{u}$ *proizvoljno gomilište niza* (u_n), *onda je* $\hat{u} \in \operatorname{argmin}\limits_{u\in[a,b]} f(u)$;

(iii) *Ako je* $u^\star = \operatorname{argmin}\limits_{u\in[a,b]} f(u)$ *(jednočlan), onda cijeli niz* (u_n) *konvergira prema* $u^\star$.



Slika 17: Monotonost i ograničenost niza $(P_n(u_{n+1}))$

Dokaz. Dokažimo najprije da je niz $(P_n(u_{n+1}))$ konvergentan. On je monotono rastući jer je (vidi Sliku 17)

$$\begin{aligned} P_{n-1}(u_n) &= \min_{u\in[a,b]} P_{n-1}(u) && \text{[po definiciji]} \\ &\le P_{n-1}(u_{n+1}) && \text{[po definiciji minimuma funkcije na segmentu]} \\ &\le P_n(u_{n+1}). && \text{[prema Lemi 1}\,\textit{(ii)}\text{]} \end{aligned}$$

Niz $(P_n(u_{n+1}))$ ograničen je odozgo jer vrijedi (vidi Sliku 17)

$$\begin{aligned} P_n(u_{n+1}) &= \min_{u\in[a,b]} P_n(u) && \text{[po definiciji]} \\ &\le \min_{u\in[a,b]} f(u) && \text{[prema Lemi 1}\,\textit{(iii)}\text{]} \\ &= f(u^\star) = f^\star. && \text{[prema pretpostavci } \textit{(b)}\text{]} \end{aligned}$$

Dakle, niz $(P_n(u_{n+1}))$ je konvergentan i postoji $\lim\limits_{n\to\infty} P_n(u_{n+1}) =: P^\star$, a prema prethodnoj nejednakosti slijedi

$$P^\star \le f^\star. \tag{12}$$

Pokažimo da je $P^\star = f^\star$. U tu svrhu najprije pokažimo da vrijedi nejednakost

$$0 \le f(u_i) - P_n(u_{n+1}) \le L|u_i - u_{n+1}|, \quad \forall n \in \mathbb{N} \text{ i } \forall i = 0,\ldots,n. \tag{13}$$

Naime, kako je $P_n(u_{n+1}) = \min\limits_{u\in[a,b]} P_n(u) \le P_n(u_i)$, za proizvoljni $n \in \mathbb{N}$ i proizvoljni $(0 \le i \le n)$ vrijedi

$$P_n(u_{n+1}) \le P_n(u_i) = f(u_i) \quad i = 0,1,\ldots,n \quad \text{[Lema 1}\,\textit{(iv)}\text{]},$$

iz čega slijedi

$$\begin{aligned} 0 &\leq f(u_i) - P_n(u_{n+1}) \\ &= P_n(u_i) - P_n(u_{n+1}) && \text{[prema Lemi 1\textit{(iv)}]} \\ &\leq L|u_i - u_{n+1}| && \text{[prema Lemi 1\textit{(i)}]} \end{aligned}$$

Kako je niz $(u_n) \subset [a,b]$ ograničen, postoji barem jedno gomilište $\hat{u}$ i podniz (u_{n_k}) koji konvergira prema tom gomilištu.

Primijetite da zbog $n_{k-1} < n_k$ vrijedi $n_{k-1} \leq n_k - 1$. Zato iz nejednakosti (13) za $n := n_k - 1$ i $i := n_{k-1}$ dobivamo

$$0 \leq f(u_{n_{k-1}}) - P_{n_k-1}(u_{n_k}) \leq L|u_{n_{k-1}} - u_{n_k}|,$$

odakle zbog neprekidnosti funkcije f i konvergentnosti nizova $(P_{n_k-1}(u_{n_k}))$ i (u_{n_k}) za $k \to \infty$ dobivamo

$$0 \leq f(\hat{u}) - P^\star \leq 0 \quad \Longrightarrow \quad f(\hat{u}) = P^\star.$$

Kako je $f^\star \leq f(\hat{u})$, iz prethodne jednakosti i nejednakosti (12) dobivamo

$$f^\star \leq f(\hat{u}) = P^\star \leq f^\star,$$

tj. $P^\star = \lim\limits_{n\to\infty} P_n(u_{n+1}) = f^\star = f(\hat{u})$. Time su tvrdnje *(i)* i *(ii)* dokazane.

Tvrdnja *(iii)* slijedi iz prve dvije. □

Primjedba 6. U svakom koraku *Metode slomljenih pravaca* treba riješiti minimizacijski problem za po dijelovima linearnu funkciju P_n (korak 6. u Algoritmu 4). To je moguće vrlo efikasno napraviti pretraživanjem vrhova funkcije P_n. U tu svrhu Algoritam 4 treba implementirati tako da u svakoj iteraciji ažurira listu koordinata vrhova

$$V_n = \{(v_1, z_1), \dots (v_m, z_m)\},$$

koji predstavljaju lokalne minimume funkcije P_n. Primjerice, za točku $u_0 \in (a,b)$ u početnoj iteraciji skup V_0 sastoji se od dva vrha: $(a, K(a; u_0))$ i $(b, K(b; u_0)$, gdje je $K(u; u_0)$ slomljeni pravac definiran u (6). Listu V_n treba organizirati tako da vrijedi $z_1 \geq \cdots \geq z_m$, iz čega slijedi da je točka globalnog minimuma u_n funkcije P_n jednostavno v_m.

U sljedećoj iteraciji promatramo funkciju $P_{n+1}(u) = \max\{K(u; u_n), P_n(u)\}$ koja se od P_n razlikuje u dva nova vrha, osim u slučaju kada je $u_n = a$ ili $u_n = b$. Lista vrhova ažurira se tako da se iz nje izbaci posljednji vrh $(v_m, z_m) = (u_n, P_n(u_n))$ i dodaju dva nova vrha. Označimo nove vrhove kao lijevi (v_l, z_l) i desni (v_r, z_r) tako da je $v_l < u_n < v_r$. Uočimo da je točka (v_l, z_l) sjecište funkcije $K(u; u_n)$ sa susjednim lijevim slomljenim pravcem $K(u; u_l)$ koji je određen nekim čvorom koji ćemo označiti s u_l. Sada ćemo pokazati da je (v_l, z_l) moguće izračunati bez poznavanja vrijednosti u_l.

S obzirom da je u_l lijevo od u_n, točka (v_l, z_l) mora biti sjecište pravaca $u \mapsto f(u_n) + L(u - u_n)$ i $u \mapsto f(u_l) - L(u - u_l)$. Lako je izračunati da je tada

$$\begin{aligned} v_l &= \frac{1}{2}(u_l + u_n) + \frac{1}{2L}(f(u_l) - f(u_n)), \\ z_l &= K(v_l; u_n) = \frac{1}{2}(f(u_l) + f(u_n)) - \frac{L}{2}(u_n - u_l). \end{aligned} \tag{14}$$

Uočimo nadalje da za posljednji vrh u listi vrhova V_n vrijedi $(v_m, z_m) = (u_n, P_n(u_n)) = (u_n, K(u_n; u_l))$. Stoga (14) možemo zapisati kao

$$\begin{aligned} v_l &= u_n + \frac{1}{2L}(f(u_l) - L(u_n - u_l) - f(u_n)) \\ &= u_n + \frac{1}{2L}(K(u_n; u_l) - f(u_n)) \\ &= v_m + \frac{1}{2L}(z_m - f(v_m)), \\ z_l &= \frac{1}{2}(f(u_l) - L(u_n - u_l) + f(u_n)) \\ &= \frac{1}{2}(K(u_n; u_l) + f(u_n)) \\ &= \frac{1}{2}(z_m + f(v_m)). \end{aligned}$$

Analogno se može pokazati da se desni vrh (v_r, z_r) može izračunati kao

$$\begin{aligned} v_r &= v_m - \frac{1}{2L}(z_m - f(v_m)), \\ z_r &= \frac{1}{2}(z_m + f(v_m)). \end{aligned}$$

Dakle, u svakoj iteraciji treba ažurirati listu vrhova izbacivanjem posljednjeg vrha (v_m, z_m) i dodavanjem dva nova vrha čije se koordinate jednostavno izračunavaju iz v_m, z_m i $f(v_m)$. Nakon toga dobivenu listu treba sortirati po drugoj komponenti što daje listu V_{n+1} na kojoj se cijeli postupak ponavlja u sljedećoj iteraciji. Ukoliko je u nekoj iteraciji $v_m = a$, tada se dodaje samo desni vrh (v_r, z_r), a ako je $v_m = b$, onda se dodaje samo (v_l, z_l).

Ovakva implementacija Algoritma 4 napravljena je u modulu `Pijavskij` (vidi t.**??**, str.**??**).

Primjedba 7. Najpoznatije varijante *Metode slomljenih pravaca* su: *Algoritam Pijavskog*, *Shubertov algoritam* i *Algoritam* `DIRECT` (t.**??**, str.**??**). *Algoritam* `DIRECT` može se generalizirati za funkcije više varijabli [1, 2, 4], dok za *Algoritam Pijavskog* i *Shubertov algoritam* to nije moguće. *Metoda slomljenih pravaca* ima niz prednosti:

1. Minimizirajuća funkcija može imati više točaka globalnog minimuma, a *Metoda slomljenih pravaca* sve će ih detektirati. Primjerice, funkcija $x \to |\sin 30x|$ na intervalu $[0, 1]$ postiže globalni minimum u 10 točaka.

2. Minimizirajuća funkcija ne mora biti derivabilna u svim točkama intervala $[a, b]$, već samo ispunjavati Lipschitzov uvjet (4).

3. Za proizvoljni $u_0 \in [a, b]$ metoda konvergira prema globalnom minimumu, a za njenu implementaciju potrebno je „samo” poznavanje Lipschitzove konstante L.

4. U svakom koraku metode treba riješiti minimizacijski problem za po dijelovima linearnu funkciju P_n, što se svodi na ispitivanje njenih poznatih vrhova. Pri tome se P_n od P_{n-1} može razlikovati u najviše dva nova vrha (vidi Primjedbu 6, str.22).

5. Može se pokazati da je *Metoda slomljenih pravaca* bliska optimalnoj strategiji traženja globalnog minimuma Lipschitz-neprekidne funkcije.

6. Metoda se može poopćiti za slučaj funkcija koje imaju konačno mnogo prekida prve vrste i koje su definirane na nepovezanom području [9].

Nedostaci *Metode slomljenih pravaca* su:

1. Određivanje Lipschitzove konstante L može biti složeno i numerički zahtjevno (vidi [11**?**]).

2. Općenito, metoda brzo dolazi do točke globalnog minimuma, ali nakon toga često se dešava da sporo popravlja točnost (vidi primjerice [6]). Razlozi za to su sljedeći.

   - Velike vrijednost Lipschitzove konstante $L > 0$. U graničnom slučaju kada $L \to \infty$, algoritam degenerira u izračunavanje vrijednosti funkcije na diskretnom skupu.
   - Svojstva algoritma da pronalazi sve točke globalnog minimuma, što može značajno opteretiti računski postupak.

3. Određivanje kriterija zaustavljanja nije jasno definirano u literaturi. Mogućnosti su sljedeće:

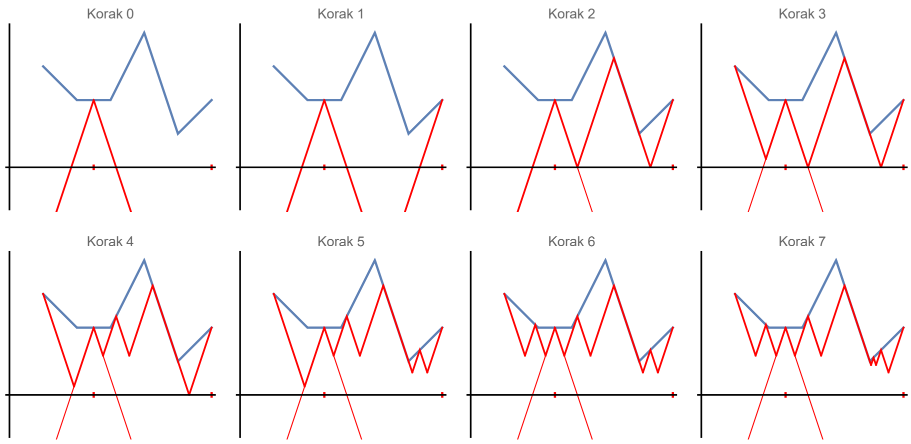
   - Unaprijed odrediti maksimalni broj iteracija;
   - Postupak završiti u k-tom koraku ako je $|u_{k+1} - u_k| < \epsilon$, gdje je $\epsilon > 0$ unaprijed zadan;
   - Postupak završiti u k-tom koraku ako je $|f(u_{k+1}) - f(u_k)| < \epsilon$, gdje je $\epsilon > 0$ unaprijed zadan.

Primjedba 8. Metodu možemo modificirati tako da na početku izaberemo $n + 1$ ekvidistantnih čvorova $a = u_0 < u_1 < \cdots < u_n = b$ u kojima definiramo:

$$
\begin{aligned}
&K(u; u_i) := f(u_i) - L|u - u_i|, \quad i = 0, 1, \ldots, n \\
&P_n(u) := \max_{i=0,1,\ldots,n} K(u; u_i) \\
&u_{n+1} = \operatorname*{argmin}_{u \in [a,b]} P_n(u) \qquad\qquad \text{[pretraživanje } n+1 \text{ čvorova]}
\end{aligned}
$$

Dalje postupak može teći kao u Algoritmu 4.

Primjer 9. *Metodu slomljenih pravaca ilustrirat ćemo na traženju globalnog minimuma funkcije iz Primjera 8, str.14.*



Slika 18: *Metoda slomljenih pravaca*

Iterativni proces može se pratiti u niže navedenoj tablici i na Slici 18 (pogledati također t.**??**, str.**??**). Može se vidjeti da je nakon 12 iteracija udaljenost susjednih čvorova, kao i apsolutna razlika funkcijskih vrijednosti pala na .05.

k	0	1	2	3	4	5	6	7	8	9	10	11	12
u_k	2.5	6	4.25	1	3.17	5.33	1.92	5.11	5.56	5.037	5.19	5.01	5.06
$\lvert u_{k+1} - u_k\rvert$	-	3.5	1.75	3.25	2.17	2.17	3.42	3.19	0.44	0.52	0.15	0.17	0.05
$f(u_k)$	2	2	3.25	3	2.33	1.33	2.08	1.11	1.56	1.04	1.19	1.01	1.06
$\lvert f(u_{k+1}) - f(u_k)\rvert$	-	0	1.25	0.25	0.67	1.	0.75	0.97	0.44	0.52	0.15	0.17	0.05

Zadatak 19. *Za funkciju* $f\colon [0,9] \to \mathbb{R}$ *zadanu s*

$$f(x) = \begin{cases} 5 - x, & 0 \le x \le 3, \\ 4 - \frac{1}{2}(x-5)^2, & 3 < x \le 7, \\ \frac{1}{2}(x-9)^2, & 7 < x \le 9, \end{cases}$$

Metodom slomljenih pravaca *odredite* u_3 *uz početnu točku* $u_0 = 4$.

Zadatak 20. *Za* $u_0 = -6$ Metodom slomljenih pravaca *odredite prve tri aproksimacije točke globalnog minimuma funkcije* $f\colon [-6,12] \to \mathbb{R}$ *zadane s*

$$f(x) = \min\left\{|x-1|+1, |x-8|+2\right\}.$$

Literatura

[1] D. E. Finkel, *DIRECT Optimization Algorithm User Guide*, Center for Research in Scientific Computation. North Carolina State University, 2003, `http://www4.ncsu.edu/definkel/research/index.html`.

[2] J. M. Gablonsky, *Direct version 2.0*, Technical report, Center for Research in Scientific Computation. North Carolina State University, 2001.

[3] F. Jarre, J. Stoer, *Optimierung*, Springer Verlag, Berlin, Heidelberg, 2004.

[4] D. R. Jones, C. D. Perttunen, B. E. Stuckman, *Lipschitzian optimization without the Lipschitz constant*, Journal of Optimization Theory and Applications, **79**(1993) 157–181.

[5] J. M. Ortega, W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, SIAM, Philadelphia, 2000.

[6] R. Paulavičius, Y. Sergeyev, D. Kvasov, J. Žilinskas, *Globally-biased DISIMPL algorithm for expensive global optimization*, Journal of Global Optimization, **59**(2014) 545–567.

[7] R. Scitovski, *Numerička matematika*, Odjel za matematiku, Sveučilište u Osijeku, 3, izdanje, 2015, `http://www.mathos.unios.hr/index.php/odjel/nasa-izdanja?getBook=541`.

[8] J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, New York, 2002.

[9] R. G. Strongin, Y. D. Sergeyev, *Global Optimization with Non-Convex Constraints: Sequential and Parallel Algorithms*, Kluwer Academic Publishers, Dordrecht, 2000.

[10] Šime Ungar, *Matematička analiza 3*, Sveučilište u Zagrebu, Prirodoslovno–matematički fakultet, Matematički odjel, 2002.

[11] G. R. Wood, B. P. Zhang, *Estimation of the Lipschitz constant of a function*, Journal of Global Optimization, **8**(1996) 91–103.