

Odjel za matematiku

Sveučilište J. J. Strossmayera u Osijeku

Zajednički sveučilišni poslijediplomski doktorski studij matematike

Neki specijalni optimizacijski problemi

Domagoj Matijević, Kristian Sabo, Rudolf Scitovski

Globalna optimizacija Lipschitz-neprekidnih funkcija

1 Lipschitz-neprekidne funkcije

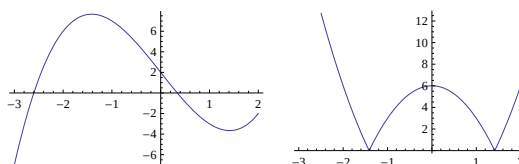
U ovom odjeljku razmatrat ćemo jednu metodu traženja globalnog minimuma realne funkcije jedne varijable, koja na intervalu $[a, b]$ zadovoljava Lipschitzov uvjet. Metoda potječe od ruskog matematičara Pijavskog (Danilin, 1967; Pijavskij, 1972). Vidi također (Vasilev, 1974).

Definicija 1. Kažemo da funkcija $f : [a, b] \rightarrow \mathbb{R}$ *Lipschitz-neprekidna* (zadovoljava *Lipschitzov uvjet*) s konstantom $L > 0$ na $[a, b]$ i pišemo $f \in \text{Lip}_L[a, b]$ ako za svaki $x, y \in [a, b]$ vrijedi

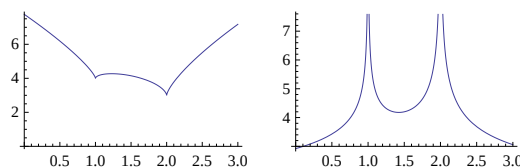
$$|f(x) - f(y)| \leq L|x - y|. \quad (1)$$

Primjedba 1. Svaka Lipschitz-neprekidna funkcija je neprekidna funkcija. Ako je $f \in C^1[a, b]$, onda je $f \in \text{Lip}_L[a, b]$ (Lagrangeov teorem o srednjoj vrijednosti), ali obrat ne vrijedi (vidi Primjer 1 i Primjer 3). Ako je $f : [a, b] \rightarrow \mathbb{R}$ neprekidna funkcija, koja ima po dijelovima neprekidnu derivaciju, onda primjenom Lagrangeovog teorema o srednjoj vrijednosti dobivamo konstantu $L = \sup_{x \in [a, b]} |f'(x)|$ (Primjer 3).

Primjer 1. Treba odrediti Lipschitzovu konstantu funkcije $f : [-3, 2] \rightarrow \mathbb{R}$, $f(x) = x^3 - 6x + 2$. Iz Slike 1 može se vidjeti da je $L = |f'(-3)| = 21$.



Slika 1: Funkcija $f(x) = x^3 - 6x + 2$ i $x \mapsto |f'(x)|$

Slika 2: Funkcija $f(x) = 1 + 2\sqrt[3]{(x-1)^2} + 3\sqrt[3]{(x-2)^2}$ i $x \mapsto |f'(x)|$

Primjer 2. Treba odrediti Lipschitzovu konstantu funkcije $f: [0, 3] \rightarrow \mathbb{R}$, $f(x) = 1 + 2\sqrt[3]{(x-1)^2} + 3\sqrt[3]{(x-2)^2}$. Njena derivacija

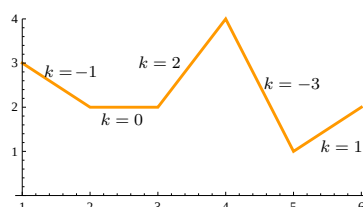
$$f'(x) = \frac{4}{3\sqrt[3]{x-1}} + \frac{2}{\sqrt[3]{x-2}}$$

nije definirana u točkama iz $\{1, 2\}$, a Lipschitzova konstanta bila bi $L = \infty$.

Primjer 3. Treba odrediti Lipschitzovu konstantu funkcije $f: [1, 6] \rightarrow \mathbb{R}$,

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

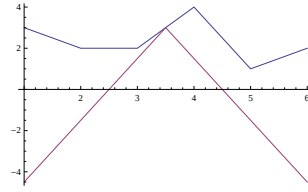
čiji graf je prikazan na Slici 3. Očigledno je $L = 3$.

Slika 3: Koeficijenti smjera po dijelovima linearne funkcije f

2 Metoda slomljenih pravaca

Pretpostavimo da postoji $u^* \in [a, b]$, na kome se postiže globalni minimum funkcije $f: [a, b] \rightarrow \mathbb{R}$, tj. postoji $u^* = \underset{u \in [a, b]}{\operatorname{argmin}} f(u)$. Sukladno (Pijavskij, 1972) za fiksni $u_0 \in [a, b]$ definiramo funkciju $K: [a, b] \rightarrow \mathbb{R}$,

$$K(u, u_0) = f(u_0) - L|u - u_0|, \quad u \in [a, b], \quad (2)$$



Slika 4: Lipschitz-donja ograda funkcije

koja ima sljedeća svojstva:

$$u \mapsto K(u, u_0) \quad \text{po dijelovima je linearna funkcija,} \quad (3)$$

$$K(u_0, u_0) = f(u_0), \quad (4)$$

$$K(u, u_0) \leq f(u), \quad \forall u \in [a, b], \quad (5)$$

Svojstvo (5) slijedi iz

$$(K(u, u_0) - f(u) = f(u_0) - f(u) - L|u - u_0| \leq |f(u_0) - f(u)| - L|u - u_0| \leq L|u - u_0| - L|u - u_0| = 0).$$

Algoritam 1. (“Slomljeni pravci”)

Korak 0: Izabrati $u_0 \in [a, b]$ i definirati $K(u, u_0) := f(u_0) - L|u - u_0|$;

Odrediti $u_1 = \operatorname{argmin}_{u \in [a, b]} K(u, u_0)$ (Očigledno je $u_1 = a$ ili $u_1 = b$).

Korak 1: Definirati $K(u, u_1) := f(u_1) - L|u - u_1|$ i

$$P_1(u) := \max_{i=0,1} K(u, u_i) = \max\{K(u, u_1), P_0(u)\}, \quad P_0(u) := K(u, u_0);$$

Odrediti $u_2 = \operatorname{argmin}_{u \in [a, b]} P_1(u)$.

Korak 2: Definirati $K(u, u_2) := f(u_2) - L|u - u_2|$ i

$$P_2(u) := \max_{i=0,1,2} K(u, u_i) = \max\{K(u, u_2), P_1(u)\};$$

Odrediti $u_3 = \operatorname{argmin}_{u \in [a, b]} P_2(u)$.

Korak n: Uz pretpostavku da su $u_0, u_1, \dots, u_n$, $K(u, u_0), K(u, u_1), \dots, K(u, u_n)$ ($n \geq 1$) poznati, definirati

$$P_n(u) = \max_{i=0, \dots, n} K(u, u_i) = \max\{K(u, u_n), P_{n-1}(u)\}, \quad P_0(u) = K(u, u_0); \quad (6)$$

Odrediti $u_{n+1} = \operatorname{argmin}_{u \in [a, b]} P_n(u)$. (Ako se $\min_{u \in [a, b]} P_n(u)$ postiže u više točaka, onda za u_{n+1} biramo bilo koju od njih).

Sljedeća lema daje svojstva nizova (u_n) i (P_n) iz Algoritma 1 i $f \in \operatorname{Lip}_L[a, b]$.

Lema 1. Neka je niz realnih brojeva (u_n) i niz funkcija (P_n) definiran kao u Algoritmu 1. Tada vrijedi:

- (i) P_n je neprekidna po dijelovima linearna funkcija, čiji dijelovi leže na pravcima s koeficijentima smjera L ili $(-L)$;
- (ii) $P_n(u) \leq P_{n+1}(u)$, $\forall u \in [a, b]$ i $\forall n = 0, 1, \dots$;
- (iii) $P_n(u) \leq f(u)$, $\forall u \in [a, b]$ i $\forall n = 0, 1, \dots$;
- (iv) $P_n(u_i) = f(u_i)$, $\forall i = 0, 1, \dots, n$
- (v) $P_n \in Lip_L[a, b]$.

Dokaz. Tvrdnje (i) i (ii) su očigledne. Tvrdnja (iii) slijedi induktivno na osnovi (5):

$$\begin{aligned}
 P_0(u) &= K(u, u_0) \leq f(u), \\
 P_1(u) &= \max\{K(u, u_1), P_0(u)\} \leq \max\{f(u), f(u)\} = f(u), \\
 &\dots \\
 P_n(u) &= \max\{K(u, u_n), P_{n-1}(u)\} \leq \max\{f(u), f(u)\} = f(u).
 \end{aligned}$$

Tvrdnja (iv) slijedi iz (4), (6) i (iii):

$$f(u_i) = K(u_i, u_i) \leq P_n(u_i) \leq f(u_i).$$

Tvrdnja (v) slijedi iz (i). □

Teorem 1. Neka za funkciju $f \in Lip_L[a, b]$ postoji $u^* \in \underset{u \in [a, b]}{\operatorname{argmin}} f(u)$. Neka je niz realnih brojeva (u_n) i niz funkcija (P_n) definiran kao u Algoritmu 1. Tada vrijedi:

- (i) $\lim_{n \rightarrow \infty} P_n(u_{n+1}) = \inf_{u \in [a, b]} f(u) = f(u^*) =: f^*$;
- (ii) Ako je $\bar{u}$ proizvoljno gomilište niza (u_n) , onda je $\bar{u} \in \underset{u \in [a, b]}{\operatorname{argmin}} f(u)$;
- (iii) Ako je $u^* = \underset{u \in [a, b]}{\operatorname{argmin}} f(u)$ (jednočlan), onda cijeli niz (u_n) konvergira prema u^* .

Dokaz. Dokažimo najprije da je niz $(P_n(u_{n+1}))$ monotonno rastuć

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

i ograničen odozgo:

$$\begin{aligned}
 P_n(u_{n+1}) &= \min_{u \in [a, b]} P_n(u) && \text{po definiciji} \\
 &\leq P_n(u^*) && \text{po definiciji} \\
 &\leq f(u^*) && \text{prem Lemi 1(iii)}
 \end{aligned}$$

Dakle, postoji $\lim_{n \rightarrow \infty} P_n(u_{n+1}) =: P^* \leq f^*$ (Lema 1(iii) i neprekidnost funkcija P_n i f).

Pokažimo da je $P^* = f^*$. Kako je $P_n(u_{n+1}) = \min_{u \in [a, b]} P_n(u)$, za proizvoljni $n \in \mathbb{N}$ i proizvoljni i ($0 \leq i \leq n$) vrijedi

$$P_n(u_{n+1}) \leq P_n(u_i) \leq f(u_i) \quad (\text{Lema 1(iii)}),$$

iz čega slijedi

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

Kako je niz $(u_n) \subset [a, b]$ ograničen, postoji barem jedno gomilište $\bar{u}$ i podniz (u_{n_k}) , koji konvergira prema tom gomilištu, pri čemu možemo pretpostaviti da je $n_1 < n_2 < \dots < n_k < \dots$. Specijalno, iz prethodnih nejednakosti za $n := n_k - 1$ i $i := n_{k-1}$ ($n_{k-1} \leq n_k - 1$), 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 za $k \rightarrow \infty$ dobivamo

$$f(\bar{u}) = \lim_{k \rightarrow \infty} P_{n_k-1}(u_{n_k}).$$

Kako je $f^* \leq f(\bar{u})$ i $P^* \leq f^*$, onda iz prethodne jednakosti dobivamo

$$f^* \leq f(\bar{u}) = P^* \leq f^*,$$

tj. $P^* = \lim_{n \rightarrow \infty} P_n(u_{n+1}) = f^* = f(\bar{u})$. Time su tvrdnje (i) i (ii) dokazane.

Tvrđnja (iii) slijedi iz prve dvije. □

Zadatak 1. *Napišite detaljan algoritam za globalnu optimizaciju Lipschitzove funkcije uz primjenu metode slomljenih pravaca (Pijavskij, 1972) i odgovarajući kompjuterski program. Program testirajte na nekoliko primjera.*

Primjedba 2. Najpoznatije varijante metode slomljenih pravaca su *Metoda Pijavskog* (Danilin, 1967; Pijavskij, 1972), *Shubertova metoda* (Shubert, 1972), koje se nažalost ne mogu direktno generalizirati na više dimenzija i Algoritam **DIRECT** (Jones et al., 1993; Kvasov and Sergeyev, 2009), koji je konstruiran u svrhu generalizacije na više dimenzija (Finkel and Kelley, 2004, 2006; Vanderbei, 1999). Metoda slomljenih pravaca ima niz prednosti:

1. Minimizirajuća funkcija f može imati proizvoljno stacionarnih točaka i lokalnih minimuma, ne mora biti derivabilna u svim točkama intervala $[a, b]$, već samo mora ispunjavati Lipschitzov uvjet (1);
2. Za proizvoljni $u_0 \in [a, b]$ metoda konvergira prema globalnom minimumu, a za njenu implementaciju potrebno je samo poznavati Lipschitzovu konstantu L ;

3. 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 P_n se od P_{n-1} može razlikovati za najviše dva nova vrha;
4. Može se pokazati (Vasilev, 1974) da je metoda slomljenih pravaca bliska optimalnoj strategiji traženja globalnog minimuma Lipschitzove funkcije.
5. Metodu možemo modificirati tako da niz $(P_n(u_{n+1}))$ bude strogo monotono rastuć (???); Ovo mi se čini teško za postići, jer kad god podignemo donji vrh funkcije P_n na krivulju dobijemo dva nova donja vrha s istom vrijednosti funkcije P_n .
6. Metodu možemo modificirati tako da na početku izaberemo $n + 1$ ekvidistantnih čvorova $a = u_0 < u_1 < \dots < u_n = b$ u kojima definiramo:

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

Dalje procedura može teći kao u Algoritmu 1;

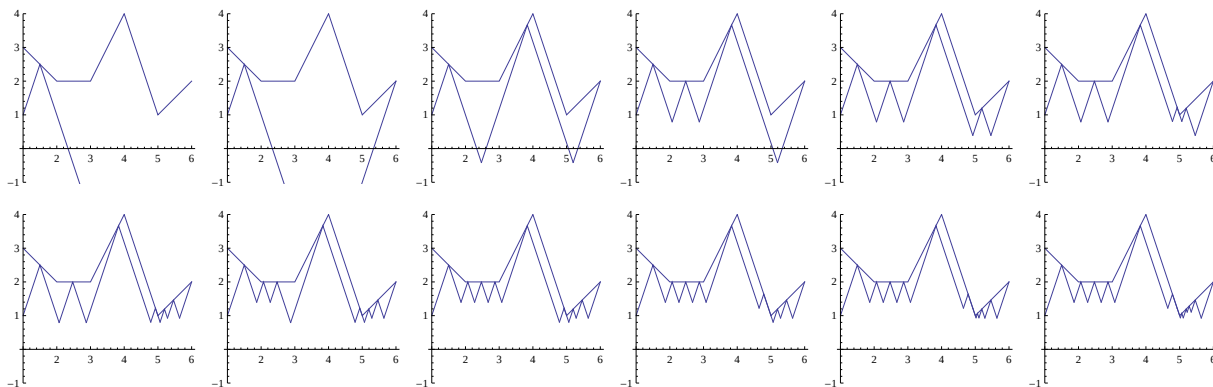
7. Metoda se može poopćiti za slučaj funkcija koje imaju konačno prekida prve vrste i koje su definirane na nepovezanom području (Strongin and Vasiljev, 1984),

i nedostataka:

1. Određivanje Lipschitzove konstante L može biti složeno. O jednoj metodi može se vidjeti u (Wood and Zang, 1996);
2. Niska brzina konvergencije. Čim je L veći, brzina je niža – u graničnom slučaju kada $L \rightarrow \infty$ algoritam degenerira u izračunavanje vrijednosti funkcije na diskretnom skupu;
3. Računska kompleksnost u više dimenzija.

Primjer 4. Metodu slomljenih pravaca ilustrirat cemo na traženju globalnog minimuma funkcije iz Primjera 3. Iterativni proces može se pratiti na Slici 5 i u niže navedenoj tablici.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
u_i	1.5	6	3.83	2.47	5.19	4.93	5.46	2.07	2.88	4.79	5.06	5.28	5.64	5.02	5.11	5.007
$f(u_i)$	2.5	2	3.67	2.	1.19	1.22	1.46	2.	2.	1.64	1.06	1.28	1.64	1.02	1.11	1.007



Slika 5: Metoda slomljenih pravaca

3 Shubertova metoda

Lipschitzov uvjet (1) možemo zapisati u obliku

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

čime je relativna promjena (kvocijent diferencija) funkcije f određena između $-L$ i L . Ako za y stavimo a , onda je $x - a \geq 0$, pa je

$$|f(x) - f(a)| \leq L(x - a), \quad \forall x \in [a, b],$$

odakle dobivamo

$$f(x) \geq f(a) - L(x - a). \quad (7)$$

Ako za y stavimo b , onda je $x - b \leq 0$, pa je

$$|f(x) - f(b)| \leq -L(x - b), \quad \forall x \in [a, b],$$

odakle dobivamo

$$f(x) \geq f(b) + L(x - b). \quad (8)$$

Desne strane u (7) i (8) su pravci koji se sijeku u točki (U, B) , gdje je

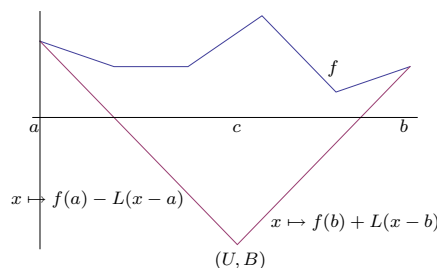
$$U(a, b, f, L) = \frac{1}{2}(a + b) + \frac{1}{2L}(f(a) - f(b)), \quad (9)$$

$$B(a, b, f, L) = \frac{1}{2}(f(a) + f(b)) - \frac{L}{2}(b - a), \quad (10)$$

a graf funkcije f je iznad tog slomljenog pravca. Naime, zbrajanjem (7) i (8) dobivamo

$$f(x) \geq \frac{1}{2}(f(a) + f(b)) - \frac{L}{2}(b - a) = B(a, b, f, L).$$

Algoritam 2. (Shubertova metoda, (Shubert, 1972))

Slika 6: Donja ograda funkcija f na intervalu $[a, b]$

Korak 1: Odrediti pravce $x \mapsto f(a) - L(x - a)$ i $x \mapsto f(b) + L(x - b)$ i točku njihovog sjecišta $(U(a, b, f, L), B(a, b, f, L))$

Korak 2: Staviti $u_1 = U(a, b, f, L)$. Od dva intervala $[a, u_1]$, $[u_1, b]$ izabrati onaj s manjom B -vrijednosti.

Korak 3: Ako je $[a, u_1]$ izabrani interval, staviti $u_2 = U(a, u_1, f, L)$.

Od tri intervala: $[a, u_2]$, $[u_2, u_1]$, $[u_1, b]$ izabrati onaj s najmanjom B -vrijednosti i na njemu ponoviti proceduru.

Primjedba 3. Primijetite da je Shubertova metoda zapravo specijalni slučaj Algoritma 1 ako za početnu aproksimaciju izaberemo jedan od rubova intervala $[a, b]$. Neka je primjerice, $u_{-1} = a$. Tada je prema Metodi slomljenih pravaca

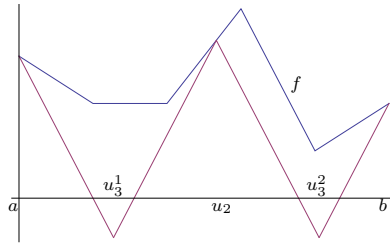
$$\begin{aligned} u_0 &= \operatorname{argmin}\{f(a) - L|u - a|\} = b, \\ P_1(u) &= \max\{f(a) - L|u - a|, f(b) - L|u - b|\}, \\ u_1 &= \operatorname{argmin} P_1(u). \end{aligned}$$

Nakon toga definiramo

$$\begin{aligned} P_2(u) &= \max\{P_1(u), f(u_1) - L|u - u_1|\}, \\ u_2 &= \operatorname{argmin} P_2(u), \end{aligned}$$

koji se postiže u jednoj od dviju vrijednosti u_2^1 ili u_2^2 , za koju je vrijednost funkcije f manja. (vidi Sliku 7). Očigledno se odluka može donijeti poznavanjem B -vrijednosti funkcije f na intervalima $[a, u_1]$ i $[u_1, b]$. Ovaj kriterij podudara se s kriterijem navedenim u *Koraku 2* Shubertovog algoritma.

Zadatak 2. Napišite detaljan algoritam za globalnu optimizaciju Lipschitzove funkcije uz primjenu Shubertove metode (Jones et al., 1993) i odgovarajući kompjuterski program. Program testirajte na nekoliko primjera.



Slika 7: Veza Shubertove metode i Metode slomljenih pravaca

4 DIRECT optimizacijski algoritam

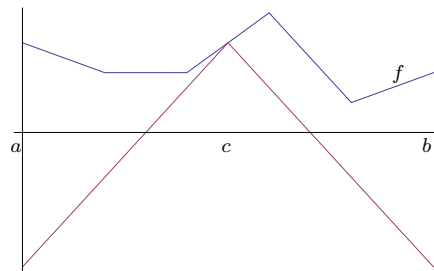
Prema (Jones et al., 1993) Shubertovu metodu modificirat ćemo tako da je moguća generalizacija na višedimenzionalni slučaj. Naziv dolazi od: DI(viding)RECT(angles). Slično kao što smo izveli formule (9) i (10), za $c = \frac{a+b}{2}$ dobivamo

$$f(x) \geq f(c) + L(x - c), \quad x \leq c, \quad (11)$$

$$f(x) \geq f(c) - L(x - c), \quad x \geq c \quad (12)$$

Graf funkcije f nalazi se iznad slomljene linearne funkcije čija je najmanja vrijednost na intervalu $[a, b]$ zadana s (vidi Sliku 8)

$$\mathcal{B}(c) = f(c) - L \frac{b-a}{2}. \quad (13)$$

Slika 8: Donja ograda funkcije f na intervalu $[a, b]$

Primijetite da uz poznavanje Lipschitzove konstante donja ograda $\mathcal{B}$ funkcije f ovisi samo o vrijednosti funkcije f u centru intervala $[a, b]$ i širini intervala. Ovo će biti važno prilikom konstrukcije optimizacijskog algoritma, ali posebno i kod generalizacije na višedimenzionalni slučaj.

Algoritam 3. (DIRECT algoritam, (Jones et al., 1993))

Korak 1: Interval $[a, b]$ podijeliti na $n = 3$ podintervala $[a_1, b_1]$, $[a_2, b_2]$, $[a_3, b_3]$, jednake duljine $d = \frac{b-a}{3}$ s centrima u točkama

$$c_1 = a + \frac{d}{2}, \quad c_2 = a + 3\frac{d}{2}, \quad c_3 = a + 5\frac{d}{2};$$

Prema (13) izračunati $\mathcal{B}$ -vrijednosti

$$\mathcal{B}(c_i) = f(c_i) - L\frac{b_i - a_i}{2}, \quad i = 1, \dots, n;$$

Odrediti i_0 , tako da bude $f(c_{i_0}) = \min_{i=1, \dots, n} f(c_i)$ i staviti $\min = \{c_{i_0}, f(c_{i_0})\}$;

Korak 2: Odrediti $j \in \{1, \dots, n\}$ tako da bude $\mathcal{B}(c_j) = \min_{i=1, \dots, n} \mathcal{B}(c_i)$;

Korak 3: Interval s centrom u c_j podijeliti na 3 podintervala jednake duljine $\frac{d}{3}$;

Na taj način umjesto intervala duljine d s centrom u c_j dobivamo tri nova podintervala duljine $\frac{d}{3}$ s centrima u $c_j - \frac{d}{3}$, c_j , $c_j + \frac{d}{3}$;

Korigirati $\min$ s vrijednosti funkcije u dva nova centra $c_j - \frac{d}{3}$, $c_j + \frac{d}{3}$;

Izračunati $\mathcal{B}$ -vrijednosti za tri nova podintervala. **Primijetite da je za to potrebno poznavati samo centre i širine podintervala!**;

Staviti $n = n + 2$ i prijeći na *Korak 2*;

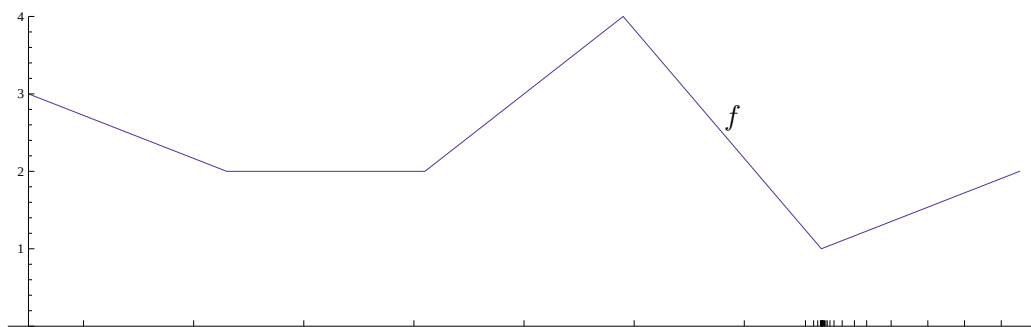
Primjedba 4. Prema Teoremu 1 proizvoljno gomilište niza (c_n) je točka u kojoj funkcija f postiže svoj globalni minimum.

Iterativni postupak može se ubrzati koristeći činjenicu da se u svakom koraku od postojećeg niza podintervala ispušta jedan i dodaju tri nova podintervala:

- U *Koraku 1* i *Koraku 3* intervale sortirati u opadajućem poretку prema njihovim $\mathcal{B}$ -vrijednostima.
- Tri nova podintervala dodati postojećem nizu tako da novi niz podintervala bude sortirati u opadajućem poretку prema njihovim $\mathcal{B}$ -vrijednostima. U *Koraku 2* tada je $j = n$.

Primjer 5. *Primjenom DIRECT optimizacijskog algoritma odredit ćemo globalni minimum funkcije $f: [1, 6] \rightarrow \mathbb{R}$ zadane u Primjeru 3. Na Slici 9 prikazan je niz aproksimacija (c_n) globalnog minimuma, koji se postiže u točki $x^* = 5$.*

Zadatak 3. *Napišite detaljan algoritam za globalnu optimizaciju Lipschitzove funkcije uz primjenu DIRECT metode (Jones et al., 1993) i odgovarajući kompjuterski program. Program testirajte na nekoliko primjera.*

Slika 9: Niz aproksimacija (c_n) globalnog minimuma funkcije iz Primjera 3

4.1 Optimizacija DIRECT algoritma

Primijetimo još jednom važnu činjenicu da smo podintervale $[a_1, b_1], \dots, [a_n, b_n]$ u algoritmu DIRECT reprezentirali smo njihovim centrima c_i , $i = 1, \dots, n$ i poluširinama intervala $d_i = \frac{b_i - a_i}{2}$. Zato ćemo svakom intervalu $[a_i, b_i]$ s centrom c_i pridružiti točku

$$T_i = \left(\frac{b_i - a_i}{2}, f(c_i) \right).$$

Donja ograda ($\mathcal{B}$ -vrijednost)

$$\mathcal{B}(c_i) = f(c_i) - L \frac{b_i - a_i}{2},$$

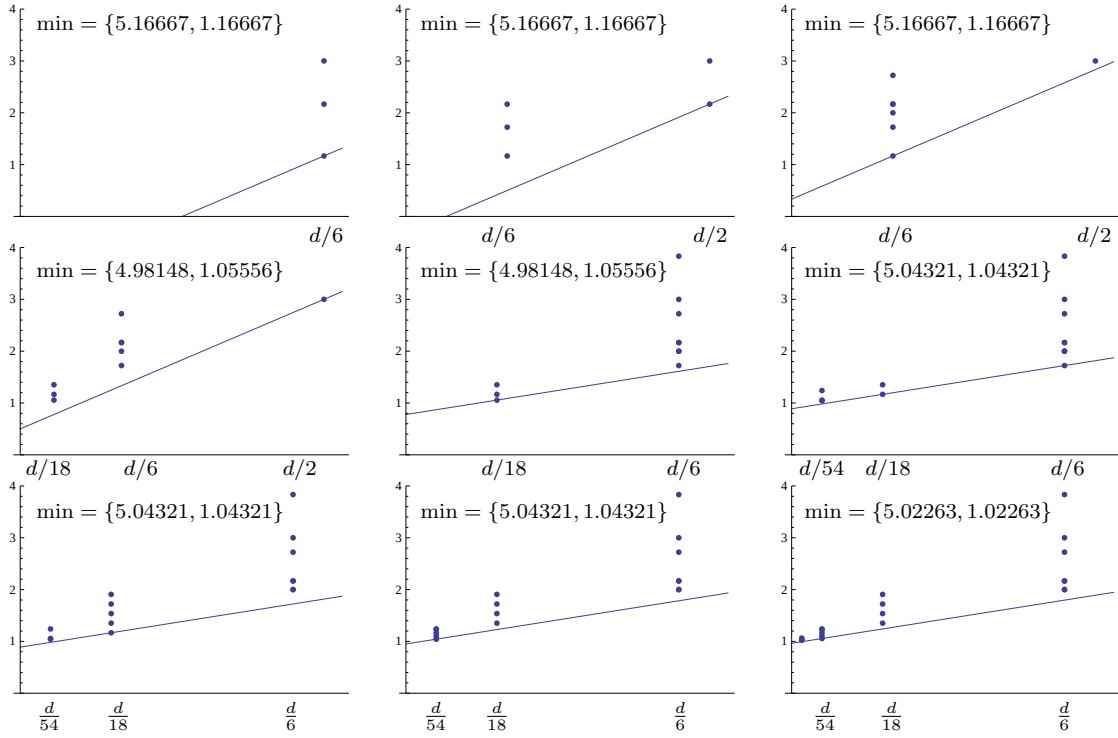
intervala $[a_i, b_i]$ može se interpretirati kao odsječak na ordinati linearne funkcije l_i čiji graf prolazi točkom T_i i ima koeficijent smjera L

$$l_i(t) = Lt + \left(f(c_i) - L \frac{b_i - a_i}{2} \right). \quad (14)$$

Svaki grafikon na Slici 10 prikazuje točke T_i , $i = 1, \dots, n$ dobivene u nekoj iteraciji algoritma DIRECT i pravac l_j kroz točku T_j i koeficijentom smjera L , pri čemu je j izabran tako da bude $\mathcal{B}(c_j) = \min_{i=1, \dots, n} \mathcal{B}(c_i)$ (interval s najmanjom $\mathcal{B}$ -vrijednosti – najmanjim odsječkom na ordinati). U prvoj iteraciji pojavljuje se $n = 3$ točke, u drugoj $n = 5$ točaka, u trećoj $n = 7$ točaka itd. Pri tome neke točke se podudaraju.

Već na drugom grafikonu Slike 10 vidi se da je bilo moguće izabrati “bolji pravac” – donju ogradu točaka T_i , $i = 1, \dots, n$ izborom nove točke T_j i manjeg koeficijenta smjera $0 < \hat{L} \leq L$.

Općenito, to znači da među intervalima $[a_i, b_i]$, $i = 1, \dots, n$ s centrima c_i , $i = 1, \dots, n$ treba izabrati *potencijalno optimalan interval* $[a_j, b_j]$ s centrom c_j . Geometrijski gledano, interval $[a_j, b_j]$ bit će potencijalno optimalan ako birajući $\hat{L} > 0$, mogu pronaći pravac kroz točku T_j koji je ujedno donja ograda za sve druge točke T_i , $i = 1, \dots, n$. Znat ćemo



Slika 10: Vrijednosti funkcije u ovisnosti o udaljenosti centra do ruba podintervala

da je taj pravac donja ograda ako je njegov odsječak na ordinati manji od odgovarajućih odsječaka za druge točke T_i .

Definicija 2. Interval $[a_j, b_j]$ s centrom c_j je potencijalno optimalan ako postoji $\hat{L} > 0$, tako da bude

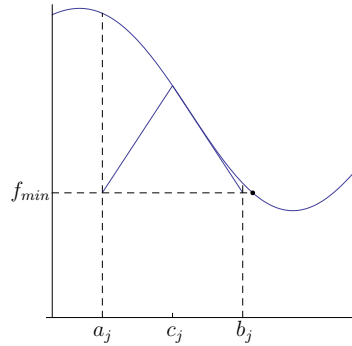
$$f(c_j) - \hat{L} \frac{b_j - a_j}{2} \leq f(c_i) - \hat{L} \frac{b_i - a_i}{2}, \quad \forall i = 1, \dots, n, \quad (15)$$

i da postoji $\epsilon > 0$, tako da bude

$$f(c_j) - \hat{L} \frac{b_j - a_j}{2} \leq f_{\min} - \epsilon |f_{\min}|, \quad (16)$$

gdje je $f_{\min} = \min_{i=1, \dots, n} f(c_i)$.

Primijetite da iz definicije $\mathcal{B}$ -vrijednosti slijedi $\mathcal{B}(c_j, d_j) = f(c_j) - \hat{L} \frac{b_j - a_j}{2} \leq f_{\min}$. Uvjet (16) garantira da se među potencijalno optimalnim intervalima neće pojaviti interval čija $\mathcal{B}$ -vrijednost ($\mathcal{B}(c_j, d_j) = f(c_j) - \hat{L} \frac{b_j - a_j}{2}$) nije manja od trenutnog minimuma funkcije $f_{\min}$ barem za $\epsilon > 0$. Na Slici 11 prikazan je interval $[a_j, b_j]$ s centrom u točki c_j , čija se $\mathcal{B}$ -vrijednost podudara s trenutnim minimumom funkcije $f_{\min}$. Takav interval ne spada u potencijalno optimalne. Geometrijsko objašnjenje uvjeta (16) može se također vidjeti na Fig 2.2 kod (Finkel and Kelley, 2004).

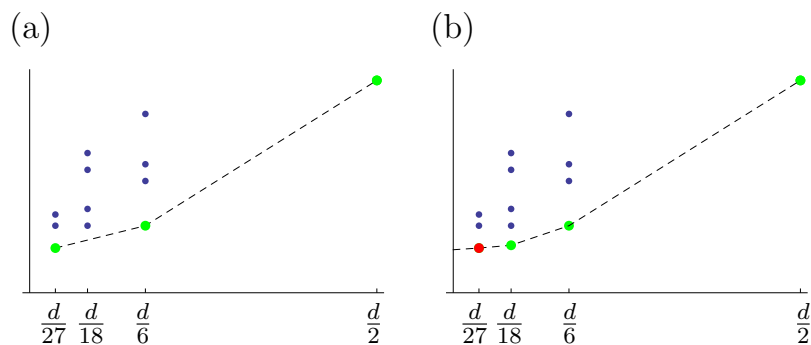


Slika 11: Interval koji nije potencijalno optimalan

Izborom različitih vrijednosti konstante $\hat{L} > 0$ u svakoj iteraciji može postojati više potencijalno optimalnih intervala.

Jedan razlog za traženje manje vrijednosti za konstantu L leži u činjenici što se za veliku vrijednost Lipschitzove konstante, postupak može nepotrebno produljiti.

Ipak, pravi razlog za uvođenje konstante (stope rasta) $\hat{L}$ leži u činjenici da je često puta teško odrediti Lipschitzovu konstantu (Wood and Zang, 1996). Umjesto toga, za razne vrijednosti od $\hat{L}$ trebalo provjeravati uvjete (15)-(16) i tražiti potencijalno optimalne intervale. Na sreću, čak ni to neće biti potrebno. Kao što se vidi na Slici 10, a još jasnije na Slici 12 potencijalno optimalni intervali odgovaraju točkama T_i na vrhovima donjeg ruba konveksne ljuske $\text{conv}\{T_1, \dots, T_n\}$. Za promatrani primjer u prvoj i drugoj iteraciji (drugi i treći grafikon na Slici 10) dobivamo samo dva potencijalno optimalna intervala (odgovaraju najnižim točkama za apscise $\frac{d}{2}$ i $\frac{d}{6}$), a u petoj iteraciji (šesti grafikon na Slici 10) dobivamo tri potencijalno optimalna intervala (odgovaraju najnižim točkama s apscisama $\frac{d}{6}$, $\frac{d}{18}$ i $\frac{d}{54}$).



Slika 12: Veće zelene točke odgovaraju potencijalno optimalnim intervalima

Potencijalno optimalni intervali na Slici 12a prikazani su zelenim točkicama. Nasuprot tome, interval, koji je na Slici 12b predstavljen crvenom točkicom, zbog uvjeta (16) neće

biti svrstan u potencijalno optimalne zato što se njegova $\mathcal{B}$ -vrijednost gotovo ne razlikuje od trenutno minimalne vrijednosti funkcije f_{min} .

4.2 Traženje potencijalno optimalnih intervala

Geometrijski gledano, točka na Slici 12 predstavljat će potencijalno optimalan interval onda ako se kroz nju može provući pravac s nekim koeficijentom smjera $\hat{L} > 0$, iznad kojeg (ili na kojem) se nalaze sve druge točke. To znači da je njegova $\mathcal{B}$ -vrijednost (tj. njegov odsječak na ordinati) manji od $\mathcal{B}$ -vrijednosti svih drugih točaka. Pri tome, ta $\mathcal{B}$ -vrijednost (tj. taj odsječak) mora se barem za $\epsilon > 0$ razlikovati od trenutno minimalne vrijednosti funkcije f_{min} (uvjet (16)).

U nastavku ćemo operacionalizirati opisani princip. Označimo $I = \{1, \dots, n\}$, $d_i = \frac{b_i - a_i}{2}$. Primijetimo najprije da se za izabrani $j \in I$ (a time i izabrani d_j) na Slici 12 može pojaviti jedna ili više točaka. Sve točke s apscisama d_j reprezentiraju sve intervale širine $2d_j$. Označimo nadalje:

- $I_1 = \{i \in I : d_i < d_j\}$ – indeksi točaka koje se na grafikonu (12) nalaze lijevo od točaka s apscisom d_j (primijetite da će za neki j skup I_1 biti prazan);
- $I_2 = \{i \in I : d_i > d_j\}$ – indeksi točaka koje se na grafikonu (12) nalaze desno od točaka s apscisom d_j (primijetite da će za neki j skup I_2 biti prazan);
- $I_3 = \{i \in I : d_i = d_j\}$ – indeksi točaka s apscisom d_j .

Primjerice, ako je j indeks točke T_j sa Slike 12a koja ima apscisu $\frac{d}{2}$ ($d = \frac{b-a}{3}$), imamo: $|I_1| = 11$, $|I_2| = 0$, $|I_3| = 1$. Ako je j indeks točke T_j koja ima apscisu $\frac{d}{6}$, imamo: $|I_1| = 7$, $|I_2| = 1$, $|I_3| = 4$. Ako je j indeks točke T_j koja ima apscisu $\frac{d}{27}$, imamo: $|I_1| = 0$, $|I_2| = 9$, $|I_3| = 3$. Formula (15) znači da je za dani $\hat{L} > 0$ koeficijent smjera pravca (14), koji prolazi točkom T_j manji ili jednak od koeficijenata smjera pravaca koji prolaze drugim točkama T_i , $i \in I$.

Ako (15) zapišemo kao

$$f(c_j) - f(c_i) \leq \hat{L}(d_j - d_i), \quad \forall i \in I, \quad (17)$$

onda za sve $i \in I_3$ uvjeti (15) jednostavno znače da točka T_j mora biti najniža između svih točaka apscise d_j , tj.

$$f(c_j) \leq f(c_i), \quad \forall i \in I_3. \quad (18)$$

Za $i \in I_1$ ($d_j - d_i > 0$) uvjete (15), odnosno (17) možemo zapisati

$$\frac{f(c_j) - f(c_i)}{d_j - d_i} \leq \max_{s \in I_1} \frac{f(c_j) - f(c_s)}{d_j - d_s} \leq \hat{L}, \quad \forall i \in I_1. \quad (19)$$

Za $i \in I_2$ ($d_j - d_i < 0$) uvjete (15), odnosno (17) možemo zapisati

$$\hat{L} \leq \min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s} \leq \frac{f(c_j) - f(c_i)}{d_j - d_i}, \quad \forall i \in I_2. \quad (20)$$

Zato uvjete (15), odnosno (17) možemo pisati

$$f(c_j) \leq f(c_i), \quad \text{if } i \in I_3 \quad (21)$$

$$\max_{s \in I_1} \frac{f(c_j) - f(c_s)}{d_j - d_s} \leq \hat{L} \leq \min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s}, \quad \text{if } I_1 \neq \emptyset \text{ and } I_2 \neq \emptyset \quad (22)$$

$$\max_{s \in I_1} \frac{f(c_j) - f(c_s)}{d_j - d_s} \leq \hat{L}, \quad \text{if } I_1 \neq \emptyset \text{ and } I_2 = \emptyset \quad (23)$$

$$\hat{L} \leq \min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s}, \quad \text{if } I_1 = \emptyset \text{ and } I_2 \neq \emptyset. \quad (24)$$

Primijetite da je slučaj $I_1 = I_2 = \emptyset$ ne treba posebno promatrati jer je uključen u (21).

Uvjet (23) znači da mora postojati $\hat{L} > 0$, tako da je $\max_{s \in I_1} \frac{f(c_j) - f(c_s)}{d_j - d_s} \leq \hat{L}$, a praktično se može provjeriti tako da ispitamo je li $\max_{s \in I_1} \frac{f(c_j) - f(c_s)}{d_j - d_s} < +\infty$.

Uvjet (24) znači da mora postojati $\hat{L} > 0$, tako da je $\hat{L} \leq \min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s}$, a praktično se može provjeriti tako da ispitamo je li $\min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s} > 0$!

Dakle, na taj način smo uvjete (15), odnosno (17), zapisali bez korištenja konstante $\hat{L}$!!!

Proanalizirajmo još uvjet (16). Ako je $f_{\min} \neq 0$ i $I_2 \neq \emptyset$, iz (16) koristeći (24) dobivamo

$$\epsilon \leq \frac{f_{\min} - f(c_j)}{|f_{\min}|} + \hat{L} \frac{d_j}{|f_{\min}|} \leq \frac{f_{\min} - f(c_j)}{|f_{\min}|} + \frac{d_j}{|f_{\min}|} \min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s},$$

a ako je $f_{\min} \neq 0$ i $I_2 = \emptyset$, iz (16) i $\hat{L} \leq L$ dobivamo

$$\epsilon \leq \frac{f_{\min} - f(c_j)}{|f_{\min}|} + \hat{L} \frac{d_j}{|f_{\min}|} \leq \frac{f_{\min} - f(c_j)}{|f_{\min}|} + \frac{d_j}{|f_{\min}|} L.$$

Ako je $f_{\min} = 0$ i $I_2 \neq \emptyset$, iz (16) koristeći (24) dobivamo

$$f(c_j) \leq \hat{L} d_j \leq d_j \min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s},$$

a ako je $f_{\min} = 0$ i $I_2 = \emptyset$, iz (16) i $\hat{L} \leq L$ dobivamo

$$f(c_j) \leq \hat{L} d_j \leq L d_j.$$

Zato uvjet (16) možemo zapisati

$$\epsilon \leq \frac{f_{\min} - f(c_j)}{|f_{\min}|} + \frac{d_j}{|f_{\min}|} K, \quad \text{if } f_{\min} \neq 0 \quad (25)$$

ili

$$f(c_j) \leq d_j K, \quad \text{if } f_{\min} = 0. \quad (26)$$

gdje je

$$K = \begin{cases} \min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s}, & \text{if } I_2 \neq \emptyset, \\ L, & \text{if } I_2 = \emptyset. \end{cases} \quad (27)$$

Dakle, na taj način smo i uvjete (16) zapisali bez korištenja konstante $\hat{L}$!!!
Na taj način možemo napisati uvjete ekvivalente uvjetima (15)–(16) bez korištenja konstante $\hat{L}$.

Lema 2. (*Gablonsky and Kelley, 2001; Gablonsky, 2001*) – *korigirana*

Neka je $L > 0$ Lipschitzova konstanta funkcije $f: [a, b] \rightarrow \mathbb{R}$, $\epsilon > 0$, $I = \{1, \dots, n\}$, $f_{\min} = \min_{i \in I} f(c_i)$. Interval s oznakom $j \in I$ je potencijalno optimalan u smislu Definicije 2 ako

$$(i) \quad f(c_j) \leq f(c_i), \quad \text{if } i \in I_3 \quad (28)$$

$$(ii) \quad \max_{s \in I_1} \frac{f(c_j) - f(c_s)}{d_j - d_s} \leq \min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s}, \quad \text{if } I_1 \neq \emptyset \text{ and } I_2 \neq \emptyset \quad (29)$$

$$(iii) \quad \max_{s \in I_1} \frac{f(c_j) - f(c_s)}{d_j - d_s} < +\infty, \quad \text{if } I_1 \neq \emptyset \text{ and } I_2 = \emptyset \quad (30)$$

$$(iv) \quad \min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s} > 0, \quad \text{if } I_1 = \emptyset \text{ and } I_2 \neq \emptyset \quad (31)$$

$$(v) \quad \epsilon \leq \frac{f_{\min} - f(c_j)}{|f_{\min}|} + \frac{d_j}{|f_{\min}|} K, \quad \text{if } f_{\min} \neq 0 \quad (32)$$

ili

$$(vi) \quad f(c_j) \leq d_j K, \quad \text{if } f_{\min} = 0, \quad (33)$$

gdje je

$$K = \begin{cases} \min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s}, & \text{if } I_2 \neq \emptyset, \\ L, & \text{if } I_2 = \emptyset. \end{cases} \quad (34)$$

Programsku konstrukciju uvjeta (i)-(vi) isprogramirao sam ovako:

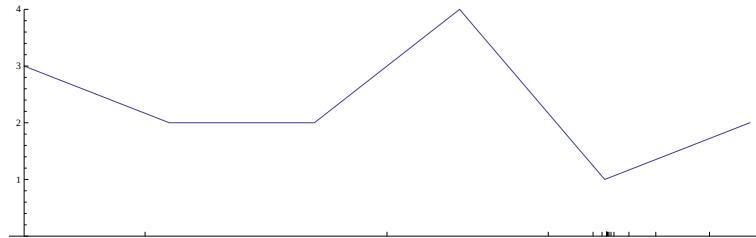
```

false = 0;
Do[If[f[cj] > f[ci], false = 1], {i, I3}];
If[false = 0, Print["(i) je isp." ];
  If[(I1 ≠ ∅ and I2 ≠ ∅ and KL ≤ KD) or (I1 ≠ ∅ and I2 = ∅ and KL < +∞) or (I1 = ∅ and I2 ≠ ∅ and KD > 0),
    Print["(ii-iv) je isp." ];
    If[If[I2 = ∅, K = L, K = KD]; (fmin = 0 and f[cj] ≤ Kdj) or (fmin ≠ 0 and ε ≤  $\frac{f_{min} - f[c_j]}{|f_{min}|} + \frac{d_j}{|f_{min}|} K$ ),
      Print["(v-vi) je isp." ]; S=Append[S,j],
      Print["(v-vi) nije isp." ]
    ], Print["(ii-iv) nije isp." ]
  ], Print["(i) nije isp." ]
], {j, n}]

```

$$KL = \max_{s \in I_1} \frac{f(c_j) - f(c_s)}{d_j - d_s}, \quad KD = \min_{s \in I_2} \frac{f(c_j) - f(c_s)}{d_j - d_s}$$

Primjer 6. Na Slici 13 prikazan je iterativni postupak *DIRECT* s biranjem potencijalno optimalnih intervala u smislu Definicije 2.



Slika 13: Primjena algoritma *DIRECT* s izborom potencijalno optimalnih intervala

Primjedba 5. Možemo promatrati funkciju

$$\Phi(\lambda) = \text{med}_i (y_i - \lambda)^2, \quad (35)$$

koja se pojavljuje kod Least Median of Squares. Vjerojatno bi u ovom slučaju *DIRECT* slabo funkcionirao jer je Lipschitzova konstanta ove funkcije vrlo velika. Mislim da ovo možemo razmatrati kao problem optimizacije funkcije sa skokovima prve vrste: postoje $\lim_{t \rightarrow x_0-} f(t)$ i $\lim_{t \rightarrow x_0+} f(t)$, ali su međusobno različiti. U (Strongin and Vasiljev, 1984) opisan je postupak koji rješava ovakvu situaciju za jednu drugu metodu, ali se spominje da bi se na sličan način moglo pristupiti i metodi slomljenih pravaca (Pijavskij, 1972).

Zadatak 4. Zadani su realni brojevi $y_1, \dots, y_m \in \mathbb{R}$, $m \geq 1$ i funkcija $\Phi: \mathbb{R} \rightarrow \mathbb{R}_+$,

$$\Phi(t) = \operatorname{med}_i (y_i - t)^2. \quad (36)$$

Je li funkcija Φ Lipschitzova? Ako je, pokušajte odrediti Lipschitzovu konstantu L .

Za podatke 90, 93, 86, 92, 95, 83, 75, 40, 88, 80 definirajte funkciju Φ i pronađite njezin globalni minimum.

4.3 Diskretni DIRECT

Neka je $f: [a, b] \rightarrow \mathbb{R}$, Lipschitzova funkcija s konstantom $L > 0$ čije su vrijednosti poznate na diskretnom skupu $U = \{a \leq u_1 < \dots < u_n \leq b\}$. Treba odrediti

$$\operatorname{argmin}_{u \in U} f(u). \quad (37)$$

Neka je $[\alpha, \beta]_U = \{u_i \in U: \alpha \leq u_i \leq \beta\}$ neki “interval” na U .

- Ako $(\alpha, \beta)_U = \{u_i \in U: \alpha < u_i < \beta\}$ ima manje od k_1 članova, izračunat ćemo $\min_{u_i \in (\alpha, \beta)_U} f(u_i)$ i cijeli interval zamijeniti jednom točkom $\operatorname{argmin}_{u_i \in (\alpha, \beta)_U} f(u_i)$;
- Ako $(\alpha, \beta)_U$ ima više od k_1 članova, odredit ćemo

$$c \in \left(\operatorname{Med}_{\alpha < u_i < \beta} u_i \right) \cap U \quad (38)$$

Slično kao kao kod DIRECT metode (Jones et al., 1993), dobivamo

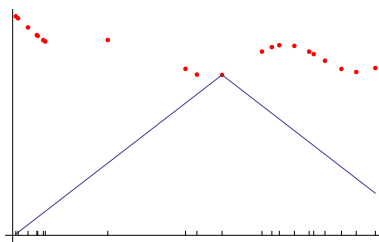
$$f(x) \geq f(c) + L(x - c), \quad x \leq c, \quad (39)$$

$$f(x) \geq f(c) - L(x - c), \quad x \geq c \quad (40)$$

Graf funkcije f nalazi se iznad slomljene linearne funkcije čija je najmanja vrijednost na “intervalu” $[\alpha, \beta]_U$ zadana s (vidi Sliku 14)

$$\mathcal{B}(\alpha, \beta) = \min\{f(c) + L(\alpha - c), f(c) - L(\beta - c)\}. \quad (41)$$

Dalje poduzmemo strategiju kao kod (Pijavskij, 1972) ili (Shubert, 1972) ili (Jones et al., 1993), a dokaz konvergencije imamo kao kod (Vasilev, 1974).

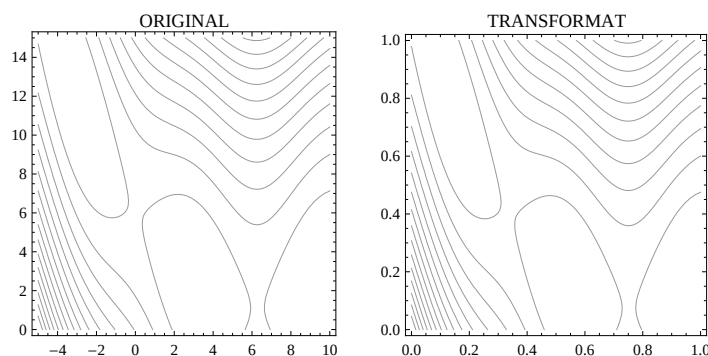
Slika 14: Donja ograda funkcija f na intervalu $[\alpha, \beta]$

5 DIRECT algoritam za globalnu optimizaciju funkcija više varijabli

Primjer 7. Promatramo Branin funkciju, koja se razmatra kao test primjer u (Jones et al., 1993) $g: [-5, 10] \times [0, 15] \rightarrow \mathbb{R}$

$$g(x_1, x_2) = \left(x_2 - \frac{5.1}{4\pi^2} x_1^2 + \frac{5}{\pi} x_1 - 6 \right)^2 + 10 \left(1 - \frac{1}{8\pi} \right) \cos x_1 + 10$$

Ova funkcija postiže globalni minimum $g_{\min} = 0.398$ u tri različite točke. Jedna od njih je $x_1^* = (3.14159, 2.275)$.



Slika 15: ContourPlot of the Branin function

Za optimizaciju algoritmom DIRECT najprije treba odrediti preslikavanje koje pravokutnik $[-5, 10] \times [0, 15]$ prevodi u $[0, 1] \times [0, 1]$. Općenito, preslikavanje koje pravokutnik $[a, b] \times [c, d]$ prslikava u $[0, 1] \times [0, 1]$ zadano je s $T(x) = A(x - u)$, gdje je

$$A = \begin{bmatrix} \frac{1}{b-a} & 0 \\ 0 & \frac{1}{d-c} \end{bmatrix}, \quad u = \begin{bmatrix} a \\ c \end{bmatrix}.$$

Na taj način umjesto optimizacije funkcije $g: [a, b] \times [c, d] \rightarrow \mathbb{R}$ provodimo optimizaciju funkcije $f = g \circ T^{-1}: [0, 1] \times [0, 1] \rightarrow \mathbb{R}$, gdje je $T^{-1}(x) = A^{-1}x + u$.

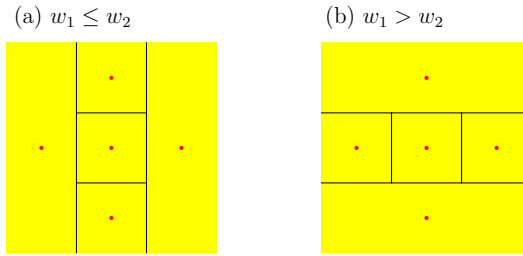
Postupak počinje dijeljenjem kvadrata $[0, 1] \times [0, 1]$ s centrom u točki $c = (\frac{1}{2}, \frac{1}{2})$ na način opisan na Slici 16.a ako je

$$w_1 := \min\{f(c - \delta e_1), f(c + \delta e_1)\} \leq \min\{f(c - \delta e_2), f(c + \delta e_2)\} =: w_2, \quad (42)$$

odnosno na način opisan na Slici 16.b ako je

$$w_1 > w_2, \quad (43)$$

Tako dobivamo skup od $n = 5$ pravokutnika $P = \{P_1, \dots, P_n\}$ s centrima



Slika 16: Dijeljenje kvadrata $[0, 1] \times [0, 1]$

$$c_1 = \left(\frac{1}{2}, \frac{1}{2}\right), c_2 = \left(\frac{1}{2} - \delta, \frac{1}{2}\right), c_3 = \left(\frac{1}{2} + \delta, \frac{1}{2}\right), c_4 = \left(\frac{1}{2}, \frac{1}{2} - \delta\right), c_5 = \left(\frac{1}{2}, \frac{1}{2} + \delta\right), \quad (44)$$

gdje je $\delta = \frac{1}{3}$. Također, svakom pravokutniku pridružujemo duljine njegovih stranica u slučaju (42) (Slika 16.a)

$$h^{(1)} = (\delta, \delta), h^{(2)} = (\delta, 1), h^{(3)} = (\delta, 1), h^{(4)} = (\delta, \delta), h^{(5)} = (\delta, \delta), \quad (45)$$

odnosno duljine njegovih stranica u slučaju (43) (Slika 16.b)

$$h^{(1)} = (\delta, \delta), h^{(2)} = (\delta, \delta), h^{(3)} = (\delta, \delta), h^{(4)} = (1, \delta), h^{(5)} = (1, \delta). \quad (46)$$

Kao mjeru veličine pravokutnika uzimamo

$$d_i = \max\{h_1^{(i)}, h_2^{(i)}\}, \quad i = 1, \dots, n. \quad (47)$$

Tako u slučaju (42) (Slika 16.a) imamo

$$d_1 = \delta, \quad d_2 = 1, \quad d_3 = 1, \quad d_4 = \delta, \quad d_5 = \delta,$$

a u slučaju (43) (Slika 16.b) imamo

$$d_1 = \delta, \quad d_2 = 1, \quad d_3 = \delta, \quad d_4 = 1, \quad d_5 = 1.$$

Primjedba 6. Primijetite da je svaki pravokutnik P_i određen svojim centrom c_i i mjerom njegove veličine d_i . Također važno je primijetiti da smo za mjeru veličine pravokutnika mogli uzeti i nešto drugo (primjerice, poluduljinu maksimalne stranice ili udaljenost centra do vrha) – to nam dozvoljavaju zapisi uvjeta (ii), (iii) u Lemi 2! To će biti važno kod trokuta.

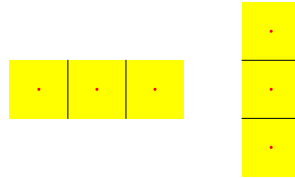
Kao početnu točku globalnog minimuma uzimamo

$$\{\hat{c}, f_{\min}\}, \text{ gdje je } f_{\min} = f(\hat{c}) = \min_{i=1,\dots,n} f(c_i). \quad (48)$$

Nadalje, primjenom Leme 2 pronađemo skup potencijalno optimalnih pravokutnika

$$S = \{P_\mu, \dots\} \subseteq P$$

Skup potencijalno optimalnih pravokutnika S sadrži kvadrate ili paralelepipedu. Ovisno tome provest ćemo njihovo dijeljenje na sljedeći način:



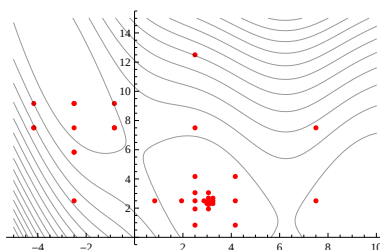
Slika 17: Dijeljenje pravokutnika

- Ako je potencijalno optimalni pravokutnik $P_\mu(c_\mu, d_\mu)$, kvadrat ($h_1^{(\mu)} = h_2^{(\mu)}$), dijelimo ga dalje prema (42) (Slika 16.a) ili prema (43) (Slika 16.b) ovisno o njegovim brojevima w_1, w_2 definiranim kao u (42). Primijetite da na taj način umjesto kvadrata $P_\mu(c_\mu, d_\mu)$ dobivamo manji kvadrat s istim centrom $P'_\mu(c_\mu, \frac{d_\mu}{3})$ i 4 nova pravokutnika. Dakle, ukupni broj pravokutnika postaje $n := n + 4$;
- Ako je potencijalno optimalni pravokutnik $P_\mu(c_\mu, d_\mu)$, pravokutnik ($h_1^{(\mu)} \neq h_2^{(\mu)}$), dijelimo ga dalje na tri jednaka dijela kao na Slici 17. Primijetite da na taj način umjesto pravokutnika $P_\mu(c_\mu, d_\mu)$ dobivamo manji kvadrat s istim centrom $P'_\mu(c_\mu, \frac{d_\mu}{3})$ i 2 nova kvadrata. Dakle, ukupni broj pravokutnika postaje $n := n + 2$;

Primjer 8. Za funkciju iz Primjera 7 Nakon 6 iteracija (41 centar) dobivamo $\hat{x} = (0.541152, 0.154321)$ i $f_{\min} = 0.40$.

Primjedba 7. Primijetite da u slučaju $N = 3$ kocku dijelimo na $2N + 1$ paralelepipedu na sljedeći način (vidi (Jones et al., 1993), Fig. 8, str. 169):

- po prvoj izabranoj koordinati (izbor ovisi o brojevima w_1, w_2, w_3 analogno definiranih kao u (42)) kocku dijelimo na 3 jednaka paralelepipedu (sloja);



Slika 18: The Branin function and DIRECT-approxiamtions

- po drugoj izabranoj koordinati centralni paralelepiped (sloj) dijelimo na 3 jednaka paralelepipeda (sloja);
- po trećoj koordinati centralni paralelepiped (sloj) dijelimo na 3 jednake kocke.

Nadalje uvijek samo kocku dijelimo na na ovaj način. Paralelepipede, kakvi god bili, uvijek dijelimo na tri jednaka dijela.

6 Stronginova metoda za traženje globalnog minimuma

Prema Strongin (1978) navest cemo jednu metodu zasnovanu na informacijsko-statističkim pristupu, koja vodi prema točki globalnog minimuma.

Neka je $f : [a, b] \rightarrow \mathbb{R}$, koja na $[a, b]$ može primiti više lokalnih minimuma.

Algoritam 4. (Strongin)

K_0 : Izabrati $\epsilon > 0$, $u_0 = a$, $u_1 = b$ i izračunati $M = \frac{|f(u_1) - f(u_0)|}{u_1 - u_0}$

K_1 : Izračunati $\mu = \begin{cases} rM, & M > 0 \\ 1, & M = 0 \end{cases}$, gdje je $r > 0$ konstanta (parametar algoritma, primjerice $r = 2$)

K_2 : Izračunati sljedeću točku $u_2 = \frac{1}{2}(u_1 + u_0) - \frac{f(u_1) - f(u_0)}{2\mu}$.

K_3 : Uz pretpostavku da su određene točke $a = u_0 < u_1 < \dots < u_{k-1} < u_k = b$ odrediti veličine

$$M = \max_{1 \leq s \leq k} \frac{|f(u_s) - f(u_{s-1})|}{u_s - u_{s-1}}, \quad \mu = \begin{cases} rM, & M > 0 \\ 1, & M = 0 \end{cases}$$

i definirati funkciju cjelobrojnog argumenta $s = 1, \dots, k$

$$R_k(s) = \mu(u_s - u_{s-1}) + \frac{|f(u_s) - f(u_{s-1})|^2}{u_s - u_{s-1}} - 2(f(u_s) - f(u_{s-1}))$$

Odrediti $s_k \in \{1, \dots, k\}$ iz uvjeta $R_k(s_k) = \max_{1 \leq s \leq k} R_k(s)$.

Ako je $u_{s_k} - u_{s_k-1} \leq \epsilon$, proces završavamo u točki u_{s_k} ;

Ako je $u_{s_k} - u_{s_k-1} > \epsilon$, stavimo

$$u_{k+1} = \frac{1}{2}(u_{s_k} + u_{s_k-1}) - \frac{f(u_{s_k}) - f(u_{s_k-1})}{2\mu}.$$

Iz definicije broja μ vidi se da je $u_{s_k-1} < u_{k+1} < u_{s_k}$.

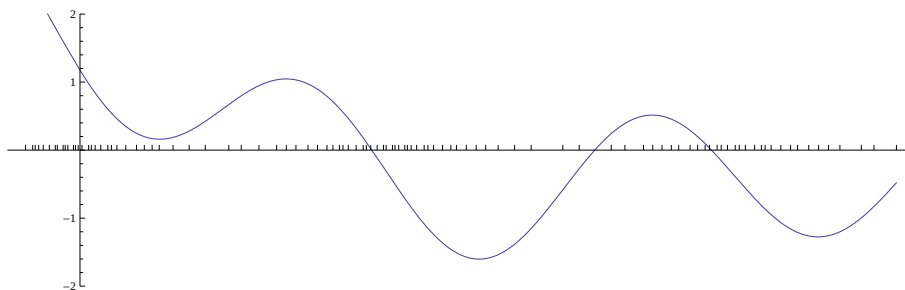
Sada treba prenumerirati niz $u_0, u_1, \dots, u_k, u_{k+1}$ u rastućem poretku i prijeći na definiranje sljedeće točke u_{k+2} na analogan način.

```
In[1]:= f[x_]:= If[x <= 2, -x + 4,
               If[x <= 3, 2,
                 If[x <= 4, 2 x - 4,
                   If[x <= 5, -3 x + 16, x - 4]]]]
a = 1.; b = 6.; ccc = 0; ddd = 4;

In[2]:= x= {a, b}; It=100;
Do[
  x = Sort[x]; k = Length[x];
  M = Max[Table[Abs[(f[x[[i]]]-f[x[[i-1]])]/(x[[i]]-x[[i-1]])], {i,2, k}]];
  m = If[M == 0, 1, r M];
  R = Table[m (x[[i]]-x[[i-1]])+(f[x[[i]]]-f[x[[i-1]])^2/(m(x[[i]]-x[[i-1]]))
            -2(f[x[[i]]]-f[x[[i-1]]])), {i,2,k}];
  t = Position[R, Max[R]][[1]][[1]];
  x = Append[x, (x[[t+1]]+x[[t]])/2 - (f[x[[t+1]]]-f[x[[t]])]/(2 m)];
  , {n,It}]
Print[It, " ", x]
Plot[f[x], {x,a,b}, PlotRange -> {ccc, ddd}, AspectRatio -> .3,
      Ticks -> {Table[{x[[i]], ""}, {i, k+1}], Automatic}, ImageSize -> 700]
```

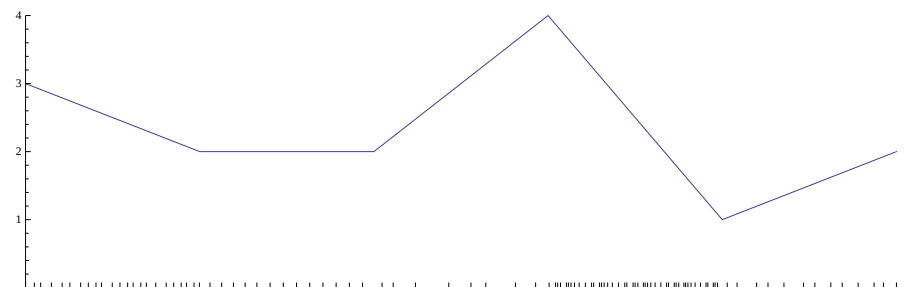
Primjer 9. *Potražimo globalni minimum funkcije*

$$f : [2.7, 7.5) \in \mathbb{R}, \quad f(x) = \sin x + \sin \frac{10}{3}x + \ln x - 0.84x + 3.$$



Slika 19: Niz aproksimacija globalnog minimuma

Primjer 10. *Potražimo globalni minimum funkcije iz Primjera 3.*



Slika 20: Niz aproksimacija globalnog minimuma

Literatura

- Z. Bohte, *Numerično reševanje nelinearnih enačb*, Društvo matematikov, fizikov in astronomov Slovenije, Ljubljana, 1993.
- L. Chiter, *New interpretation of the DIRECT algorithm*, Applied Mathematical Sciences **2**(2008), 529–543
- Ю. М. Данилин, Ц. А. Пиявский, Об одном алгоритме отыскания абсолютного минимума, Тп. семинара “Теория оптимальных решений”, Киев Изд-во АН УССР **2**(1967), 25–37.
- В. Ф. Демьянов, Ф. П. Василев, Недифференцируемая оптимизация, Наука, Москва, 1981.
- J. E. Dennis, Jr, R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, 1996.
- D. E. Finkel, *DIRECT optimization algorithm user guide*, 2003
- D. E. Finkel, C. T. Kelley, *Convergence analysis of the DIRECT algorithm*, Optimization Online (2004)
- D. E. Finkel, C. T. Kelley, *Additive scaling and the DIRECT algorithm*, Journal of Global Optimization **36**(2006), 597–608
- C. A. Floudas, C. E. Gounaris, *A review of recent advances in global optimization*, Journal of Global Optimization **45**(2009), 3–38
- R. Fletcher, *Practical Methods of Optimization*, Wiley, New York, 1987.
- J. M. Gablonsky, C. T. Kelley, *A Locally-Biased form of the DIRECT Algorithm*, Journal of Global Optimization **21**(2001), 27–37
- J. M. Gablonsky, *DIRECT version 2.0 user guide*, 2001
- M. Gaviano, D. Lera, *A global minimization algorithm for Lipschitz functions*, Optimization Letters **2**(2008), 1–13

- E. Gourdin, B. Jaumard, R. Ellaia, *Global Optimization of Holder Functions*, Journal of Global Optimization **8**(1996), 323–348
- J. D. Griffin, T. G. Kolda, *Asynchronous parallel hybrid optimization combining DIRECT and GSS*, Optimization Methods and Software **25**(2010), 797–817
- P. E. Gill, W. Murray and M. H. Wright, *Practical Optimization*, Academic Press, 1981.
- E. Hansen, G. W. Walster, *Global optimization using interval analysis, Second Edition*, Marcel Dekker, 2004
- P. Hansen, B. Jaumard, S.-H. Lu, *On Timonov's algorithm for global optimization of univariate Lipschitz functions*, Journal of Global Optimization **1**(1991), 37–46
- P. Hansen, B. Jaumard, S.-H. Lu, *On the number of iterations of Piyavskii's global optimization algorithm*, Mathematics of Operations Research **16**(1991), 334–350
- P. Hansen, B. Jaumard, S.-H. Lu, *Global optimization of univariate Lipschitz functions: I. Survey and properties*, Mathematical Programming **55**(1992), 251–272
- P. Hansen, B. Jaumard, S.-H. Lu, *Global optimization of univariate Lipschitz functions: II. New algorithms and computational comparison*, Mathematical Programming **55**(1992), 273–292
- D. R. Jones, C. D. Perttunen, B. E. Stuckman, *Lipschitzian optimization without the Lipschitz constant*, JOTA **79**(1993), 157–181
- F. Jarre, J. Stoer *Optimierung*, Springer Verlag, Berlin Heidelberg, 2004.
- D. E. Kvasov, Y. D. Sergeyev, *A univariate global search working with a set of Lipschitz constants for the first derivative*, Optimization Letters **3**(2009), 303–318
- J. M. Ortega, W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
- С. А. Пиявский, Один алгоритм отыскания абсолютного экстремума функции, Ж. вычисл. матем. и матем. физ. **12**(1972), 888–896.
- W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge, 1989.
- H. R. Schwarz, *Numerische Mathematik*, Teubner, Stuttgart, 1986.
- Y. D. Sergeyev, D. Famularo, P. Pugliese, *Index branch-and-bound algorithm for Lipschitz univariate global optimization with multiextremal constraints*, Journal of Global Optimization **21**(2001), 317–341
- B. Shubert, *A sequential method seeking the global maximum of a function*, SIAM Journal on Numerical Analysis, **9**(1972), 379–388
- J. Stoer, *Numerische Mathematik I*, Springer-Verlag, Berlin, 1994.
- J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, New York, 1993.

- Р. Г. Стронгин, Численные методы в многоэкстремальных задачах, Наука, Москва, 1978.
- Р. Г. Стронгин, Ф. П. Василев, —, Ж. вычисл. матем. и матем. физ. , 1984.
- R. J. Vanderbei, *Extension of Piyavskii's algorithm to continuous global optimization*, Journal of Global Optimization **14**(1999), 205–216
- Ф. П. Василев, Лекции по методам решения экстремальных задач, Издательство Московского университета, Москва, 1974.
- G. R. Wood, B. P. Zhang, *Estimation of the Lipschitz constant of a function*, Journal of Global Optimization **8**(1996), 91–103