Intervalna analiza - metoda za globalnu minimizaciju realne funkcije jedne varijable

1 Uvod

Analizirat ćemo jednu numeričku metodu za globalnu minimizaciju realne funkcije f jedne varijable definirane na segmentu $[a, b]$. Metoda je zasnovana na intervalnoj analizi, a u različitim varijantama može se naći u radovima (Hansen, 1979, 1992; Casado at al., 2002). U svrhu primjene ove metode zahtijevat ćemo samo neprekidnost funkcije f na intervalu $[a, b]$. Pokazat ćemo da je metodu moguće značajno ubrzati uz pretpostavku da je funkcija $f \in C^2([a, b])$. U tom slučaju vidjet ćemo kako se poznata Newtonova metoda za traženej stacionarnih točaka funkcije jedne varijable može proširiti na intervale. U sljedećem odjeljku dat ćemo osnove intervalne analize, pri čemu najvećim dijelom koristimo terminologiju iz (Hansen, 1992). Spomenimo još da se ova metoda može prirodno proširiti i na rješavanje problema globalne minimizacije neprekidne realne funkcije više varijabli. Više detalja o toj generalizaciji može se naći u radu (Hansen, 1980).

2 Intervalna aritmetika

Neka su $a, b \in \mathbb{R}$, $a \leq b$. Skup $X = [a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ zovemo interval. Specijalno svaki realni broj x možemo prikazati u obliku tzv. degeneriranog intevala $[x, x]$. Ako je $X = [a, b]$, često koristimo oznake $a = \underline{X}$ te $b = \overline{X}$.

U nastavku definirat ćemo osnovne pojmove intervalne analize.

Definicija 1. *Za interval $X = [a, b]$ kažemo da je*

1) pozitivan, ako je $a > 0$;

2) nenegativan, ako je $a \geq 0$;

3) negativan, ako je $b < 0$;

4) nepozitivan, ako je $b \leq 0$.

Definicija 2. *Za dva intervala $X = [a, b]$ te $Y = [c, d]$ kažemo da su jednaki ako je $a = c$ i $b = d$.*

Definicija 3. *Kažemo da je interval $X = [a, b]$ lijevo od intervala $Y = [c, d]$ i pišemo $X < Y$ ako je $b < c$.*

Promatrajmo skup svih intervala $\mathcal{I} \subseteq \mathbb{R}$. Na skupu $\mathcal{I}$ definiramo operacije zbrajanja, oduzimanja, množenja i dijeljenja na sljedeći način:

Neka je $\diamond$ jedna od operacija $+, -, \cdot, /$ te $X = [a, b], Y = [c, d]$, onda je

$$X \diamond Y = \{x \diamond y : x \in X, y \in Y\}.$$

Može se pokazati da vrijedi

$$X + Y = [a + c, b + d] \tag{1}$$

$$X - Y = [a - d, b - c] \tag{2}$$

$$X \cdot Y = [\min\{ac, ad, bc, bd\}, \max\{ac, ad, bc, bd\}] \tag{3}$$

$$X/Y = [\min\{a/c, a/d, b/c, b/d\}, \max\{a/c, a/d, b/c, b/d\}], \tag{4}$$

pod pretpostavkom da interval $[c, d]$ ne sadrži 0 tj. da je $d < 0$ ili $c > 0$.

Specijalno za $c, d \neq 0$ definiramo

$$1/[c, 0] = [-\infty, 1/c], c < 0 \quad 1/[0, d] = [1/d, +\infty], d > 0,$$

te ako je $c \leq 0 \leq d$ vrijedi

$$1/[c, d] = [-\infty, +\infty].$$

Aritmetika nad neomeđenim intervalima je znatno složenija te izlazi izvan okvira ovih materijala. Više detalja o aritmetici nad neomeđenim intervalima može se pronaći u (Hansen, 1992).

Za prirodni broj n te interval $X = [a, b]$ definiramo $n-$tu potenciju intervala X na sljedeći način:

$$X^n = \begin{cases} [1, 1], & \text{ako je } n = 0; \\ [a^n, b^n], & \text{ako je } a \geq 0 \text{ ili je } n \text{ neparan;} \\ [b^n, a^n], & \text{ako je } b \leq 0 \text{ i } n \text{ je paran;} \\ [0, \max\{a^n, b^n\}], & \text{ako je } a \leq 0 \leq b \text{ i } n \text{ je paran;} \end{cases} \tag{5}$$

Ako je $\alpha > 1$, te $X = [a, b]$, onda definiramo

$$\alpha^X = [\alpha^a, \alpha^b],$$

$$\log_\alpha X = [\log_\alpha a, \log_\alpha b], \quad a > 0$$

Slično za $\alpha < 1$, definiramo

$$\alpha^X = [\alpha^b, \alpha^a],$$

$$\log_\alpha X = [\log_\alpha b, \log_\alpha a], \quad a > 0.$$

Nadalje ako je $X \subseteq \left[-\frac{\pi}{2} + 2k\pi, \frac{\pi}{2} + 2k\pi\right]$, $k \in \mathbb{Z}$, onda definiramo

$$\sin X = [\sin a, \sin b]$$

te ako je $X \subseteq \left[\frac{\pi}{2} + 2k\pi, \frac{3\pi}{2} + 2k\pi\right]$, $k \in \mathbb{Z}$, definiramo

$$\sin X = [\sin b, \sin a],$$

dok inače uzimamo $\sin X = [-1, 1]$.

Ako je $X \subseteq [2k\pi, \pi + 2k\pi]$, $k \in \mathbb{Z}$, onda definiramo

$$\cos X = [\cos b, \cos a]$$

te ako je $X \subseteq [\pi + 2k\pi, 2\pi + 2k\pi]$, $k \in \mathbb{Z}$, definiramo

$$\cos X = [\cos a, \cos b],$$

dok inače uzmimamo $\cos X = [-1, 1]$.

Intervalnu aritmetiku moguće je proširiti i na druge elementarne funkcije. U konkretnim situacijama moguće je koristiti mogućnosti programskog paketa *Mathematica*.

Primjer 1. *Ako je* $X = [a, b]$*, onda je* $X - X = [a - b, b - a]$.

Premda bismo očekivali da je $X - X = [0, 0]$, općenito smo dobili znatno širi interval $[a - b, b - a]$, koji sadrži element 0. Ovakva pojava u intervalnoj aritmetici naziva se problem zavisnosti (Hansen, 1992). Analogna pojava se događa u sljedećem primjeru:

Primjer 2. *Neka je* $X = [-1, 2]$. *Izračunajmo* X^2 *na dva različita način*

a) Preko (5): $[-1, 2]^2 = [0, 4]$.

b) Preko (3): $[-1, 2]^2 = [-1, 2] \cdot [-1, 2] = [-2, 4]$.

Primijetimo da je $[0, 4] \subset [-1, 2]$*, te da smo primjenom (5) dobili upravo skup* $\{x^2 : x \in [-1, 2]\}$.

Ova dva primjera pokazuju da se primjenom intervalne aritmetike mogu dobiti intervali koji su širi od očekivanih te da rezultat ovisi o izboru pravila za računanje. Općenito ako promatramo funkciju $f : \mathcal{D} \to \mathbb{R}$, gdje je $\mathcal{D} \subseteq \mathbb{R}$, takav da je $X \subseteq \mathcal{D}$, onda skup $f(X)$ sadrži sliku funkcije, odnosno skup

$$R(f, X) := \{f(x) : x \in X\} \subseteq f(X)$$

Može se pokazati, vidi (Hansen, 1992) da ako se interval X u izrazu $f(X)$ pojavljuje točno jednom, onda će $f(X)$ biti jednak skupu $R(f, X)$. U svrhu ilustracije navodimo sljedeća dva primjera.

Primjer 3. *Zadana je funkcija* $f : [-2, 2] \to \mathbb{R}$, $f(x) = x^2 - 1$. *Lako se vidi da je slika funkcije* $R(f, [-2, 2]) = [-1, 3]$. *Računamo li* $f(X)$ *primjenom intervalne aritmetike, ali tako da za potenciranje koristimo pravilo (5) dobivamo*

$$\begin{aligned} f(X) &= X^2 - 1 = [-1, 3]^2 - [1, 1] \\ &= [0, 4] - [1, 1] = [-1, 3]. \end{aligned}$$

Uočimo da se u izrazu $X^2 - 1$ *interval* X *pojavljuje točno jednom. Međutim, ako izraz* $x^2 - 1$ *napišemo u obliku* $(x + 1)(x - 1)$, *te potražimo* $(X + 1)(X - 1)$, *dobivamo*

$$(X + 1) \cdot (X - 1) = ([-2, 2] + [1, 1]) \cdot ([-2, 2] - [1, 1]) = [-1, 3] \cdot [-3, 1] = [-9, 3].$$

Uočimo da se u izrazu $(X - 1) \cdot (X + 1)$, *interval* X *pojavljuje dva puta.*

Primjer 4. *Zadana je funkcija* $f : [0, 2] \to \mathbb{R}$, $f(x) = x^2 - 5x + 6$. *Lako se vidi da je slika funkcije* $R(f, [0, 2]) = [0, 6]$. *Računamo li* $f(X)$ *primjenom intervalne aritmetike, ali tako da za potenciranje koristimo pravilo (5) dobivamo*

$$\begin{aligned} f(X) &= X^2 - 5X + 6 = [0, 2]^2 - 5[0, 2] + [6, 6] \\ &= [0, 4] - [0, 10] + [6, 6] = [-4, 10]. \end{aligned}$$

Pretpostavimo da smo prije primjene intervalne aritmetike funkciju f *zapisali na sljedeći način* $f(x) = \left(x - \frac{5}{2}\right)^2 - \frac{1}{4}$ *te onda primijeniti intervalnu aritmetiku. U tom slučaju dobivamo*

$$f(X) = \left([0, 2] - \left[\frac{5}{2}, \frac{5}{2}\right]\right)^2 - \left[\frac{1}{4}, \frac{1}{4}\right] = [0, 6].$$

Uočimo da se u izazu $\left(X - \frac{5}{2}\right)^2 - \frac{1}{4}$, *interval* X *pojavljuje samo jednom, do se u izrazu* $X^2 - 5X + 6$, *interval* X *pojavljuje dva puta.*

Sljedećim primjerom ilustrirat ćemo još jedno svojstvo intervalne aritmetike.

Primjer 5. *Neka su* $A = [-1, 1], B = [2, 3]$, $X = [0, 1]$. *Promatrajmo sljedeća dva izraza*

1) $(A + B)X = ([-1, 1] + [2, 3]) \cdot [0, 1] = [1, 4] \cdot [0, 1] = [0, 4]$

2) $AX + BX = [-1, 1] \cdot [0, 1] + [2, 3] \cdot [0, 1] = [-1, 1] + [0, 3] = [-1, 4]$.

Prethodni primjer pokazuje da je množenje u odnosu na zbrajanje u intervalnoj aritmetici nije ditributivno. Može se pokazati da za intervale X, Y, Z vrijedi

$$X \cdot (Y + Z) \subseteq X \cdot Y + X \cdot Z. \tag{6}$$

Svojstvo (6) zovemo sub-distributivnost. Također se može pokazati da operacije zbrajanja i množenja imaju svojstvo komutativnosti i asocijativnosti.

Spomenimo još neke pojmove koji se koriste intervalnoj aritmetici. Neka je $X = [a, b]$, onda

- $m(X) = \frac{b-a}{2}$, zovemo sredina intervala X,
- $w(X) = b - a$, zovemo širina intervala X,
- $\mu(X) = \max\{|a|, |b|\}$, zovemo magnituda intervala X,
- $\nu(X) = \begin{cases} a, & a > 0 \\ -b, & b < 0 \\ 0, & \text{inače} \end{cases}$

3 Prva metoda za globalnu minimizaciju funkcije

U ovom odjlejku analizirat ćemo metodu za globalnu minimizaciju realne funkcije jedne varijable definirane na intervalu prema (Hansen, 1979), a koja se zasniva na intervalnoj analizi. Neka je $f : [A, B] \to \mathbb{R}$ realna funkcija.

Prisjetimo se sljedećih definicija.

Definicija 4. *Kažemo da funkcija $f : [A, B] \to \mathbb{R}$ u točki x^* postiže*

a) lokalni minimum *ako postoji $\delta > 0$ takav da je*

$$f(x^*) \le f(x), \qquad \text{za svaki } x \in \langle x^* - \delta, x^* + \delta\rangle \cap [A, B].$$

b) globalni minimum *ako vrijedi*

$$f(x^*) \le f(x), \ \text{za svaki } x \in [A, B].$$

U nastavaku ćemo opisati postupak na osnovi kojeg je primjenom intervalne analize moguće iz domene funkcije eliminirati one podintervale u kojima se ne nalaze točke globalnog minimuma.

Pretpostavimo da je $X \subset [A, B]$ te da želimo utvrditi sadrži li X točku globalnog minimuma funkcije f. Neka je x_0 točka iz intervala $[A, B]$, često se uzima

$$x_0 = m([A, B]) = \frac{A + B}{2},$$

koju možemo shvatiti kao početnu aproksimaciju, te $\overline{f} := f(x_0)$ trenutno poznata najmanja vrijednost funkcije. Sve intervale na kojima je vrijednost funkcije veća od $\overline{f}$ možemo eliminirati, jer sigurno ne sadrže točku globalnog minimuma. Formalno koristimo sljedeći test zasnovan na intervalnoj analizi:

Test 1 ako je $\underline{f(X)} > \overline{f}$, onda X ne zadrži točku globalnog minimuma funkcije f, pa X možemo eliminirati iz područja u kojem tražimo točku globalnog minimuma

Primijetimo da u slučaju $\underline{f(X)} \le \overline{f}$, prehodni Test 1 ne govori ništa, osim da X nije moguće eliminirati na ovaj način. U svrhu ilustracije navodimo sljedeći primjer.

Primjer 6. *Zadana je funkcija $f : [-8, 10] \to \mathbb{R}$ formulom $f(x) = x^4 - 4x^2$, koja globalni minimum postiže u točkama $x_{1,2} = \pm\sqrt{2}$. Pretpostavimo da smo izračunali vrijednost funkcije u točki $x_0 = 1$, što iznosi $f(1) = -3$. S obzirom da zasada poznajemo vrijednost funkcije u jednoj točki, trenutno najmanja vrijednost funkcije f iznosi $\overline{f} = -3$. Želimo utvrditi sadrži li interval X točku globalnog minimuma funkcije f. Neka je*

a) $X = X_1 = [3, 4]$. *Računamo*

$$f([3,4]) = [3,4]^4 - 4[3,4]^2 = [81,256] - 4[9,16] = [81,256] - [36,64] = [17,220],$$

te kako je $\underline{f(X_1)} > \overline{f}$, sukladno testu zaključujemo da X sigurno ne sadrži točku globalnog minimuma. Također primijetimo da niti jedan interval $Y > X_1$ ne sadrži točku globalnog minimuma funkcije f.

b) $X = X_2 = [2, 3]$. *Računamo*

$$f([2,3]) = [2,3]^4 - 4[2,3]^2 = [16,81] - 4[4,9] = [16,81] - [16,36] = [-20,65],$$

te sukladno testu ne možemo zaključiti ništa. Međutim zapišemo li funkciju u ekvivalentnom obliku $f(x) = \left(x^2 - 2\right)^2 - 4$, odakle je $f([2,3]) = [0,45]$, te odavde slijedi da niti interval $[2,3]$ ne sadrži točku globalnog minimuma.

Navodimo opći algoritam za globalnu minimizaciju, koja je zasnovan na intervalnoj analizi.

Opći algoritam (OA)

Neka je $\varepsilon > 0$. Postavi $R := \{[A, B]\}$, $Q := \{\}$
`while` $R \neq \{\}$
  odaberi X iz R i obriši ga iz R
    `if` X zadovoljava uvjet za eliminaciju, eliminiraj X
    `else` podijeli X u podintervale X_i, $i = 1, \ldots, n$
      `for` $i = 1, \ldots, n$
        `if` X_i zadovoljava uvjet za eliminaciju, eliminiraj X_i
        `else`
          `if` $\overline{X_i} - \underline{X_i} < \varepsilon$, pohrani X_i u Q
          `else` pohrani X_i u R
`return` Q

3.1 Prva metoda

Navodimo najjednostavniju metodu za eliminaciju intervala koji ne sadrži točku globalnog minimuma funkcije f, a koja se zasnova na Testu 1. Na početku postavimo $\overline{f} := f\left(\frac{A+B}{2}\right)$. Odabrani interval X eliminiramo, ako je $\underline{f(X)} > \overline{f}$. Ako X nije moguće eliminirati, onda predefiniramo $\overline{f}$ na sljedeći način:

$$\overline{f} = \min\{\overline{f}, f(m(X))\}, \quad m(X) = \frac{\underline{X} + \overline{X}}{2}.$$

Primjer 7. *Promatramo funkciju f iz Primjera 6. Vidjeli smo da niti niti interval $[3,4]$, niti bilo koji interval $Y > [3,4]$ ne sadrži točku globalnog minimuma funkcije f. Nadalje, kako je funkcija parna dovoljno je tražiti točku globalnog minimuma funkcije f u intervalu $[0,3]$. Primijenimo Algoritam (OA) na ovaj primjer. U tu svrhu pretpostavimo da je $\varepsilon = 0.1$, te da je broj podintervala na koji ćemo dijeliti interval u koraku 5 Algoritma (OA) jednak $n = 2$. Postavimo $\overline{f} = f(m([3,4])) = f(1.5) = -63/16 = -3.9375$.*

Iteracija 0 $R = \{[0,3]\}$, $Q = \{\}$, $\overline{f} = -3.9375$

Iteracija 1 Odaberemo $X = [0,3]$. Kako je

$$f([0,3]) = [0,3]^4 - 4[0,3]^2 = [0,81] - 4[0,9] = [0,81] - [0,36] = [-36,81],$$

nije moguće eliminirati X te ga dijelimo na $X_1 = [0,1.5]$ te $X_2 = [1.5,3]$. Imamo

$i = 1$ $f(X_1) = f([0,1.5]) = [-9., 5.0625]$, te ga nije moguće eliminirati;
$\overline{f} = \min\{\overline{f}, f(m(X_1))\} = \min\{-3.9375, -1.93359\} = -3.9375$;
$\overline{X_1} - \underline{X_1} > \varepsilon$ te ga nije moguće pohraniti u Q;

$i = 2$ $f(X_2) = f[1.5,3] = [-30.9375, 72]$, te ga nije moguće eliminirati;
$\overline{f} = \min\{\overline{f}, f(m(X_1))\} = \min\{-3.9375, 5.37891\} = -3.9375$;
$\overline{X_2} - \underline{X_2} > \varepsilon$ te ga nije moguće pohraniti u Q;

Prema tome rezultat ove iteracije 1. je

$$R = \{[0,1.5],[1.5,3]\},\ Q = \{\}, \overline{f} = -3.9375$$

Iteracija 2 Odaberimo $X = [0,1.5]$. Kako je $f([0,1.5]) = [-9, 5.0625]$, nije ga moguće eliminirati te ga dijelimo na $X_1 = [0,0.75]$ te $X_2 = [0.75,1.5]$. Imamo

$i = 1$ $f(X_1) = f([0,0.75]) = [-2.25, 0.316406]$, te ga je moguće eliminirati;

$i = 2$ $f(X_2) = f[0.75,1.5] = [-8.68359, 2.8125]$, te ga nije moguće eliminirati;
$\overline{f} = \min\{\overline{f}, f(m(X_2))\} = \min\{-3.9375, -3.46069\} = -3.9375$;
$\overline{X_2} - \underline{X_2} > \varepsilon$ te ga nije moguće pohraniti u Q;

Prema tome rezultat ove iteracije je

$$R = \{[0.75,1.5],[1.5,3]\},\ Q = \{\}, \overline{f} = -3.9375$$

Iteracija 3 Odaberimo $X = [0.75,1.5]$. Kako je $f([0.75,1.5]) = [-8.68359, 2.8125]$, nije ga moguće eliminirati te ga dijelimo na $X_1 = [0.75,1.125]$ te $X_2 = [1.125,1.5]$. Imamo

$i = 1$ $f(X_1) = f([0.75, 1.125]) = [-4.74609, -0.648193]$, te ga nije moguće eliminirati;
$\overline{f} = \min\{\overline{f}, f(m(X_1))\} = \min\{-3.9375, -2.74315\} = -3.9375$;
$\overline{X_1} - \underline{X_1} < \varepsilon$, te ga je moguće pohraniti u Q;

$i = 2$ $f(X_2) = f[(1.125, 1.5)] = [-7.39819, 0]$, te ga nije moguće eliminirati;
$\overline{f} = \min\{\overline{f}, f(m(X_2))\} = \min\{-3.9375, -2.20482\} = -3.9375$;
$\overline{X_2} - \underline{X_2} < \varepsilon$, te ga je moguće pohraniti u Q;

Prema tome rezultat ove iteracije je

$$R = \{[1.5, 3]\},\ Q = \{[0.75, 1.125], [1.125, 1.5]\}, \overline{f} = -3.9375$$

Iteracija 4 Odaberimo $X = [1.5, 3]$. Kako je $f([1.5, 3]) = [-30.9375, 72]$, nije ga moguće eliminirati te ga dijelimo na $X_1 = [1.5, 2.25]$ te $X_2 = [2.25, 3]$. Imamo

$i = 1$ $f(X_1) = f([1.5, 2.25]) = [-15.1875, 16.6289]$, te ga nije moguće eliminirati;
$\overline{f} = \min\{\overline{f}, f(m(X_1))\} = \min\{-3.9375, -1.70288\} = -3.9375$;
$\overline{X_1} - \underline{X_1} > \varepsilon$, te ga nije moguće pohraniti u Q;

$i = 2$ $f(X_2) = f([2.25, 3]) = [-10.3711, 60.75]$, te ga nije moguće eliminirati;
$\overline{f} = \min\{\overline{f}, f(m(X_2))\} = \min\{-3.9375, 19.9182\} = -3.9375$;
$\overline{X_2} - \underline{X_2} < \varepsilon$, te ga nije moguće pohraniti u Q;

Prema tome rezultat ove iteracije je

$$R = \{[1.5, 2.25], [2.25, 3]\},\ Q = \{[0.75, 1.125], [1.125, 1.5]\}, \overline{f} = -3.9375$$

Iteracija 5 Odaberimo $X = [1.5, 2.25]$. Kako je $f([1.5, 2.25]) = [-15.1875, 16.6289]$, nije ga moguće eliminirati te ga dijelimo na $X_1 = [1.5, 1.875]$, i $X_2 = [1.875, 2.25]$. Imamo

$i = 1$ $f(X_1) = f([1.5, 1.875]) = [-9., 3.35962]$, te ga nije moguće eliminirati;
$\overline{f} = \min\{\overline{f}, f(m(X_1))\} = \min\{-3.9375, -3.28148\} = -3.9375$;
$\overline{X_1} - \underline{X_1} < \varepsilon$, te ga je moguće pohraniti u Q;

$i = 2$ $f(X_2) = f([1.875, 2.25]) = [-7.89038, 11.5664]$, te ga nije moguće eliminirati;
$\overline{f} = \min\{\overline{f}, f(m(X_2))\} = \min\{-3.9375, 1.08009\} = -3.9375$;
$\overline{X_2} - \underline{X_2} < \varepsilon$, te ga je moguće pohraniti u Q;

Prema tome rezultat ove iteracije je

$$R = \{[2.25, 3]\},\ Q = \{[0.75, 1.125], [1.125, 1.5], [1.5, 1.875], [1.875, 2.25]\}, \overline{f} = -3.9375$$

Iteracija 6 Odaberimo $X = [2.25, 3]$, Kako je $f([2.25, 3]) = [-10.3711, 60.75]$, nije ga moguće eliminirati te ga dijelimo na $X_1 = [2.25, 2.625]$ te $X_2 = [2.625, 3]$. Imamo

$i = 1$ $f(X_1) = f([2.25, 2.625]) = [-1.93359, 27.2307]$ te ga je moguće eliminirati.

$i = 2$ $f(X_2) = f([2.625, 3]) = [11.4807, 53.4375]$, te ga je moguće eliminirati.

Prema tome rezultat ove iteracije je

$$R = \{\}, \quad Q = \{[0.75, 1.125], [1.125, 1.5], [1.5, 1.875], [1.875, 2.25]\}, \overline{f} = -3.9375.$$

Točka u kojoj funkcija f postiže globalni minimum nalazi se u uniji intervala iz skupa Q, a što je interval $[0.75, 2.25]$.

Zadatak 1. *Primijenite opći algoritam na funkciju iz Primjera 7, ali tako da funkciju zapišete na način da se varijabla x, u zapisu funkcije f pojavljuje samo jednom.*

Zadatak 2. *Primijenite opći algoritam na funkciju iz Primjera 7, ali tako da interval umjesto na dva dijela dijelite na tri dijela, neka je pri tome $\varepsilon = 0.2$.*

Primijetimo da smo ispitivanje mogućnosti izbacivanja nekih intervala provodili dva puta. U svrhu izbjegavanja dvostrukog ispitivanja intervala Opći algoritam može se zapisati na sljedeći način, koji je u praktičnom smislu značajno efikasniji.

Opći algoritam (OA-1)

1. Neka je $\varepsilon > 0$. Postavi $R := \{[A, B]\}$, $Q := \{\}$
2. `while` $R \neq \{\}$
3. odaberi X iz R i obriši ga iz R
4. `if` $\underline{f(X)} > \overline{f}$, eliminiraj X
5. `else` predefiniraj $\overline{f} = \min\{f(m(X), \overline{f})\}$
6. `if` $\overline{X} - \underline{X} < \varepsilon$, spremi X u Q
8. `else` podijeli X u podintervale X_i i spremi ih u R
9. `return` Q

3.2 Druga metoda

Prema (Hansen, 1979), navodimo drugu složeniju i efikasniju metodu za eliminaciju intervala koji ne sadrže točku globalnog minimuma dva puta neprekidno diferencijabilne funkcije f. Osnovna ideja ove metode je izbaciti one podintervale domene funkcije $[A, B]$

i) na kojima je vrijednost funkcije veće od trenuto poznate najmanje vrijednost funkcije (kao u Prvoj metodi);

ii) na kojima je funkcija konkavna;

iii) koji ne sadrže stacionarne točke funkcije f.

3.2.1 Newtonova metoda

Prije same metode prisjetimo se Newtonove metode za traženje lokalnih minimuma dva puta neprekidnos diferencijabilne funkcije. Pretpostavimo da funkcija $f : [a, b] \to \mathbb{R}$ na (a, b) postiže jedinstvenu točku lokalnog minimuma. Neka je $x_0 \in (a, b)$ neka točka. Razvoj funkcije f u Taylorov red u okolini točke x_0 glasi

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \frac{f''(\xi)}{6}(x - x_0)^3,$$

gdje je ξ točka između x_0 i x. Neka je g kvadratna aproksimaciju funkcije f:

$$g(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2,$$

Ako je $f''(x_0) > 0$, onda funkcija g postiže lokalnog minimum u točki x za koju vrijedi $g'(x) = 0$, odakle je $x = x_0 + \frac{f'(x_0)}{f''(x_0)}$. Na taj način definiran je iterativni postupak

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)},\ n = 0, 1, 2$$

koji uz određene uvjete na funkciju f te početnu aproksimaciju x_0 konvergira ka jedinstvenoj točki lokalnog minimuma funkcije f na (a, b).

3.2.2 Druga derivacija i intervalna analiza

Pretpostavit ćemo da je funkcija f dva puta neprekidno diferencijabilna na intervalu X. Razvoj funkcije f u Taylorov red u okolini točke $x_0 \in X$ glasi

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(\xi_1),$$

pri čemu je ξ_1 neki broj između x_0 i x. Ako je $x_0 \in X$ zadan, onda je

$$f(x) \in f(x_0) + (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(X), \tag{7}$$

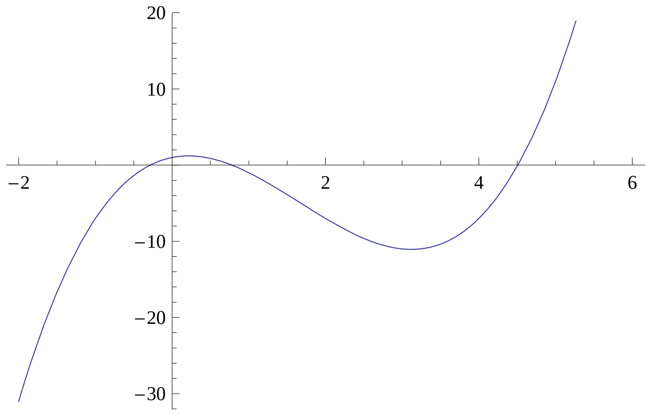
za svaki $x \in X$. Analogno za zadani $x_0 \in X$ vrijedi

$$f'(x) = f'(x_0) + (x - x_0)f''(\xi),$$

gdje je ξ neki element između x_0 i x, što možemo pisati u obliku

$$f'(x) \in f'(x_0) + (x - x_0)f''(X).$$

Prema tome na osnovi poznavanja $f''(X)$ te na osnovi poznavanja $f(x_0)$ i $f'(x_0)$ u jednoj točki $x_0 \in X$ moguće je odrediti intervale koji sadrže $\{f(x) : x \in X\}$ te $\{f'(x) : x \in X\}$.



Slika 1: Graf funkcije f

Primjer 8. *Zadana je funkcija* $f : \mathbb{R} \to \mathbb{R}$, *formulom* $f(x) = x^3 - 5x^2 + 2x + 1$. *Nije teško vidjeti (Slika 1) da je*

$$\{f(x) : x \in [0,5]\} = \left[-\frac{19}{27}(7 + 2\sqrt{19}), 11\right] \approx [-11.0607, 11].$$

Primijenimo li intervalnu aritmetiku na funkciju f *dobivamo* $f(X) = [-124, 136]$. *Istovremeno primijenit ćemo formulu (7). U tu svrhu izračunajmo* $f''(x) = 6y - 10$, *te odaberimo točku* $x_0 \in [0,5]$, *primjerice sredinu* $x_0 = \frac{5}{2}$. *Prema formuli (7) dobivamo*

$$f(x) \in f\left(\frac{5}{2}\right) + ([0,5] - \left[\frac{5}{2}, \frac{5}{2}\right])f'\left(\frac{5}{2}\right) + \frac{1}{2}([0,5] - \left[\frac{5}{2}, \frac{5}{2}\right])^2 f''([0,5]) = [-51.5, 63.5],$$

što je očigledno skup koji je sadržan u $f(X)$ *te sadrži* $\{f(x) : x \in [0,5]\}$.

3.2.3 Sužavanje područja primjenom svojstva konkavnosti

Analogno kao u prethodnom slučaju podijelimo u tu svrhu interval $[A, B]$ na podintervale. Pretpostavimo da je $X = [a, b]$, jedan od takvih podintervala. Primjenom intervalne aritmetike računamo $f''(X) = [\underline{f''(X)}, \overline{f''(X)}]$. Primijetimo

- Ako je $\overline{f''(X)} < 0$, onda je f konkavna na X, onda X ne sadrži točku u kojoj se postiže globalni minimum funkcije f na $[A, B]$.

Primjer 9. *Promatrajmo funkciju* f *iz Primjera 7. Lako se vidi da je* $f''(x) = 12x - 8$, *odnosno* $f([-0.5, 0.5]) = [-14., -2.]$, *odakle slijedi da se točka globalnog minimuma sigurno ne nalazi u* $[-0.5, 0.5]$.

Nakon što smo odbacili intervale na osnovi konkavnosti, za traženje točke globalnog minimuma primijenit ćemo metodu koja je opisana u sljedećem pododjeljku.

3.2.4 Newtonova metoda u intervalnoj aritmetici

Neka je X neki podinterval od $[A, B]$ te neka je $x_0 \in X$. Primjerice možemo uzeti da je $x_0 = m(X)$. Ako je $x \in X$ stacionarna točka od f, onda je $f'(x) = 0$ te je x rješenje jednadžbe

$$f'(x_0) + (x - x_0)f''(\xi) = 0,$$

za neki $\xi \in X$. Prema tome, svaka stacionarna točka $x \in X$ sadržana je u skupu

$$S'' = \{x : f'(x_0) + (x - x_0)f''(\xi') = 0, \xi' \in X\}.$$

Pretpostavimo da je $f''(\xi') \neq 0$ za svaki $\xi' \in X$ te neka je $x \in S''$. Onda je

$$x \in S' = x_0 - \frac{f'(x_0)}{f''(X)},$$

odakle slijedi da skup S' sadrži skup S''. Dakle, ako pretpostavimo da $0 \notin f''(X)$, dobro je definiran iterativni postupak:

$$N(X_n) = x_n - \frac{f'(x_n)}{f''(X_n)} \tag{8}$$

$$X_{n+1} = X_n \cap N(X_n),\ n = 0, 1, 2, \ldots \tag{9}$$

pri čemu je $x_n = m(X^n)$.

Iterativni postupak (9) zovemo Newtonova metoda u intervalnoj aritmetici. Moě se pokazati da ako interval $X_0 := X$, sadrži stacionarnu točku, te ako je $0 \notin f''(X)$, onda iterativni postupak (9) konvergira ka degeneriranom intervalu $[x', x']$ gdje je x' upravo stacionarna točka od f.

Prema tome, vratimo se jednoj iteraciji općeg algoritma. Pretpostavmo da je X_i takav da je $\overline{f''(X_i)} > 0$, inače je odbačen na osnovi konkavnosti. Primijetimo da je nužno $\overline{f''(X_i)} > 0$ te $\underline{f''(X_i)} > 0$, jer bi inače bilo $0 \in f''(X_i)$, što smo zasada isključili. To znači da je funkcija f konkveksna na X_i pa se mogu dogoditi dvije mogućnosti:

1) funkcija f ima lokalni minimum u točki x' na X_i, no to zbog konveksnosti zanči da je taj lokalni minimum jedinstven na X_i te ćemo tu točku naći primjenom iterativnog postupka (9). Ostale dijelove intervala X_i odbacujemo, dok interval $[x', x']$, odnosno točku x' spremamo u Q.

2) funkcija f na X_i ne sadrži točku lokalnog minimuma, onda će iterativni postupak (9) konvergirati prema $\emptyset$, što znači da cijeli X_i ne možemo odbaciti te ga vraćamo u R.

Sljedeći korak je analiza slučaja $0 \notin f''(X)$. Označimo zbog jednaostavnosti u tu svrhu

$$X = [x_L, x_R] \quad f''(X) = [u, v].$$

Ako je $u \neq 0$ definiramo
$$c = x - \frac{f'(x)}{u},$$
te ako je $v \neq 0$, definiramo
$$d = x - \frac{f'(x)}{v}.$$
Pretpostavljamo da je $u \leq 0 \leq v$, jer u suprotnom možemo primijeniti Newtonovu metodu. Ako je $z \in S'' \cap X$, onda se može pokazati da je $z \in S_1 \cup S_2$, gdje su

1) za $f'(x) \geq 0$
$$S_1 = \begin{cases} [x_L, d], & \text{ako je } v > 0 \,\&\, d \geq x_L \\ \emptyset, & \text{ako je } v = 0 \text{ ili } v > 0 \,\&\, d < x_L \end{cases}$$
$$S_2 = \begin{cases} [c, x_R], & \text{ako je } u < 0 \,\&\, c \leq x_R \\ \emptyset, & \text{ako je } u = 0 \text{ ili } u < 0 \,\&\, c > x_R \end{cases}$$

2) za $f'(x) \leq 0$
$$S_1 = \begin{cases} [x_L, c], & \text{ako je } u < 0 \,\&\, c \geq x_L \\ \emptyset, & \text{ako je } u = 0 \text{ ili } u < 0 \,\&\, c < x_L \end{cases}$$
$$S_2 = \begin{cases} [d, x_R], & \text{ako je } v > 0 \,\&\, d \leq x_R \\ \emptyset, & \text{ako je } v = 0 \text{ ili } v > 0 \,\&\, d > x_R \end{cases}$$

Prisjetimo se općeg algoritma, te u ovom slučaju iz intervala X brišemo komplement skupa $S_1 \cup S_2$, dok intervale S_1 i S_2 vraćamo u listu R

3.3 Popravljeni Test 1

Podijelimo interval $[A, B]$ na podintervale. Pretpostavimo da je X, jedan od takvih podintervala. Želimo odbaciti sve one $x \in X$ koji imaju svojstvo da je $f(x) \geq \overline{f}$, pri čemu je $\overline{f}$ najmanja trenutno poznata vrijednost funkcije f. Neka je $x = m(X)$ sredina intervala X, te predefiniramo $\overline{f} = \min\{\overline{f}, f(x)\}$. Razvijemo li funkciju f u Taylorov red u okolini točke x_0 dobivamo
$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(\xi)(x - x_0)^2,$$
gdje je ξ neki broj između x_0 i x. Prema tome
$$f(x) \in f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(X)(x - x_0)^2.$$
Kako je funkcija $z \mapsto f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} z(x - x_0)^2$, rastuća problem se svodi na odbacivanje onih $x \in X$ za koje
$$f'(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^2 \underline{f''(X)} \geq E, \tag{10}$$

gdje je $E = \overline{f} - f(x_0) \leq 0$. Promatrajmo nejednakost (10) zapisanu na sljedeći način

$$\frac{1}{2}(x-x_0)^2\underline{f''(X)} + f'(x)(x-x_0) - E \geq 0. \tag{11}$$

Diskriminanta kvadratna nejednadžbe (11) glasi

$$\Delta = (f'(x_0))^2 + 2E\underline{f''(X)}.$$

Pri tome može nastupiti jedna od sljedeće tri mogućnosti

1) Ako je $\Delta < 0$, onda nejednadnakost (11) vrijedi za svaki $x \in X$, pa cijeli X možemo odbaciti.

2) Ako je $\underline{f''(X)} = 0$, onda nejednakost (11) vrijedi za svaki $x \in X$ za koji je

$$x \geq x_0 + \frac{E}{f'(x_0)}, \text{ ako je } f'(x_0) > 0,$$
$$x \text{ proizvoljan ako je } f'(x_0) = 0,$$
$$x \leq x_0 + \frac{E}{f'(x_0)}, \text{ ako je } f'(x_0) < 0.$$

3) Ako je $\underline{f''(X)} \neq 0$ i $\Delta \geq 0$, onda kvadratna jednadžba $\frac{1}{2}(x-x_0)^2\underline{f''(X)} + f'(x_0)(x-x_0) - E = 0$ po x ima dva realna rješenja

$$r_{1,2} = x_0 + \frac{-f'(x_0) \pm \sqrt{\Delta}}{\underline{f''(X)}}.$$

te pri tome treba odbaciti one $x \in X$ za koje je vrijedi

$$x \in \begin{cases} [r_1, r_2], & \underline{f''(X)} < 0 \\ [\underline{X}, r_1] \cup \left[r_2, \overline{X}\right], & \underline{f''(X)} > 0 \end{cases}$$

Literatura

L. G. Casado, J. A. Martinez, I. Garcia, Ya. D.Sergeyev, *New interval analysis support functions using gradient information in a global minimization algorithm*, Journal of Global Optimization, ,**22**(2002), 359–376

E. R. Hansen, *Global Optimization Using Interval Analysis: The One-Dimensional Case*, JOTA **29**(1979), 331–344

E. R. Hansen, *Global Optimization Using Interval Analysis: The Multi-Dimensional Case*, Numer. Math. **34**(1980), 247–270

E. R. Hansen G. W. Walster, *Global Optimization Using Interval Analysis*, Marcel Dekker, INC., New York-Basel, 1992