Matematički praktikum
Predavanje 4 – 6 *

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

18. studenoga 2020.

Sadržaj

*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)

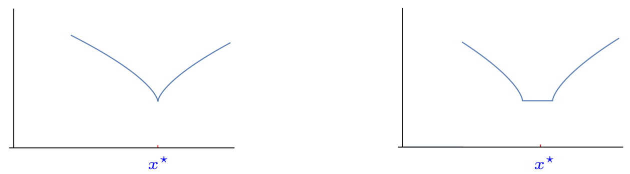
†scitowsk@mathos.hr, ksabo@mathos.hr, dgrahova@mathos.hr, mugrica@mathos.hr

1 Optimizacijski problem

Definicija 1. [13] Kažemo da funkcija $f\colon \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$ u točki $x^\star \in \mathcal{D}$ postiže lokalni minimum ako postoji okolina $\mathcal{O}(x^\star)$ takva da je $f(x) \geq f(x^\star)$ za sve $x \in \mathcal{O}(x^\star) \cap \mathcal{D}$. Točku $x^\star$ zovemo točka lokalnog minimuma ili lokalni minimizator funkcije f.

Kažemo da je $x^\star$ točka strogog lokalnog minimuma ili strogi lokalni minimizator funkcije f ako postoji okolina $\mathcal{O}(x^\star)$ takva da je $f(x) > f(x^\star)$ za sve $x \in \mathcal{O}(x^\star) \cap \mathcal{D} \setminus \{x^\star\}$.

(a) Strogi lokalni minimizator (b) Lokalni minimizator



Slika 1: Lokalni minimizator i strogi lokalni minimizator

Definicija 2. Kažemo da funkcija $f\colon \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$ u točki $x^\star \in D$ postiže globalni minimum na $\mathcal{D}$ ako je

$$f(x^\star) \leq f(x), \quad \forall x \in \mathcal{D}.$$

Točku $x^\star \in \mathcal{D}$ zovemo točka globalnog minimuma ili globalni minimizator funkcije f na $\mathcal{D}$. Skup svih točaka globalnog minimuma funkcije f označavamo s

$$\operatorname*{argmin}_{x \in \mathcal{D}} f(x),$$

vrijednost

$$f(\hat{x}), \quad \text{za} \quad \hat{x} \in \operatorname*{argmin}_{x \in \mathcal{D}} f(x),$$

zovemo globalni minimum funkcije f na $\mathcal{D}$, a funkciju f nazivamo *minimizirajuća funkcija.*

Primjedba 1. Analogno definiciji globalnog minimuma može se definirati pojam *globalnog maksimuma* i *točka globalnog maksimuma* funkcije f, ali kako je

$$(i) \quad \max_{x \in \mathcal{D}} f(x) = -\min_{x \in \mathcal{D}} \big(- f(x) \big),$$
$$(ii) \quad \operatorname*{argmax}_{x \in \mathcal{D}} f(x) = \operatorname*{argmin}_{x \in \mathcal{D}} \big(- f(x) \big),$$

dovoljno je proučavati samo problem globalnog minimuma.

2 Konveksne funkcije

Za funkciju $f\colon \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$, promatrajmo optimizacijski problem definiran u t.**??**, str.**??**. U nastavku navodimo definicije konveksnog skupa i konveksne funkcije te dajemo odgovarajuće karakterizacije (vidi primjerice [4, 6, 8]).

Definicija 3. Kažemo da je skup $\mathcal{D} \subseteq \mathbb{R}^n$ konveksan ako je za sve $x, y \in \mathcal{D}$, $\alpha x + (1-\alpha)y \in \mathcal{D}$ za svaki $\alpha \in [0,1]$. Za funkciju $f\colon \mathcal{D} \to \mathbb{R}$, definiranu na konveksnom skupu $\mathcal{D} \subseteq \mathbb{R}^n$ kažemo da je konveksna ako za sve $x, y \in \mathcal{D}$ vrijedi

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y), \quad \forall \alpha \in [0,1]. \tag{1}$$

Ako u prethodnoj nejednakosti vrijedi znak „$<$" za svaki $\alpha \in (0,1)$ i $x \neq y$, onda kažemo da je funkcija f strogo konveksna.

U sljedećim lemama navedena su neka važna svojstva konveksnih funkcija [1, 4, 6, 7].

Lema 1. *Neka je $f\colon \mathcal{D} \to \mathbb{R}$, konveksna funkcija definirana na konveksnom skupu $\mathcal{D} \subseteq \mathbb{R}^n$. Onda vrijedi:*

(i) nivo-skup (level-set) $\mathcal{D}_\lambda(f) = \{x \in \mathcal{D}\colon f(x) \le \lambda\}$ je konveksan skup za svaki $\lambda \in \mathbb{R}$;

(ii) ako je $\mathcal{D}$ otvoren skup, onda je funkcija f neprekidna na $\mathcal{D}$.

Lema 2. *Neka je $f\colon [a,b] \to \mathbb{R}$ konveksna derivabilna funkcija. Tada:*

(i) $f(u) \ge f(v) + f'(v)(u-v), \quad \forall u, v \in [a,b]$ (gradijentna nejednakost)

(ii) Derivacija f' ne opada na $[a,b]$.

Dokaz. U svrhu dokaza tvrdnje (i) nejednakost (1) za $\alpha \in (0,1)$ zapišimo u obliku

$$\alpha f(u) \ge f(v + \alpha(u-v)) - (1-\alpha) f(v),$$

što nakon dijeljenja s α možemo zapisati kao:

$$f(u) \ge f(v) + \frac{f(v+\alpha(u-v)) - f(v)}{\alpha(u-v)}(u-v).$$

Za $\alpha \to 0+$ dobivamo nejednakost (i).

Dokažimo tvrdnju (ii). U gradijentnoj nejednakosti (i) varijable u i v su ravnopravne. Zamijenimo li ih, iz (i) dobivamo $f(v) \ge f(u) + f'(u)(v-u)$. Zbrajanjem ove nejednakosti s gradijentnom nejednakosti (i) dobivamo

$$\left(f'(u) - f'(v)\right)(u-v) \ge 0,$$

iz čega slijedi tražena tvrdnja (vidi također Zadatak 1). $\square$

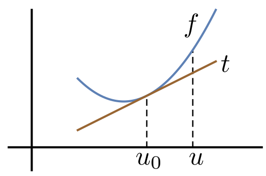
Zadatak 1. *Pokažite da je funkcija $f\colon [a,b] \to \mathbb{R}$ rastuća onda i samo onda ako vrijedi*

$$(x_2 - x_1)(f(x_2) - f(x_1)) \geq 0, \quad \forall x_1, x_2 \in [a,b],\, x_1 < x_2,$$

a da je padajuća onda i samo onda ako vrijedi

$$(x_2 - x_1)(f(x_2) - f(x_1)) \leq 0, \quad \forall x_1, x_2 \in [a,b],\, x_1 < x_2.$$

Geometrijski smisao gradijentne nejednakosti je da njena tangenta $t(u) = f(u_0) + f'(u_0)(u - u_0)$ u proizvoljnoj točki $u_0 \in [a,b]$ leži ispod grafa te funkcije (vidi Sliku 2).



Slika 2: Ilustracija gradijentne nejednakosti

Tvrdnje sljedećih zadataka u sebi sadrže nekoliko kriterija konveksnosti, ovisno o pretpostavkama koje funkcija zadovoljava.

Zadatak 2. *Neka je $f\colon \mathcal{D} \to \mathbb{R}$ konveksna derivabilna funkcija definirana na otvorenom konveksnom skupu $\mathcal{D} \subseteq \mathbb{R}$. Dokažite da ako je $x^\star \in \mathcal{D}$ takav da je $f'(x^\star) = 0$, tada je $x^\star$ globalni minimizator funkcije f.*

Zadatak 3. *Neka je $f\colon \mathcal{D} \to \mathbb{R}$ dva puta derivabilna funkcija definirana na otvorenom konveksnom skupu $\mathcal{D} \subseteq \mathbb{R}$. Dokažite:*

(a) Funkcija f je konveksna ako i samo ako je $f''(x) \geq 0$ za svaki $x \in \mathcal{D}$.

(b) Ako je $f''(x) > 0$ za svaki $x \in \mathcal{D}$, onda je f strogo konveksna.

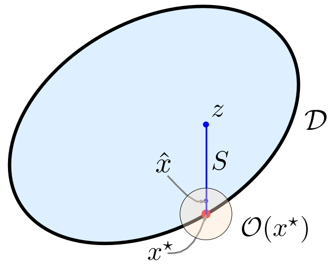
(c) Na primjeru funkcije $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^4$ pokažite da obrat tvrdnje pod (b) ne vrijedi.

Konveksnost minimizirajuće funkcije značajno pojednostavljuje optimizacijski problem zbog činjenice da lokalna optimalnost rješenja garantira globalnu optimalnost o čemu govori sljedeća lema.

Lema 3. *Neka je $f\colon \mathcal{D} \to \mathbb{R}$ konveksna funkcija definirana na konveksnom skupu $\mathcal{D} \subseteq \mathbb{R}^n$. Ako je $x^\star$ lokalni minimizator funkcije f, tada je $x^\star$ i globalni minimizator funkcije f.*

Dokaz. Kako je $x^\star$ lokalni minimizator funkcije f, po Definiciji 1 postoji okolina $\mathcal{O}(x^\star)$ takva da je $f(x) \geq f(x^\star)$ za sve $x \in \mathcal{O}(x^\star) \cap \mathcal{D}$. Pretpostavimo da $x^\star$ nije globalni minimizator funkcije f. Po Definiciji 2 postoji $z \in \mathcal{D}$, takav da je $f(z) < f(x^\star)$. Koristeći ovu nejednakost i svojstvo konveksnosti funkcije f definirane na konveksnom skupu $\mathcal{D}$, za proizvoljni $\alpha \in (0,1]$ dobivamo

$$f(\alpha z + (1-\alpha)x^\star) \leq \alpha f(z) + (1-\alpha)f(x^\star) < \alpha f(x^\star) + (1-\alpha)f(x^\star) = f(x^\star).$$



Slika 3: Globalni minimum konveksne funkcije na konveksnom skupu $\mathcal{D}$

To znači da za svaki x iz spojnice $S = \{\alpha z + (1-\alpha)x^\star \in \mathcal{D} \colon \alpha \in (0,1]\}$ vrijedi $f(x) < f(x^\star)$. Zato postoji $\hat{x} \in S \cap \mathcal{O}(x^\star)$ za koji je također $f(\hat{x}) < f(x^\star)$, što se protivi pretpostavci da je $x^\star$ lokalni minimizator funkcije f. □

Zadatak 4. *Pokažite da je funkcija $F_1 \colon \mathbb{R}^2 \to \mathbb{R}$ definirana u Primjeru* **??** *konveksna.*

Zadatak 5. *Dokažite da je konveksna funkcija $f : [a,b] \to \mathbb{R}$ neprekidna u svim unutrašnjim točkama intervala $[a,b]$ i da ima konačne derivacije slijeva i sdesna, pri čemu vrijedi $f'(x-) \le f'(x+)$, $x \in (a,b)$.*

Uputa: Dokazati da vrijedi

$$\frac{f(u)-f(u-\tau)}{\tau} \le \frac{f(u)-f(u-h)}{h} \le \frac{f(u+h)-f(u)}{h} \le \frac{f(u+\tau)-f(u)}{\tau}, \tag{2}$$

gdje je $0 < h < \tau$; $\quad u, u \pm \tau, u \pm h \in [a,b]$.

Zadatak 6. *Pokažite da za konveksnu funkciju $f : [a,b] \to \mathbb{R}$ vrijedi*

$$\frac{f(u)-f(v)}{u-v} \le \frac{f(w)-f(v)}{w-v} \le \frac{f(w)-f(u)}{w-u}, \quad a \le v < u < w \le b. \tag{3}$$

Koji je geometrijski smisao ovih nejednakosti?

Uputa. Iskoristiti konveksnost funkcije f i zapis $u = \alpha v + (1-\alpha)w$, gdje je $v < u < w$, $\alpha = \frac{w-u}{w-v}$, $\alpha \in (0,1)$.

Zadatak 7. *Dokažite da ako je $f : [a,b] \to \mathbb{R}$ konveksna funkcija, onda vrijedi*

$$f(u) \ge f(v) + \gamma_n(u-v), \qquad \forall u \in [a,b], \forall v \in (a,b), \tag{4}$$

gdje je $\gamma_n \in \mathbb{R}$ proizvoljni broj takav da je $f'(v-) \le \gamma_n \le f'(v+)$.
Uputa. Iskoristiti nejednakost (2) razmotrivši slučajeve $u > v$ i $u < v$ i nejednakost $f'(v-) \le f'(v+)$.

Zadatak 8. *Dokažite da je $x^* \in [a,b]$ točka minimuma konveksne funkcije $f : [a,b] \to \mathbb{R}$ onda i samo onda ako je $f'(x^*+) \ge 0$ i $f'(x^*-) \le 0$. Specijalno, ako $f'(x^*)$ postoji i ako je $x^* \in (a,b)$, onda je $f'(x^*) = 0$.*
Uputa. U (4) stavite $\gamma_n = 0$.

Zadatak 9. *Dokažite da je derivabilna funkcija $f : [a,b] \to \mathbb{R}$ konveksna onda i samo onda ako njezina derivacija f' ne opada na $[a,b]$ (vidi Lemu 2).*

Zadatak 10. *Dokažite da ako je funkcija $f : [a,b] \to \mathbb{R}$ monotono rastuća, a $z : [c,d] \to [a,b]$ konveksna, onda je $f \circ z : [c,d] \to \mathbb{R}$ konveksna.*

2.1 Metoda tangenti za konveksnu derivabilnu funkciju

Za konveksnu derivabilnu funkciju $f : [a, b] \to \mathbb{R}$ može se konstruirati jednostavan algoritam za traženje njenog globalnog minimuma.

Algorithm 1 **(Metoda tangenti)**

Require: f, $u_0 \in [a, b]$

1: Definirati $\Lambda(u; u_0) = f(u_0) + f'(u_0)(u - u_0)$; $\quad P_0(u) := \Lambda(u; u_0)$;

2: Odrediti $u_1 = \operatorname*{argmin}_{u\in[a,b]} P_0(u)$ (Očigledno je $u_1 = a$ ili $u_1 = b$);

3: Provjeriti je li $f'(u_1) \neq 0$ [ako je $f'(u_1) = 0$, onda funkcija f u točki u_1 postiže minimum]

4: Definirati $\Lambda(u; u_1)$ i $P_1(u) = \max_{i=0,1} \Lambda(u; u_i) = \max\{\Lambda(u; u_1), P_0(u)\}$;

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

6: Uz pretpostavku da je funkcija P_{n-1} i čvorovi $u_0, u_1, \ldots, u_{-1}$ $(n \geq 2)$ poznati, definirati $u_n = \operatorname*{argmin}_{u\in[a,b]} P_{n-1}(u)$ i provjeriti je li $f'(u_n) \neq 0$ [ako je $f'(u_n) = 0$, onda funkcija f postiže minimum u točki u_n];

7: Definirati

$$P_n(u) = \max_{i=0,\ldots,n} \Lambda(u; u_i) = \max\left\{\Lambda(u; u_n), P_{n-1}(u)\right\}; \tag{5}$$

8: Odrediti $u_{n+1} = \operatorname*{argmin}_{u\in[a,b]} P_n(u)$

Za fiksni $u_0 \in [a, b]$ definiramo funkciju

$$\Lambda(u; u_0) = f(u_0) + f'(u_0)(u - u_0), \quad u \in [a, b], \tag{6}$$

koju u algoritmu označavamo s $P_0(u)$. Zbog svojstva (i) iz Leme 2 vrijedi $f(u) \geq \Lambda(u; u_0)$, odnosno $f(u) \geq P_0(u)$, za svaki $u \in [a, b]$, tj. P_0 je donja ograda funkcije f.

Nadalje, definiramo sljedeću aproksimaciju $u_1 = \operatorname*{argmin}_{u\in[a,b]} P_0(u)$ koja se očigledno postiže na jednom od rubova intervala $[a, b]$. Nakon toga analogno definiramo $\Lambda(u; u_1) := f(u_1) + f'(u_1)(u - u_1)$ i funkciju

$$P_1(u) = \max\{\Lambda(u; u_1), \Lambda(u; u_0)\} = \max\{\Lambda(u; u_1), P_0(u)\}.$$

Sljedeća aproksimacija dobiva se kao

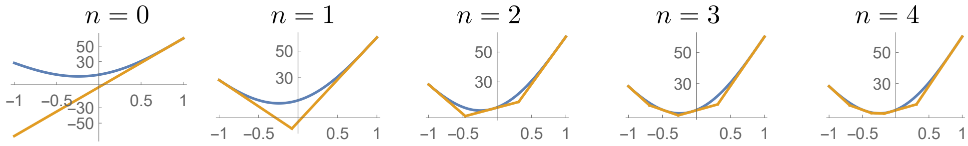
$$u_2 = \operatorname*{argmin}_{u\in[a,b]} P_1(u).$$

Iterativni postupak nastavlja se do unaprijed zadanog broja iteracija ili primjenom nekog drugog kriterija zaustavljanja (vidi primjerice Zadatak **??**).

Navodimo odgovarajući algoritam, čiju implementaciju možemo vidjeti u t.**??**, str.**??**. Funkcija P_n neprekidna je po dijelovima linearna funkcija, čiji je graf sastavljen od dijelova tangenti na graf funkcije f. Zato ovu metodu zovemo Metoda tangenti.

Primjer 1. *Metodu tangenti ilustrirat ćemo na primjeru minimizacije funkcije* $f : [-1, 1] \to \mathbb{R}$, $f(x) = 13 + 18x + 37x^2 - 2x^3 - 6x^4 + \frac{1}{4}x^6$ *koja predstavlja lokalnu LS-udaljenost točke do kubne parabole (vidi Primjer* **??**, *str.***??** *i Primjer* **??**, *str.***??***). Primijetite da je ovako definirana funkcija konveksna i derivabilna.*

Na Slici 1 dana je ilustracija Algoritma 1 primjenom *Mathematica*-modula `MeTang` (vidi t.**??**, str.**??**), a u Tablici 1 može se detaljnije pratiti iterativni postupak. Globalni minimum postiže se u točki $u^\star = -0.243094$.



Slika 4: Ilustracija *Metode tangenti* za funkciju iz Primjera 1

n	1	2	3	4	5
u_n	-1	-0.078	-0.464	-0.266	-0.172
$f(u_n)$	28.25	11.825	12.545	10.838	11.006
$\lvert u^\star - u_n \rvert$	0.757	0.165	0.221	0.023	0.071

Tablica 1: Prvih 5 iteracija *Metode tangenti* na Primjeru 1

Lema 4. *Neka je* $f\colon [a, b] \to \mathbb{R}$ *konveksna derivabilna funkcija,* u_n *niz čvorova i* (P_n) *niz funkcija definiranih kao u Algoritmu 1. Tada vrijedi (usporedite sa svojstvima koja su navedena u Lemi* **??**, *str.***??** *u* Metodi Pijavskog*):*

(i) P_n *je po dijelovima linearna funkcija;*

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

(iii) $P_n(u) \le f(u)$, $\forall u \in [a, b]$ *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. Tvrdnja (i) je očigledna, a (ii) slijedi iz nejednakosti $P_{n+1}(u) = \max\{\Lambda(u; u_{n+1}, P_n(u))\} \ge P_n(u)$.

Tvrdnja (iii) dokazuje se induktivno. Najprije zbog gradijentne nejednakosti zaključujemo da je $P_0(u) = \Lambda(u; u_0) \le f(u)$ i pretpostavimo da je $P_{n-1}(u) \le f(u)$.

Kako također prema gradijentnoj nejednakosti vrijedi $\Lambda(u; u_n) \le f(u)$, imamo

$$P_n(u) = \max\{\Lambda(u; u_n), P_{n-1}(u))\} \le \max\{f(u), f(u)\} = f(u).$$

Za dokaz tvrdnje (iv) najprije primijetimo da iz gradijentne nejednakosti slijedi

$$f(u_i) = \Lambda(u_i; u_i), \quad i = 0, 1, \ldots, n. \tag{7}$$

Nadalje, kako je $P_n(u) = \max\limits_{i=0,1,\ldots,n} \Lambda(u; u_i) \geq \Lambda(u; u_i)$, slijedi

$$P_n(u_i) \geq \Lambda(u_i; u_i), \quad i = 0, 1, \ldots, n. \tag{8}$$

Zato iz (7) slijedi

$$f(u_i) \overset{(7)}{=} \Lambda(u_i; u_i) \overset{(8)}{\leq} P_n(u_i) \overset{(iii)}{\leq} f(u_i), \quad i = 0, 1, \ldots, n,$$

a to znači da u prethodnoj nejednakosti svuda mora biti jednakost. □

Primjedba 2. Niz (u_n) dobiven Algoritmom 1 konvergira prema točki globalnog minimuma funkcije f. Dokaz ove činjenice analogan je dokazu Teorema **??**, str.**??**.

Primjedba 3. Opisana *Metoda tangenti* može se modificirati za traženje minimuma konveksne funkcije (ne nužno derivabilne), koja na intervalu $[a, b]$ postiže svoj minimum. Naime, konveksna funkcija $f : [a, b] \to \mathbb{R}$ neprekidna je na (a, b) i u svakoj točki $x \in (a, b)$ postoje konačni limesi

$$\lim_{h \to 0+} \frac{f(x+h) - f(x)}{h} =: f'(x+), \qquad \lim_{h \to 0+} \frac{f(x) - f(x-h)}{h} =: f'(x-),$$

pri čemu je $f'(x-) \leq f'(x+)$ (vidi Zadatak 5). Zato u prethodno opisanoj *Metodi tangenti* možemo staviti

$$\Lambda(u; u_n) = f(u_n) + \gamma_n (u - u_n), \quad u \in [a, b], \tag{9}$$

gdje je γ_n proizvoljna konstanta koja zadovoljava nejednakost

$$f'(u_n-) \leq \gamma_n \leq f'(u_n+).$$

Primijetite da nejednakost $f(u) \geq \Lambda(u; u_n)$, $u \in [a, b]$ pri tome ostaje sačuvana (vidi Zadatak 7). U primjeni je najjednostavnije γ_n odrediti iz uvjeta

$$|\gamma_n| = \min\{|f'(u_n-)|, |f'(u_n+)|\}.$$

Ako se za neki n pokaže da vrijedi

$$f'(u_n-) \leq 0 \leq f'(u_n+),$$

onda je u_n točka minimuma (vidi Zadatak 8) i proces je završen.

U cilju izbjegavanja mogućih degeneracija $f'(a-) = -\infty$ ili $f'(b+) = +\infty$, dovoljno je iterativni proces započeti točkama $u_0 = a + h_0, u_1 = b - h_1$, gdje h_0 i h_1 treba izabrati tako da računalo prepoznaje $f'(a + h_0+)$ i $f'(b - h_1-)$.

Primjedba 4. Ideja *Metode tangenti* za traženje globalnog minimuma konveksne derivabilne funkcije iskoristit će se u t.**??**, str.**??** za konstrukciju važne metode za traženje globalnog minimuma Lipschitz neprekidne funkcije.

Zadatak 11. *Za funkciju* $f\colon [1,4] \to \mathbb{R}$, $f(x) = \frac{x}{2} + \frac{1}{x}$, *odredite prve tri aproksimacije globalnog minimuma* Metodom tangenti *uz* $u_0 = 1$.

Zadatak 12. *Provjerite može li se za minimizaciju funkcije* $f\colon [-1,1] \to \mathbb{R}$, $f(x) = x^3 + 3x^2$ *primijeniti* Metoda tangenti *te ako može, provedite prve tri iteracije uz početnu aproksimaciju* $u_0 = -1$.

3 Kvazikonveksne funkcije

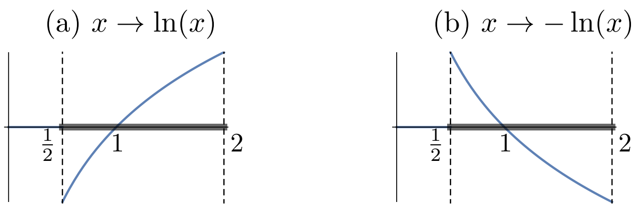
Kao što smo vidjeli iz Leme 3, konveksne funkcije imaju svojstvo izrazito korisno u optimizaciji, a koje glasi da je svaki lokalni minimizator funkcije (ako postoji) ujedno i globalni minimizator te funkcije. U mnogim optimizacijskim problemima pojavljuju se funkcije koje imaju slično svojstvo, a pri tome nisu konveksne. Prirodno je postaviti pitanje može li se oslabjeti zahtjev na konveksnost funkcije, a da se zadrži neki oblik ovog svojstva. S tim ciljem uvodimo sljedeću definiciju.

Definicija 4. Funkcija $f : \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$ je kvazikonveksna na konveksnom skupu $\mathcal{D}$ ako za sve $x, y \in \mathcal{D}$ vrijedi

$$f(\alpha x + (1-\alpha)y) \le \max\{f(x), f(y)\}, \quad \forall \alpha \in [0,1]. \tag{10}$$

Ako u (10) vrijedi stroga nejednakost za $\alpha \in (0,1)$ i $x \neq y$, onda kažemo da je f strogo kvazikonveksna na $\mathcal{D}$ (vidi [13]).

Primjedba 5. Primijetite da je svaka monotona funkcija ujedno i kvazikonveksna. Kao primjer razmotrite monotono rastuću funkciju $x \to \ln(x)$ i monotono padajuću funkciju $x \to -\ln(x)$, obje definirane na segmentu $[0.5, 2]$ (vidi Sliku 5).



Slika 5: Kvazikonveksne funkcije

Sljedeća lema daje jedan praktičan kriterij za ispitivanje kvazikonveksnosti neke funkcije.

Lema 5. *Funkcija* $f : \mathcal{D} \to \mathbb{R}$ *je kvazikonveksna na* $\mathcal{D}$ *onda i samo onda ako je za svaki* $\lambda \in \mathbb{R}$ *nivo-skup* $\mathcal{D}_\lambda(f) = \{x \in \mathcal{D} : f(x) \le \lambda\}$ *konveksan.*

Dokaz. *(Nužnost)* Pretpostavimo da je funkcija $f\colon \mathcal{D} \to \mathbb{R}$ kvazikonveksna i dokažimo da je za proizvoljni $\lambda \in \mathbb{R}$ skup $\mathcal{D}_\lambda(f)$ konveksan.

Neka su $x, y \in \mathcal{D}_\lambda(f)$. To znači da je $f(x) \le \lambda$ i $f(y) \le \lambda$. Kako je f kvazikonveksna, za proizvoljni $\alpha \in [0,1]$ vrijedi

$$f(\alpha\, x + (1-\alpha)y) \le \max\{f(x), f(y)\} \le \lambda,$$

iz čega zaključujemo da je $\alpha\, x + (1-\alpha)y \in \mathcal{D}_\lambda(f)$.

(Dovoljnost) Pretpostavimo da je za proizvoljni $\lambda \in \mathbb{R}$ skup $\mathcal{D}_\lambda(f)$ konveksan i dokažimo da je tada funkcija $f\colon \mathcal{D} \to \mathbb{R}$ kvazikonveksna.

Za proizvoljne $x, y \in \mathcal{D}$ definirajmo $\mu = \max\{f(x), f(y)\}$. Tada su $x, y \in \mathcal{D}_\mu(f)$, a kako je $\mathcal{D}_\mu(f)$ konveksan, onda je $\alpha\, x + (1-\alpha)y \in \mathcal{D}_\mu(f)$, $\forall \alpha \in [0,1]$, iz čega slijedi

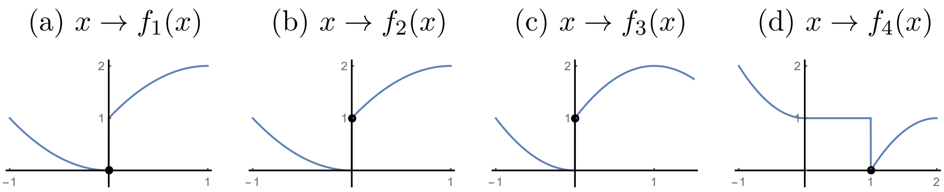
$$f(\alpha\, x + (1-\alpha)y) \le \mu = \max\{f(x), f(y)\}. \qquad \square$$

Primjedba 6. Primijetite da vrijedi [1, 13]:

(i) Svaka konveksna funkcija f je kvazikonveksna, ali obrat ne vrijedi, tj. kvazikonveksna funkcija ne mora biti konveksna. Primjerice, na osnovu Leme 1, nije teško vidjeti da su funkcije $f : \langle 0, \infty\rangle \to \mathbb{R}$ zadane s $f(x) = \ln x$ te $g\colon [-1,1] \to \mathbb{R}$, $g(x) = \sqrt[3]{x^2}$ kvazikonveksne, ali nisu konveksne.

(ii) Ako su funkcije f_i, $i = 1, \dots, m$, kvazikonveksne na $\mathcal{D}$, onda je i funkcija $f(x) = \max\limits_{i=1,\dots,m} f_i(x)$ također kvazikonveksna na $\mathcal{D}$.

(iii) Kvazikonveksna funkcija ne mora biti neprekidna. Primjerice, funkcija $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = \operatorname{sign} x$ je kvazikonveksna na $\mathbb{R}$, ali nije neprekidna. Primijetite također da ova funkcija nije strogo kvazikonveksna.



Slika 6: Koje od ovih funkcija su kvazikonveksne funkcije?

Primjer 2. *Promatramo sljedeće funkcije (vidi Sliku 6)*

$$f_1(x) = \begin{cases} x^2, & x \in [-1,0] \\ 2-(x-1)^2, & x \in (0,1] \end{cases}, \qquad f_2(x) = \begin{cases} x^2, & x \in [-1,0) \\ 2-(x-1)^2, & x \in [0,1] \end{cases},$$

$$f_3(x) = \begin{cases} x^2, & x \in [-1,0] \\ 2-(x-1)^2, & x \in (0,1.5] \end{cases}, \qquad f_4(x) = \begin{cases} x^2+1, & x \in [-1,0] \\ 1, & x \in (0,1) \\ 1-(x-2)^2, & x \in [1,2] \end{cases}.$$

Funkcija f_1 je strogo kvazikonveksna funkcija za koju je $\operatorname*{argmin}_{x\in[-1,1]} f_1(x) = \{0\}$. Funkcija f_2 je strogo kvazikonveksna funkcija za koju vrijedi $\operatorname*{argmin}_{x\in[-1,1]} f_2(x) = \emptyset$. Funkcija f_3 nije kvazikonveksna funkcija, dok je f_4 kvazikonveksna, ali ne i strogo kvazikonveksna funkcija. Nadalje, lokalni minimizator funkcije f_4 svaka je točka iz intervala $[0, 1)$, ali niti jedna od tih točaka nije globalni minimizator. Njezin globalni minimizator je $\operatorname*{argmin}_{x\in[-1,2]} f_4(x) = x^\star = 1$, a globalni minimum iznosi $f_4(x^\star) = 0$.

Primjer funkcije f_4 pokazuje kako se može dogoditi da lokalni minimizator kvazikonveksne funkcije nije ujedno globalni minimizator, odnosno da tvrdnja iz Leme 3 ne vrijedi općenito za kvazikonveksne funkcije. Međutim, analogna tvrdnja može se dokazati uz nešto jaču pretpostavku na skup lokalnih minimizatora funkcije f.

Lema 6. *Neka je $f : \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$ kvazikonveksna funkcija definirana na konveksnom skupu $\mathcal{D}$. Ako je $x^\star$ strogi lokalni minimizator funkcije f, tada je $x^\star$ strogi globalni minimizator funkcije f.*

Dokaz. Ako je $x^\star$ strogi lokalni minimizator funkcije f na skupu $\mathcal{D}$, onda postoji otvoreni skup $\mathcal{O}(x^\star)$ oko točke $x^\star$ takav da je $f(x) > f(x^\star)$, za svaki $x \in \mathcal{O}(x^\star) \cap \mathcal{D}$ i $x \neq x^\star$. Pretpostavimo da $x^\star$ nije strogi globalni minimizator od f na $\mathcal{D}$, odnosno da postoji $z \in \mathcal{D}$, $z \neq x^\star$ i $f(z) < f(x^\star)$. Zbog kvazikonveksnosti funkcije f vrijedi

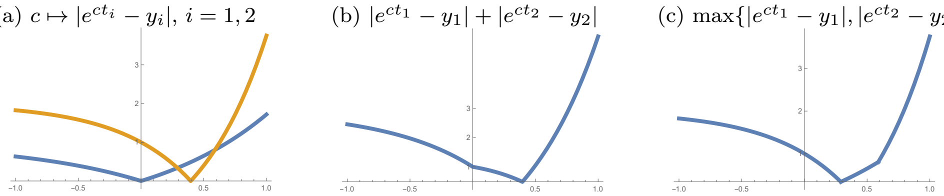
$$f(\alpha z + (1-\alpha)x^\star) \leq \max\{f(z), f(x^\star)\} = f(x^\star), \quad \forall \alpha \in (0, 1].$$

To znači (vidi Sliku 3, str. 5) da za svaki x iz spojnice $S = \{\alpha z + (1-\alpha)x^\star \in \mathcal{D} : \alpha \in (0, 1]\}$ vrijedi $f(x) \leq f(x^\star)$ što se protivi pretpostavci da je $x^\star$ strogi lokalni minimizator. $\square$

Zadatak 13. *Dokažite sljedeću varijantu Leme 6. Neka je $f : \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$ strogo kvazikonveksna funkcija definirana na konveksnom skupu $\mathcal{D}$. Ako je $x^\star$ lokalni minimizator funkcije f, tada je $x^\star$ globalni minimizator funkcije f.*

Zadatak 14. *Je li zbroj dviju kvazikonveksnih funkcija također kvazikonveksna funkcija ? Za podatke (t_i, y_i), $i = 1, \ldots, m$, funkcija $c \mapsto |e^{ct_i} - y_i|$ je kvazikonveksna. Je li i funkcija $c \mapsto \sum_{i=1}^{m} |e^{ct_i} - y_i|$ kvazikonveksna (vidi Sliku 7b)?*

(a) $c \mapsto |e^{ct_i} - y_i|$, $i = 1, 2$ (b) $|e^{ct_1} - y_1| + |e^{ct_2} - y_2|$ (c) $\max\{|e^{ct_1} - y_1|, |e^{ct_2} - y_2|\}$



Slika 7: Zbroj i maksimum dviju kvazikonveksnih funkcija

Zadatak 15. *Je li maksimum dviju kvazikonveksnih funkcija također kvazikonveksna funkcija ? Za podatke* (t_i, y_i), $i = 1, \ldots, m$, *funkcija* $c \mapsto |e^{ct_i} - y_i|$ *je kvazikonveksna. Je li i funkcija* $c \mapsto \max_{i=1}^{m} |e^{ct_i} - y_i|$ *kvazikonveksna (vidi Sliku 7c)?*

Zadatak 16. *Provjerite koje su od ovih funkcija konveksne, a koje kvazikonveksne:*

$$
\begin{array}{ll}
(a) & f_1 : [a,b] \to \mathbb{R},\ f_1(x) = x^2, \\
(b) & f_2 : [a,b] \to \mathbb{R},\ f_2(x) = |x| + x + \operatorname{sign} x, \\
(c) & f_3 : [1,2] \to \mathbb{R},\ f_3(x) = \ln x, \\
(d) & g_1 : [a,b] \to \mathbb{R},\ g_1(x) = \left\{ \begin{array}{cc} |x|, & x \neq 0 \\ -1, & x = 0 \end{array} \right., \\
(e) & g_2 : [a,b] \to \mathbb{R},\ g_2(x) = \left\{ \begin{array}{cc} |x|, & x \neq 0 \\ 1, & x = 0 \end{array} \right., \\
(f) & g_3 : [0,2] \to \mathbb{R},\ g_3(x) = \left\{ \begin{array}{cc} x^2, & x \in [0,1] \\ \sqrt{x}, & x \in [1,2] \end{array} \right. .
\end{array}
$$

Zadatak 17. *Provjerite svojstvo konveksnosti i kvazikonveksnosti sljedećih funkcija:*

$$
\begin{array}{ll}
(a) & f_1 : [a,b] \to \mathbb{R},\ f_1(x) = x^4 + 3x, \\
(b) & f_2 : (0,\infty) \to \mathbb{R},\ f_2(x) = -e^{-x}, \\
(c) & f_3 : [0,3] \to \mathbb{R},\ f_3(x) = \dfrac{-1}{x^2+1}.
\end{array}
$$

Zadatak 18. *Kako treba proširiti funkciju* $f(x) = \frac{1}{e^{1/x^2}+1}$ *u točki* $x_0 = 0$ *tako da ona bude*
(a) strogo kvazikonveksna na $[0,1]$,
(b) strogo kvazikonveksna na $[-1,1]$,
(c) strogo kvazikonveksna na $[-1,0]$*?*

Zadatak 19. *Kako treba proširiti funkciju*

$$
f(x) = \begin{cases} x, & x > 0 \\ -x + A, & x < 0 \end{cases}
$$

u točki $x_0 = 0$ *tako da ona bude*
(a) strogo kvazikonveksna na $[0,1]$,
(b) strogo kvazikonveksna na $[-1,0]$,
(c) strogo kvazikonveksna na $[-1,1]$,
(d) strogo kvazikonveksna na $[a,b]$ *za proizvoljne* $a, b \in \mathbb{R}$ *?*
Razmotrite slučajeve: $A = 0, +1, -1$.

Zadatak 20. *Ako su funkcije* f_1, f_2 *strogo kvazikonveksne na* $[a,b]$, *jesu li takve i funkcije* $f_1 + f_2$, $f_1 - f_2$, $f_1 \cdot f_2$ *te* f_1/f_2 *pod pretpostavkom da su dobro definirane ?*

Zadatak 21. *Neka je* $f : \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}$ *derivabilna funkcija na otvorenom konveksnom skupu* $\mathcal{D}$. *Pokažite da je tada* f *kvazikonveksna na* $\mathcal{D}$ *ako i samo ako za svaki* $x, y \in \mathcal{D}$ *vrijedi*

$$
f(y) \leq f(x) \Rightarrow f'(x)(y - x) \leq 0.
$$

Zadatak 22. *Neka je $f\colon \mathbb{R} \to \mathbb{R}$ dva puta derivabilna funkcija. Dokažite da je f kvazikonveksna ako i samo ako za svaki $x \in \mathbb{R}$*

$$f'(x) = 0 \Rightarrow f''(x) \geq 0.$$

Zadatak 23. *Provjerite je li funkcija $f\colon (0,\infty)^2 \to \mathbb{R},\ f(x_1,x_2) = -x_1x_2$ kvazikonveksna.*

Zadatak 24. *Neka je funkcija $f\colon \mathbb{R}^n \to \mathbb{R}$ definirana s $F(x) = \max\{i\colon x_i \neq 0\}$ i $f(\mathbf{0}) = 0$. Pokažite da je f kvazikonveksna.*

Zadatak 25. *Provjerite jesu li sljedeće funkcije kvazikonveksne.*

a) $f\colon \mathcal{D} \to \mathbb{R}$, $\mathcal{D} = \{x \in \mathbb{R}^n\colon c^Tx + d > 0\}$ *zadana s*

$$f(x) = \frac{a^Tx+b}{c^Tx+d},$$

gdje su $a, b, c, d \in \mathbb{R}^n$.

b) $f\colon \mathcal{D} \to \mathbb{R}$, $\mathcal{D} = \{x \in \mathbb{R}^n\colon \|x-a\|_2 \leq \|x-b\|_2\}$ *zadana s*

$$f(x) = \frac{\|x-a\|_2}{\|x-b\|_2},$$

gdje su $a, b \in \mathbb{R}^n$.

3.1 Unimodalne funkcije

U slučaju funkcija jedne varijable pojam „stroge kvazikonveksnosti” u literaturi se može pronaći i pod nazivom „unimodalne funkcije” (vidi primjerice [7, 13]).

Definicija 5. [7] Kažemo da je funkcija $f : [a,b] \to \mathbb{R}$ unimodalna ako postoji $x^\star \in [a,b]$ takav da je $f(x^\star) = \min\limits_{x\in[a,b]} f(x)$, a za proizvoljne $a \leq x_1 < x_2 \leq b$ vrijedi

$$\begin{aligned} x_2 \leq x^\star &\Longrightarrow f(x_1) > f(x_2),\\ x^\star \leq x_1 &\Longrightarrow f(x_2) > f(x_1). \end{aligned}$$

Primjedba 7. [7]

(i) Uočimo da svojstvo unimodalnosti povlači da je funkcija f strogo padajuća na $[a,x^\star]$ i strogo rastuća na $[x^\star,b]$ dok u točki $x^\star$ postiže globalni minimum.

(ii) Svaka je strogo monotono rastuća (padajuća) funkcija $f : [a,b] \to \mathbb{R}$ unimodalna.

(iii) Strogo konveksna funkcija je unimodalna.

(iv) Unimodalna funkcija ne mora biti neprekidna.

(v) Ako je $f : [a,b] \to \mathbb{R}$ unimodalna funkcija i $[a_1,b_1] \subset [a,b]$, onda je $f\big/_{[a_1,b_1]}$ unimodalna funkcija na $[a_1,b_1]$.

(vi) Ako je $f : [a,b] \to \mathbb{R}$ unimodalna funkcija, $x^\star = \operatorname*{argmin}_{x\in[a,b]} f(x)$ i ako je $a < x_1 < x_2 < b$, onda vrijedi

$$f(x_1) \leq f(x_2) \Longrightarrow x^\star \in [a, x_2],$$
$$f(x_1) > f(x_2) \Longrightarrow x^\star \in [x_1, b].$$

Zadatak 26. *Neka za funkciju $f : [a,b] \to \mathbb{R}$ postoji $x^\star = \operatorname*{argmin}_{x\in[a,b]} f(x)$. Pokažite da je tada funkcija f strogo kvazikonveksna na $[a,b]$ onda i samo onda ako je unimodalna.*

4 Diferencijabilne funkcije

Ako se može pretpostaviti da je promatrana funkcija diferencijabilna, pristup rješavanju optimizacijskog problema može se preciznije postaviti. Svojstvo diferencijabilnosti funkcije ima posebno veliki značaj u karakterizaciji lokalnog minimizatora funkcije te kod konstrukcije iterativne numeričke metode za rješavanje optimizacijskog problema. U nastavku navodimo definiciju i neka važna svojstva diferencijabilnih funkcija.

Definicija 6. [13] Kažemo da funkcija $f : \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$ u točki $x \in \operatorname{int} \mathcal{D}$ diferencijabilna ako postoji $a \in \mathbb{R}^n$ takav da za svaki $h \in \mathbb{R}^n$ vrijedi

$$\lim_{t\to 0} \frac{f(x+th) - f(x) - ta^T h}{t} = 0. \tag{11}$$

Ako za h redom uzmemo j-ti koordinatni vektor e^j, dobivamo

$$\lim_{t\to 0} \frac{f(x+te^j) - f(x) - ta_j}{t} = 0.$$

Dakle, komponente vektora a parcijalne su derivacije $a_j = \frac{\partial f(x)}{\partial x_j} = \partial_j f(x)$.

Ako je funkcija f diferencijabilna u točki $x \in \operatorname{int}\mathcal{D}$, onda ćemo derivaciju funkcije f u točki $x = (x_1, \ldots, x_n)^T \in \operatorname{int}\mathcal{D}$ označavati s

$$f'(x) = \left(\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n}\right) = (\partial_1 f(x), \ldots, \partial_n f(x)),$$

a gradijent funkcije f u točki x s

$$\operatorname{grad} f(x) = \nabla f(x) = f'(x)^T.$$

Nadalje, drugu derivaciju (Hessian) funkcije f u točki x (ako postoji) označavat ćemo s

$$f''(x) = H_f(x) = \nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial^2 x_n} \end{bmatrix} = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right).$$

Prisjetimo se kako skup svih p-puta neprekidno diferencijalbilnih funkcija $f\colon \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$ označavamo s $C^p(\mathcal{D})$.

Ako je $f \in C^p(\mathcal{D})$ te $x + \alpha h \in \mathcal{D}$ za $\alpha \in [0,1]$, onda za nju vrijedi Taylorova formula (vidi primjerice [13, 22]). Specijalno, za $p = 1, 2$ vrijedi

$$f(x+h) - f(x) = (\nabla f(x))^T h + o(\|h\|), \qquad p = 1,$$
$$f(x+h) - f(x) = (\nabla f(x))^T h + \frac{1}{2} h^T \nabla^2 f(x) h + o(\|h\|^2), \qquad p = 2.$$

5 Klasične metode optimizacije

U ovom poglavlju razmatramo problem traženja lokalnog minimuma funkcije $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$. Primjeri navedeni u prethodnom poglavlju ilustriraju da je teško očekivati postojanje jedne optimizacijske metode koja bi jednako uspješno rješavala optimizacijski problem u svim uvjetima. Iz tog razloga, pri izboru optimizacijske metode ključno je iskoristiti sve dostupne informacije o optimizacijskoj funkciji i strukturi njene domene.

Navest ćemo neke klasične optimizacijske metode koje se najčešće sreću u literaturi [3, 4, 6, 7, 13, 17, 18, 22]. Promatrat ćemo samo problem minimizacije jer, kao što smo spomenuli u uvodu (vidi t.1, str.2), problem maksimizacije lako se svodi na problem minimizacije. Kod tih metoda posebno važnu ulogu ima svojstvo konveksnosti i diferencijabilnosti minimizirajuće funkcije.

Definicija 7. Neka je $f\colon \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$, diferencijabilna funkcija. Kažemo da je $x^\star \in \mathcal{D}$ stacionarna točka funkcije f ako je $\nabla f(x^\star) = 0$.

Sljedeći teorem daje vezu između stacionarne točke i lokalnog minimizatora neke funkcije (vidi primjerice [7, 13, 21, 22]).

Teorem 1. *[7] Neka je $\mathcal{D} \subseteq \mathbb{R}^n$ otvoren skup. Ako je $x^\star \in \mathcal{D}$ lokalni minimizator funkcija $f \in C^1(\mathcal{D})$, onda je $x^\star$ stacionarna točka funkcije f. Ako je dodatno $f \in C^2(\mathcal{D})$, onda je Hessian pozitivno semidefinitan (s oznakom $H_f(x^\star) \geq 0$).*

Ako je $x^\star \in \mathcal{D}$ stacionarna točka funkcije $f \in C^2(\mathcal{D})$, a Hessian $H_f(x^\star)$ pozitivno definitan, onda je $x^\star$ strogi lokalni minimumizator funkcije f na $\mathcal{D}$.

5.1 Metode spusta za konveksne funkcije

Minimizacijski problem vrlo često ne možemo riješiti egzaktno te ga u tom slučaju nastojimo riješiti nekom iterativnom numeričkom metodom. U iterativnim metodama obično polazimo od neke aproksimacije x^k za koju želimo odrediti vektor $p^k \neq 0$ te $\alpha_k > 0$ tako da vrijedi

$$f(x^k + \alpha_k p^k) < f(x^k), \quad k = 0, 1, \ldots.$$

Na taj način dobivamo niz aproksimacija

$$x^{k+1} = x^k + \alpha_k p^k, \tag{12}$$

gdje $p^k \in \mathbb{R}^n$ zovemo vektor smjera, a $\alpha_k > 0$ duljina koraka u smjeru p^k. Ako u nekoj iteraciji nismo u mogućnosti pronaći vektor smjera i duljinu koraka s kojim možemo smanjiti vrijednost funkcije f, za očekivati je da je iterativni postupak pronašao stacionarnu točku funkcije f. Ako je uz to funkcija f konveksna, točka koju smo na taj način odredili ujedno je globalni minimizator funkcije f (vidi Lemu 3, str.4).

Primijetite da ako vektor smjera normiramo $\|p^k\| = 1$, onda α_k predstavlja udaljenost točaka x^{k+1} i x^k. Sljedeća lema pokazuje kako treba birati vektor smjera p^k i duljinu koraka α_k kako bismo postigli sniženje vrijednosti minimizirajuće funkcije.

Lema 7. *[13] Neka je $f\colon \mathcal{D} \to \mathbb{R}$ diferencijabilna funkcija definirana na konveksnom skupu $\mathcal{D} \subseteq \mathbb{R}^n$. Ako za neku točku $x \in \operatorname{int}(\mathcal{D})$ i neki $p \in \mathbb{R}^n$ vrijedi $(\nabla f(x))^T p < 0$, tada postoji $\delta > 0$ takav da je*

$$f(x + \alpha p) < f(x), \quad \forall\, \alpha \in (0, \delta).$$

Dokaz. Definiramo pomoćnu funkciju $\varphi : [0, +\infty\rangle \to \mathbb{R}$, $\varphi(\alpha) = f(x + \alpha p)$, za koju je

$$\varphi'(0) = \lim_{\alpha \to 0+} \frac{\varphi(\alpha) - \varphi(0)}{\alpha} = \lim_{\alpha \to 0+} \frac{f(x + \alpha p) - f(x)}{\alpha p} p = (\nabla f(x))^T p.$$

Kako je $(\nabla f(x))^T p < 0$, postoji realan broj $\delta > 0$ takav da je $x + \alpha p \in \mathcal{D}$ i $(f(x + \alpha p) - f(x)) < 0$ za sve $\alpha \in (0, \delta)$. $\square$

Prema Lemi 7, ako je vektor smjera iz točke x^k u točku x^{k+1} dobro definiran ($(\nabla f(x^k))^T p^k < 0$), onda se u iterativnom procesu (12)

$$x^{k+1} = x^k + \alpha_k p^k$$

uvijek može odabrati duljina koraka $\alpha_k > 0$ tako da $f(x^{k+1}) < f(x^k)$. Za to postoje brojne metode (vidi [3, 6, 13, 18]).

Različite strategije izbora vektora smjera p_k vode ka različitim optimizacijskim metodama o kojima ćemo nešto reći u nastavku.

5.1.1 Koordinatna relaksacija

Metoda započinje izborom onog baznog vektora $e^i \in \{e^1, \ldots, e^n\}$ za koji je

$$|(\nabla f(x^k))^T e^i| = \max_{j=1,\ldots,n} |(\nabla f(x^k))^T e^j|.$$

Primijetite da je $(\nabla f(x^k))^T e^i = (f'(x^k))_i$. Sukladno Lemi 7, vektor smjera p^k izabrat ćemo između vektora $\{e^i, -e^i\}$ za koji je

$$(\nabla f(x^k))^T p^k < 0.$$

Umjesto ortonormiranih baznih vektora $\{e^1, \dots, e^n\}$, u metodi možemo koristiti neke druge bazne vektore iz $\mathbb{R}^n$. Metoda ima nisku brzinu konvergencije i koristi se za određivanje početne aproksimacije za neku drugu iterativnu metodu [13].

5.1.2 Gradijentna metoda

Za vektor $p^k := -\nabla f(x^k)$ vrijedi

$$(\nabla f(x^k))^T p^k = -(\nabla f(x^k))^T \nabla f(x^k) = -\|\nabla f(x^k)\|^2 < 0.$$

Zato sukladno Lemi 7, kao vektor smjera ima smisla izabrati ovakav vektor p^k koji određuje gradijentnu metodu. U literaturi (vidi primjerice [5, 6, 13]) ova metoda naziva se još i *Metoda najbržeg spusta*. *Gradijentna metoda* sporo konvergira – ima linearnu brzinu konvergencije [4–6, 13].

5.1.3 Newtonova metoda za minimizaciju bez ograničenja

Za vektor $p^k := -[\nabla^2 f(x^k)]^{-1} \nabla f(x^k)$ vrijedi

$$(\nabla f(x^k))^T p^k = -(\nabla f(x^k))^T [\nabla^2 f(x^k)]^{-1} \nabla f(x^k),$$

što je negativno ako je $\nabla^2 f(x^k)$ pozitivno definitna matrica (u tom slučaju je i $[\nabla^2 f(x^k)]^{-1}$ pozitivno definitna). Zato sukladno Lemi 7, kao vektor smjera ima smisla izabrati ovakav vektor p^k.

U cilju izbjegavanja numerički nestabilnog izračunavanja inverzne matrice $[\nabla^2 f(x^k)]^{-1}$, iterativni postupak uvijek se definira na sljedeći način:

$$\begin{aligned} & x^{k+1} = x^k + p^k, \quad x^0 \in \mathbb{R}^n, \\ & \text{gdje je } p^k \text{ rješenje jednadžbe} \\ & \nabla^2 f(x^k)\, p = -\nabla f(x^k). \end{aligned}$$

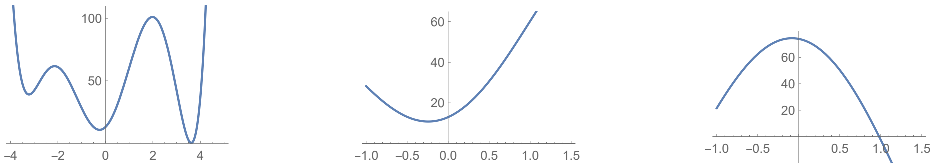
Ovako definirana metoda poznata je kao Newtonova metoda za minimizaciju bez ograničenja [4–6, 13]. Primijetite da se ovaj iterativni postupak potpuno podudara s *Newtonovom metodom* za rješavanje jednadžbe $\nabla f(x) = 0$ (vidi primjerice [17, 18]).

Algorithm 2 **(Newtonova metoda)**

Require: $f\colon \mathbb{R}^n \to \mathbb{R}$;
Require: $x^0 \in \mathbb{R}^n$; $\epsilon > 0$; It, $k = 0$;
{Učitati funkciju, početnu aproksimaciju, točnost i maksimalno dozvoljeni broj iteracija;}
1: **while** $\|\nabla f(x^k)\| > \epsilon$ and $k < It$ **do**
2: Riješi sustav $\nabla^2 f(x^k) p = -\nabla f(x^k)$ i rješenje označi sa p^k;
3: Stavi $x^{k+1} = x^k + p^k$; $k = k + 1$;
4: **end while**
Ensure: $\{k, x^{k+1}\}$

Primjer 3. *Promatramo funkciju $f : \mathbb{R} \to \mathbb{R}$ zadanu s (??) iz Primjera* **??**, *str.*??, *čiji je graf prikazan na Slici 8a. Na Slici 8b prikazana je restrikcija minimizirajuće funkcije f na interval $[-1, 1.5]$, a na Slici 8c vidi se da ta restrikcija nije konveksna. Zato će trebati još suziti interval promatranja.*

(a) Minimizirajuća funkcija f (b) Restrikcija funkcije f (c) f'' na $[-1, 1.5]$



Slika 8: LS udaljenost točke T_0 do parabole q

Newtonovom metodom minimizacije potražit ćemo lokalni minimum ove funkcije na intervalu $[-1, 1.5]$. Uz odabir početne aproksimacije $x_0 = 0.6$, dobivamo niz aproksimacija (vidi Tablicu 2) u kojima funkcija f prima padajući niz vrijednosti. Vrijednost derivacije $f'(x_3) = 0.09$ pokazuje da smo se već u trećoj iteraciji približili točnosti na jednu decimalu, a u četvrtoj iteraciji točnosti na 4 decimale. Vrijednost derivacije funkcije u nekoj aproksimaciji može samo sugerirati da smo bliže ili dalje od tražene točke minimuma, ali ne može to i garantirati. Drugi indikator točnosti aproksimacije je međusobna udaljenost susjednih aproksimacija, ali naravno, ni to nije garancija točnosti aproksimacije. Egzaktni kriteriji točnosti aproksimacije obično su toliko složeni da se u praktičnim primjenama rijetko koriste (vidi primjerice [5]).

k	x_k	$f(x_k)$	$f'(x_k)$
0	.6	35.922	55.173
1	-0.718	18.339	29.641
2	-0.094	11.638	11.022
3	-0.242	10.819	0.093
4	-0.243	10.819	0.00002

Tablica 2: *Newtonova metoda minimizacije*

Uz početnu aproksimaciju $x_0 = .8$, i $f(x_0) = 47.66$ u sljedećoj aproksimaciji dobivamo: $x_1 = -2.078$, i $f(x_1) = 61.564$, dakle vrijednost funkcije ne opada.

Zato u Algoritam 2 uključujemo izračunavanje duljine koraka:

Algorithm 3 **(Izbor duljine koraka)**

Require: $f : \mathbb{R}^n \to \mathbb{R}$; $s \in \mathbb{R}^n$;
Require: $x^k \in \mathbb{R}^n$; $\tau \in \langle 0, \frac{1}{2} \rangle$; $\nu \in \langle 0, 1 \rangle$; $\lambda = 1$;
{Učitati funkciju, k-tu aproksimaciju i parametre τ, ν, λ_{min};}
Riješiti sustav: $\nabla^2 f(x^k) p = -\nabla f(x^k)$ i rješenje označiti sa p^k;
while $f(x^k + \lambda p^k) - f(x^k) > \tau \lambda \left(\nabla f(x^k) \right)^T \cdot p^k$ and $\lambda > \lambda_{min}$, **do**
  $\lambda = \lambda \cdot \nu$
end while
Ensure: λ,

k	x_k	$f(x_k)$	$f'(x_k)$	λ_k
0	.8	47.664	61.564	.5
1	-0.639	16.143	25.630	1
2	-0.160	11.071	6.099	1
3	-0.242	10.819	0.050	-

Tablica 3: *Newtonova metoda s regulacijom koraka*

Nakon toga, u Algoritmu 2 Korak 3: zamijenimo s

$$3: \qquad x^{k+1} = x^k + \lambda\, p^k, \quad k = k+1.$$

Ovako redefinirani Algoritam 2 u literaturi se naziva *Newtonova metoda s regulacijom koraka.*

U tom slučaju iterativni proces uz početnu aproksimaciju $x_0 = .8$ daje niz aproksimacija u kojima funkcija f postiže opadajuće vrijednosti, a već u trećoj iteraciji postiže točnost na jednu decimalu (vidi Tablicu 3).

Primjer 4. *Promatramo funkciju [16]* $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = 2x_1^3 + x_1 x_2^2 + 5x_1^2 + x_2^2$ *[16], čiji su gradijent i Hessian*

$$\nabla f(x_1, x_2) = \begin{bmatrix} 10x_1 + 6x_1^2 + x_2^2 \\ 2x_2 + 2x_1 x2 \end{bmatrix}; \quad \nabla^2 f(x_1, x_2) = \begin{bmatrix} 10 + 12x_1 & 2x_2 \\ 2x_2 & 2 + 2x_1 \end{bmatrix}. \tag{13}$$

Može se pokazati da ova funkcija ima četiri stacionarne točke: $S_1 = (-\frac{5}{3}, 0)$, $S_2 = (-1, -2)$, $S_3 = (-1, 2)$, $S_4 = (0, 0)$, *i da u točki* S_1 *postiže lokalni maksimum, u točki* S_4 *lokalni minimum, a* S_2 *i* S_3 *su sedlaste točke.*

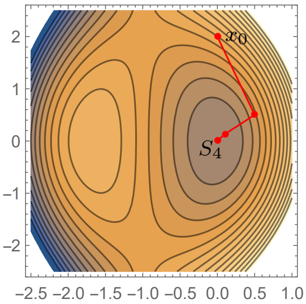
U t.**??**, str.**??** opisani su *Mathematica*-moduli `Newton` i `NewtonKorak` koji implementiraju *Newtonovu metodu* (Algoritam 2) te *Newtonovu metodu uz izbor duljine koraka* (Algoritam 3). Obje varijante Newtonove metode zahtijevaju poznavanje gradijenta i Hessiana minimizirajuće funkcije. Programski sustav *Mathematica* omogućava simboličko računanje derivacija funkcija korištenjem naredbe `D`. Tako primjerice izraze (13) možemo jednostavno dobiti sljedećim naredbama:

```
In[1]:= Clear[x1, x2];
        f[x1_, x2_] := 2*x1^3 + x1*x2^2 + 5*x1^2 + x2^2;
        D[f[x1, x2], {{x1, x2}}]
Out[1]:= {10 x1 + 6 x1^2 + x2^2, 2 x2 + 2 x1 x2}

In[2]:= D[f[x1, x2], {{x1, x2}}, {{x1, x2}}]
Out[2]:= {{10 + 12 x1, 2 x2}, {2 x2, 2 + 2 x1}}
```

Dobivene izraze moguće je direktno iskoristiti za definiranje gradijenta i Hessiana kao funkcija argumenta $(x1, x2)$, što je i napravljeno u modulima opisanim u t.**??**, str.**??**.

Newtonov iterativni postupak (Algoritam 2) uz izbor duljine koraka (Algoritam 3) i početnu aproksimaciju $x_0 = (0, 2)$ prikazan je na Slici 9. Već nakon tri iteracije dobivamo aproksimaciju stacionarne točke S_4, u kojoj funkcija f postiže lokalni (ujedno i globalni) minimum.



Slika 9: Traženje točke minimuma *Newtonovom metodom*

Zadatak 27. Newtonovom metodom *odredite točku maksimuma funkcije iz Primjera 4.*

Zadatak 28. *Odredite prve tri aproksimacije minimuma funkcije* $f(x) = x^3 - 4x + 5$ *na intervalu* $[0, 2]$ *uz početnu aproksimaciju* $x^0 = 2$ *korištenjem* Newtonove metode*.*

Zadatak 29. *Odredite prve tri aproksimacije minimuma funkcije* $f(x) = x^3 - 4x + 5$ *na intervalu* $[0, 2]$ *uz početnu aproksimaciju* $x^0 = 0.5$ *korištenjem* Newtonove metode s izborom duljine koraka *(za parametre uzmite* $\tau = 0.25$ *i* $\nu = 0.5$*).*

Zadatak 30. Newtonovom metodom *odredite prve tri aproksimacije minimuma funkcije* $F(x, y) = x^4 + y^4 + y^2$ *uz početnu aproksimaciju* $(x^0, y^0)^T = (0.5, -0.5)^T$*.*

Zadatak 31. Newtonovom metodom s izborom duljine koraka *odredite prve tri aproksimacije minimuma funkcije* $F(x, y) = e^x + x^4 + y^4$ *uz početnu aproksimaciju* $(x^0, y^0)^T = (0.5, -0.5)^T$ *(za parametre uzmite* $\tau = 0.25$ *i* $\nu = 0.5$*).*

5.1.4 Kvazi-Newtonova metoda

Za vektor $p^k := -B^{-1}\nabla f(x^k)$, gdje je $B > 0$ pozitivno definitna matrica koja „nalikuje” na $\nabla^2 f(x^k)$, vrijedi

$$(\nabla f(x^k))^T p^k = -(\nabla f(x^k))^T B^{-1} \nabla f(x^k) < 0.$$

Zato sukladno Lemi 7, kao vektor smjera možemo izabrati ovakav vektor smjera p^k.

U ovisnosti o izboru matrice B dobivamo poznate kvazi-Newtonove metode: Broyden Method, Davidon–Fletcher–Powell Method (DFP), Broyden–Fletcher–Goldfarb–Shano Method (BFGS)), itd (vidi [6, 17, 18]).

5.2 Nelder-Meadova metoda

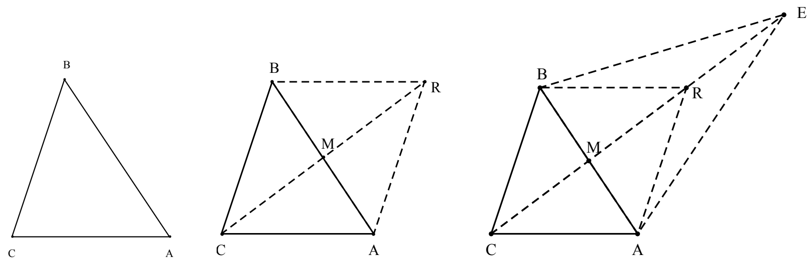
Analizirat ćemo također jednu lokalnu metodu za traženje minimuma realne funkcije $f : \mathbb{R}^n \to \mathbb{R}$ koja je u literaturi poznata kao *Nelder-Meadova metoda*. Ime je dobila po po svojim autorima R. Nelderu i J. A. Meadu koji su metodu objavili 1965. godine u radu [10]. Metoda je jednostavna i u mnogim slučajevima vrlo efikasna, posebno u malim dimenzijama. *Nelder-Meadovu metodu* moguće je primijeniti na funkciju za koju je dovoljno poznavati isključivo vrijednosti u svim točkama domene. Nažalost, u velikim dimenzijama metoda se ne pokazuje efikasnom (vidi primjerice [9, 14, 19]), osim toga, općenito ne postoji dokaz konvergencije ove metode. U radu [9] dani su uvjeti konvergencije za specijalni tip funkcije u jednoj i dvije dimenzije. Iz metodičkih razloga posebno ćemo analizirati samo metodu za funkcije dvije varijable.

5.2.1 Nelder-Meadova metoda za funkcije dviju varijabli

Pretpostavimo da je zadana funkcija $f\colon \mathbb{R}^2 \to \mathbb{R}$. Promatrajmo trokut $\triangle ABC \subset \mathbb{R}^2$ s vrhovima $A = (x_1, y_1)$, $B = (x_2, y_2)$ i $C = (x_3, y_3)$. Bez smanjenja općenitosti možemo pretpostaviti da je

$$f(A) \leq f(B) \leq f(C).$$

Pri tome ćemo točku A zvati najbolji vrh, točku C najlošiji vrh, dok ćemo točku B zvati drugi najlošiji vrh.



Slika 10: Trokut $\triangle ABC$ (lijevo) i njegova refleksija (sredina) i ekspanzija (desno)

Osnovna ideja jedne iteracije *Nelder-Meadove metode* sastoji se u tome da trokut $\triangle ABC$ zamijenimo novim trokutom u čijim vrhovima funkcija f postiže manju vrijednost od one koja se postiže u vrhovima trokuta $\triangle ABC$. U tu svrhu centroid vrhova A i B označimo s $M = \frac{1}{2}(A + B)$. Potom najlošiji vrh C preslikajmo centralno-simetrično u odnosu na točku M. Na taj način dobivamo vrh

$$R = M + (M - C) = 2M - C = A + B - C.$$

Transformaciju koja trokut $\triangle ABC$ preslikava u trokut $\triangle ABR$ zvat ćemo refleksija trokuta (vidi Sliku 10). Pri tome nas zanima je li $f(R) < f(A)$, odnosno je li vrh R bolji od najboljeg vrha A?

- Ako je odgovor pozitivan, naslućujemo da se krećemo u pravom smjeru te se u istom smjeru pomaknemo do točke

  $$E = R + (R - M) = 2R - M = 2A + 2B - 2C - \tfrac{1}{2}A - \tfrac{1}{2}B = \tfrac{3}{2}A + \tfrac{3}{2}B - 2C.$$

  Transformaciju koja trokut $\triangle ABC$ preslikava u trokut $\triangle ABE$ zovemo ekspanzija trokuta (vidi Sliku 10). Točku C zamijenit ćemo točkom E ako je $f(E) < f(R)$, odnosno točkom R ako je $f(E) \geq f(R)$.

- Ako je odgovor negativan, odnosno ako je $f(R) \geq f(A)$, točku R uspoređujemo s drugom najlošijom točkom B, odnosno postavljamo sljedeće pitanje: je li $f(R) < f(B)$?

  (1) Ako je odgovor pozitivan, točku C zamjenjujemo točkom R.

  (2) Ako je odgovor negativan, postavljamo novo pitanje: je li $f(R) < f(C)$?

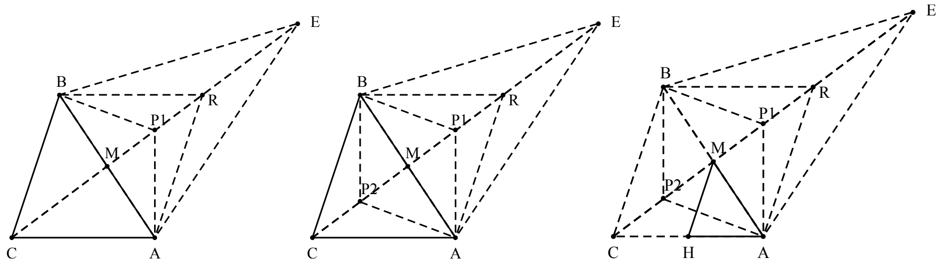
  (2a) Ako je odgovor pozitivan, definiramo novu točku (na Slici 11 označenu s P_1)

$$P = M + \tfrac{1}{2}(R - M) = \tfrac{1}{2}M + \tfrac{1}{2}R = \tfrac{3}{4}A + \tfrac{3}{4}B - \tfrac{1}{2}C.$$

  (2b) Ako je odgovor negativan, definiramo novu točku (na Slici 11 označenu s P_2)

$$P = M - \tfrac{1}{2}(M - C) = \tfrac{1}{2}M + \tfrac{1}{2}C = \tfrac{1}{4}A + \tfrac{1}{4}B + \tfrac{1}{2}C.$$

  Transformaciju koja trokut $\triangle ABC$ preslikava u trokut $\triangle ABP$ zovemo **kontrakcija trokuta** (vidi Sliku 11).



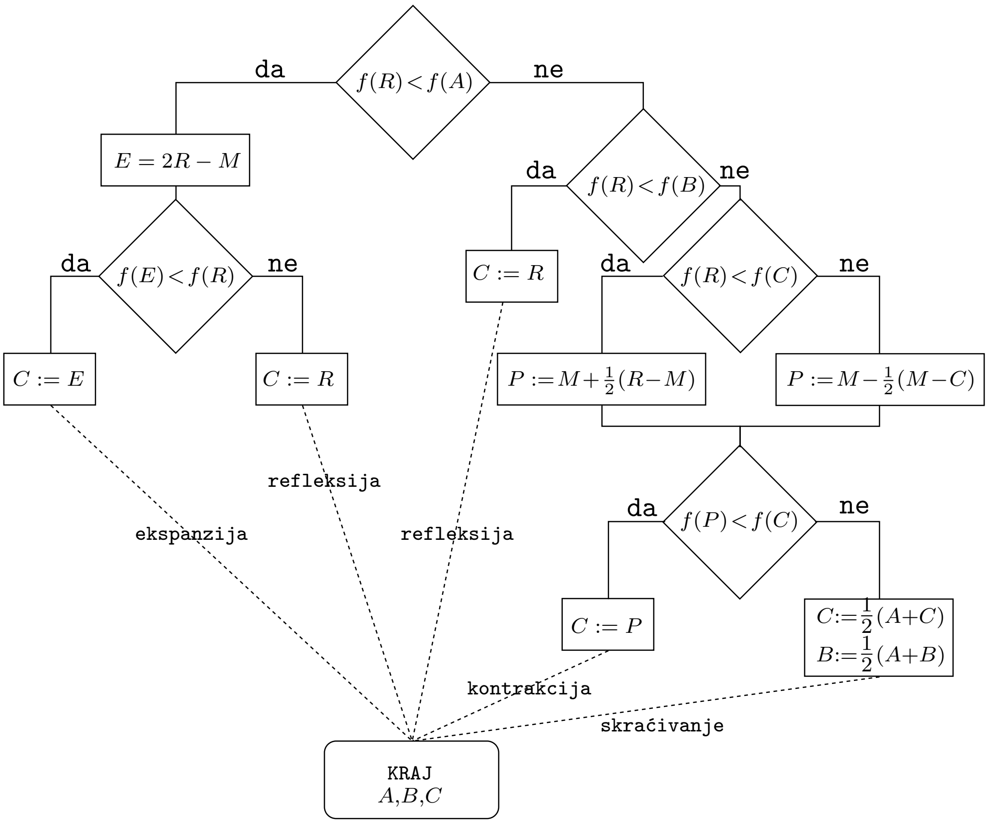
Slika 11: Kontrakcija trokuta (lijevo i u sredini) i njegovo skraćivanje (desno)

  Konačno, pitamo se je li $f(P) < f(C)$.

  (2c) Ako je odgovor pozitivan, točku C zamjenjujemo točkom P.

  (2d) Ako je odgovor negativan, definiramo $H = \frac{1}{2}(A + C)$ te točku B zamijenimo s M, a točku C zamijenimo s H.

  Transformaciju koja trokut $\triangle ABC$ preslikava u trokut $\triangle AMH$ zovemo **skraćivanje trokuta** (vidi Sliku 11).



Slika 12: Dijagram toka za *Nelder-Meadovu metodu* u $\mathbb{R}^2$

Iteracije *Nelder-Meadovog algoritma* ponavljaju se dok se ne zadovolji neki kriterij zaustavljanja. Jedna mogućnost kriterija zaustavljanja može biti udaljenost najbolje točke do težišta trokuta $\triangle ABC$, odnosno algoritam se ponavlja dok god je $d\Big(A, \frac{1}{3}(A+B+C)\Big) \geq \varepsilon$, gdje su $d:\mathbb{R}^2 \to [0,\infty\rangle$ neka funkcija udaljenosti te $\varepsilon > 0$ unaprijed zadana točnost.

U t.**??**, str.**??** opisana je uporaba *Mathematica*-modula `NelderMead` koji implementira *Nelder-Meadovu metodu* za funkcije dviju varijabli na osnovu Algoritma 4. Modul ima i mogućnost grafičkog prikaza transformacija trokuta koje se primjenjuju u iteracijama algoritma (vidjeti Primjer **??**, str.**??**).

Algorithm 4 **(Nelder-Meadov algoritam za funkcije dviju varijabli)**

Require: f, $A = (x_1, y_1)$, $B = (x_2, y_2)$, $C = (x_3, y_3)$

poredaj A, B, C tako da je $f(A) \le f(B) \le f(C)$
$M = \frac{1}{2}(A + B)$ //*centroid dobre strane*
$R = M + (M - C) = 2M - C$ //*refleksija*
if $f(R) < f(A)$ **then**
    $E = R + (R - M) = 2R - M$//*ekspanzija*
    if $f(E) < f(R)$ **then**
        $C = E$
    else
        $C = R$
    end if
else
    if $f(R) < f(B)$ **then**
        $C = R$
    else
        if $f(R) < f(C)$ **then**
            $P = M + \frac{1}{2}(R - M)$//*kontrakcija*
        else
            $P = M - \frac{1}{2}(M - C)$//*kontrakcija*
        end if
    end if
    if $f(P) < f(C)$ **then**
        $C = P$
    else
        $B = \frac{1}{2}(A + B)$//*skraćivanje*
        $C = \frac{1}{2}(A + C)$//*skraćivanje*
    end if
end if

Ensure: A, B, C

Zadatak 32. *Zadan je trokut $\triangle ABC$ s vrhovima $A = (5, 1)$, $B = 3, 4)$ i $C = (1, 2)$. Odredite vrhove trokuta nastalih refleksijom, ekspanzijom, unutarnjom i vanjskom kontrakcijom te skraćivanjem trokuta $\triangle ABC$.*

Zadatak 33. *Za funkciju $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = \max\{|x_1 + 2x_2 - 4|, |x_1 - x_2|\}$ provedite prve dvije iteracije* Nelder-Meadove metode *tako da krenete od trokuta s vrhovima*

$$(a) \quad A = (3, 0), B = (0, 3), C = (0, 0),$$
$$(b) \quad A = (3, 2), B = (1, 3), C = (3, 4).$$

U svakoj iteraciji odredite aproksimaciju minimuma kao vrijednost funkcije u centroidu trenutnog trokuta.

5.3 Coordinate Descent Algorithms

In the following section we are going to illustrate briefly a class of optimization methods which have been specially investigated in the past ten years in the context of *Big Data Analysis*, and can be found under the term *Coordinate Minimization algorithms* or *Coordinate Descent Algorithms* (see e.g. [2, 11, 12, 15]).

Instead of looking for a global minimum $\operatorname*{argmin}_{x\in\mathbb{R}^n} f(x)$ of a function $f\colon \mathbb{R}^n \to \mathbb{R}$, we will be satisfied by finding a point of local minimum or even by finding a stationary point. Let us represent vectors $x \in \mathbb{R}^n$ in the form $x = (u, v)$, $u \in \mathbb{R}^p$, $v \in \mathbb{R}^q$, $p + q = n$. Choosing some appropriate initial approximation (u_0, v_0), the iterative process can be constructed as follows

$$\begin{aligned} f(u_0, v_0) &\geq f(u_0, \operatorname*{argmin}_{v\in\mathbb{R}^q} f(u_0, v)) =: f(u_0, v_1) \\ &\geq f(\operatorname*{argmin}_{u\in\mathbb{R}^p} f(u, v_1), v_1) =: f(u_1, v_1) \geq \cdots \end{aligned}$$

or as

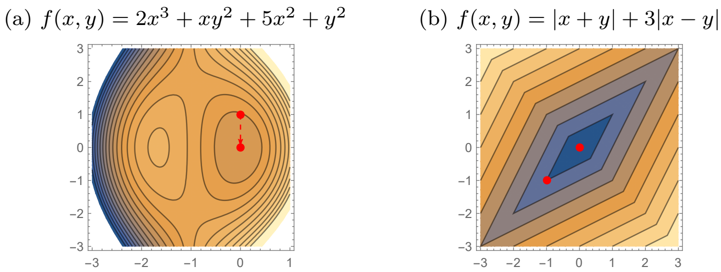
$$\begin{aligned} f(u_0, v_0) &\geq f(\operatorname*{argmin}_{u\in\mathbb{R}^p} f(u, v_0), v_0) =: f(u_1, v_0) \\ &\geq f(u_1, \operatorname*{argmin}_{v\in\mathbb{R}^q} f(u_1, v)) =: f(u_1, v_1) \geq \cdots \end{aligned}$$

This will generate a sequence of successive approximations $x^{(0)} = (u_0, v_0)$, $x^{(1)} = (u_1, v_1)$, $\ldots$, for which the sequence of function values is non-increasing. For a strictly convex differentiable function f the process will converge to the global minimum (see Example 5). As shown by Example 6, if the function f is not strictly convex, there is no guarantee that the algorithm will converge. This algorithm can be defined for differentiable but also for non-differentiable functions.

Primjer 5. *Find the local minimum of the function* $f\colon[-3,1]\times[-3,3]\to\mathbb{R}$,

$$f(x, y) = 2x^3 + xy^2 + 5x^2 + y^2, \tag{14}$$

whose `ContourPlot` *is shown in Fig. 13a. Check, using calculus, that f has four stationary points: $S_1 = (-\frac{5}{3}, 0)$, $S_2 = (-1, -2)$, $S_3 = (-1, 2)$, and $S_4 = (0, 0)$, that f attains its local maximum at S_1, its local minimum at S_4, and that S_2 and S_3 are saddle-points.*



Slika 13: Finding points of global and local minimum

Let us find the local minimum in the following way:

1. choose $y_0 = 1$ and define the function $f_x(x) := f(x, y_0) = 1 + x + 5x^2 + 2x^3$. The function f_x attains its local minimum at $x_0 = -0.10685$;

2. define the function $f_y(y) := f(x_0, y) = 0.054645 + 0.89315\,y^2$ which attains its local minimum at $y_1 = 0$;
3. define the function $f_x(x) := f(x, y_1) = 5x^2 + 2x^3$ which attains its local minimum at $x_1 = 0$;
4. define the function $f_y(y) := f(x_1, y) = y^2$ which attains its local minimum at $y_2 = 0$.

Here we stop the iterative process since the new values are going to repeat the ones we have already obtained. Thus $(0, 0)$ is the point of local minimum of the function f (see the red dots in Fig. 13a).

However, as the next example shows, the described procedure does not always perform nicely as in the previous example.

Primjer 6. *The function* $f\colon [-3, 3] \times [-3, 3] \to \mathbb{R}$, $f(x, y) = |x + y| + 3\,|x - y|$ *attains its global minimum at* $(0, 0)$ *(its* `ContourPlot` *is depicted in Fig. 13b). Try to find this minimum in a similar way as previously, starting at* $y_0 = -1$.

The function $f_x(x) := f(x, y_0) = |x - 1| + 3\,|x + 1|$ attains its minimum at $x_0 = -1$. In the next step the function $f_y(y) := f(x_0, y) = |y - 1| + 3\,|y + 1|$ attains its minimum at $y_1 = -1$. Further procedures repeat themselves, thus not producing the point of global minimum.

Primjer 7. *Traženje najboljeg LS-pravca (Primjer* **??***) možemo promatrati kao iterativni proces [13].*

Najprije za izabranu početnu aproksimaciju $k_0 \in \mathbb{R}$ *promatramo problem jednodimenzionalne globalne optimizacije*

$$\operatorname*{argmin}_{\ell\in\mathbb{R}} \sum_{i=1}^{m} (y_i - k_0 x_i - \ell)^2. \tag{15}$$

Rješenje ovog problema dobiveno prema Primjeru **??***, str.***??** *označimo s*

$$\ell_0 = \bar{y} - k_0 \bar{x}. \tag{16}$$

U sljedećem koraku promatramo problem jednodimenzionalne globalne optimizacije

$$\operatorname*{argmin}_{k\in\mathbb{R}} \sum_{x_i\neq 0} x_i^2 \left(\tfrac{y_i - \ell_0}{x_i} - k\right)^2, \tag{17}$$

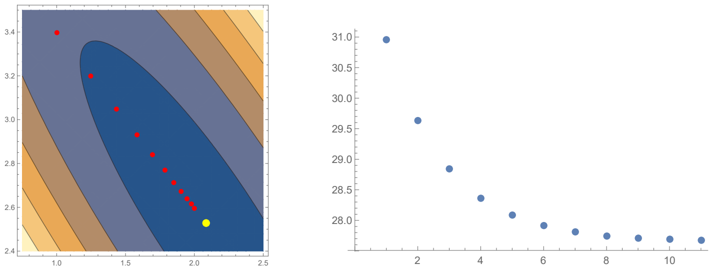
čije je rješenje ponovo prema Primjeru **??***, str.***??**

$$k_1 = \frac{1}{\sum\limits_{x_i\neq 0} x_i^2} \sum_{x_i\neq 0} x_i (y_i - \ell_0). \tag{18}$$

Ponavljajući postupak dobivamo niz (k_i, ℓ_i), $i = 0, 1 \ldots$ *koji konvergira prema točki* $(k^\star, \ell^\star)$ *iz Primjera* **??***, a granična vrijednost niza* $F_{LS}(k_i, \ell_i)$ *je vrijednost funkcije* $F_{LS}(k^\star, \ell^\star)$.

Za podatke iz Primjera **??** *na Slici 14a prikazan je* `ContourPlot` *funkcije* F_{LS}*, niz aproksimacija* (k_i, ℓ_i) *(crvene točkice) i rješenje* $(k^\star, \ell^\star)$ *(žuta točka), a na Slici 14b prikazan je niz funkcijskih vrijednosti* $(F_{LS}(k_i, \ell_i))$.

(a) `ContourPlot` (b) $(F_{LS}(k_i, \ell_i))$



Slika 14: Niz sukcesivnih aproksimacija

Zadatak 34. *Za skup podataka* $\mathcal{A} = \{a^i = (x_1, y_i) \colon i = 1, \dots, m\}$ *s težinama* $w_1, \dots, w_m > 0$ *u Primjeru* **??** *promatra se problem traženja najboljeg* ***LAD****-pravca*

$$\operatorname*{argmin}_{k,l \in \mathbb{R}} \sum_{i=1}^{m} w_i |y_i - kx_i - \ell|$$

Za podatke generirane kao u Primjeru **??** *potražite najbolji* ***LAD****-pravac primjenom* Coordinate Descent Algorithm*.*

Zadatak 35. *Za skup podataka* $\mathcal{A} = \{a^i = (x_1, y_i) \colon i = 1, \dots, m\}$ *s težinama* $w_1, \dots, w_m > 0$ *u Primjeru* **??** *promatra se problem traženja najboljeg* ***OD****-pravca.*

Za podatke generirane kao u Primjeru **??** *potražite najbolji* ***OD****-pravac u eksplicitnom obliku primjenom* Coordinate Descent Algorithm*.*
Uputa: Pravac tražite u obliku $y = kx + \ell$*. U iterativnom postupku parametar* ℓ *odredite kao težinski medijan, a parametar* k *primjenom* Nelder-Meadove *metode ili* Metodom tri točke *(vidi t.**??**, str.**??**).*

U Primjeru **??** za skup podataka $\mathcal{A} = \{a^i = (x_1, y_i) \colon i = 1, \dots, m\}$ s odgovarajućim težinama $w_1, \dots, w_m > 0$ promatrali smo problem traženja najboljeg TLS-pravca u implicitnom obliku. Ako pravac potražimo u eksplicitnom obliku $y = kx + \ell$, onda odgovarajući TLS-problem glasi

$$\operatorname*{argmin}_{k,\ell \in \mathbb{R}} \sum_{i=1}^{m} w_i \frac{(y_i - kx_i - \ell)^2}{k^2 + 1}, \tag{19}$$

a ako uvažimo činjenicu da najbolji TLS-pravac prolazi centroidom podataka (x_p, y_p), problem(19) prelazi u problem traženja globalnog minimuma derivabilne funkcije jedne varijable

$$\operatorname*{argmin}_{k \in \mathbb{R}} \sum_{i=1}^{m} w_i \frac{(y_i - y_p - k(x_i - x_p))^2}{k^2 + 1}. \tag{20}$$

O minimizaciji ovakvih funkcija govorit ćemo u nastavku.

5.4 Kvadratne interpolacijske metode za jednodimenzionalnu minimizaciju

Pretpostavimo da je zadana funkcija $f\colon \mathbb{R} \to \mathbb{R}$ koja u točki $x^\star \in \mathbb{R}$ postiže lokalni minimum. U nastavku ćemo opisati jednu klasu metoda za numeričko određivanje točke lokalnog minimuma poznatu kao kvadratne interpolacijske metode. Zajedničko kvadratnim interpolacijskim metodama je interpolacija funkcije f kvadratnom funkcijom u jednoj, dvjema ili trima točkama. Tada se umjesto traženja točke lokalnog minimuma funkcije f traži točka globalnog minimuma kvadratne funkcije te se postupak uz određena pravila iterativno nastavlja. U slučaju interpolacije u jednoj točki dobivamo najčešće korištenu *Newtonovu metodu* ili *Metodu jedne točke*, dok u slučajevima interpolacije u dvjema odnosno trima točkama, dobivamo *Metodu dvije točke*, odnosno *Metodu tri točke*. Ako je funkcija f konveksna i ako ima točku lokalnog minimuma, onda spomenute metode pronalaze globalni minimum.

5.4.1 Newtonova metoda – metoda jedne točke

U t.5.1.3, str.17 već smo razmatrali Newtonovu metodu za traženje lokalnog minimuma funkcije $f \in C^2(\mathcal{D})$, $\mathcal{D} \subseteq \mathbb{R}^n$. Ovdje ćemo specijalno za funkciju jedne varijable $f \in C^2(\mathbb{R})$ dati geometrijski smisao Newtonove metode. Neka je zadana točka $x_0 \in \mathbb{R}$, za koju pretpostavljamo da leži dovoljno blizu točke minimuma $x^\star$ koju želimo odrediti. Funkciju f u okolini točke x_0 aproksimirat ćemo kvadratnom funkcijom $g : \mathbb{R} \to \mathbb{R}$, $g(x) = \alpha x^2 + \beta x + \gamma$, pri čemu koeficijente α, β, γ određujemo iz sljedećih uvjeta:

$$\begin{aligned} g(x_0) &= \alpha x_0^2 + \beta x_0 + \gamma = f(x_0) =: f_0, \\ g'(x_0) &= 2\alpha x_0 + \beta = f'(x_0) =: f_0', \\ g''(x_0) &= 2\alpha = f''(x_0) =: f_0''. \end{aligned} \tag{21}$$

Rješavanjem sustava (21) dobivamo:

$$\alpha = \tfrac{1}{2} f_0'', \quad \beta = f_0' - f_0'' x_0.$$

Umjesto traženja točke lokalnog minimuma funkcije f potražit ćemo točku u kojoj se postiže globalni minimum njezine aproksimacije g. Kako je g kvadratna funkcija, njezin globalni minimum postiže se u točki

$$\overline{x} = -\frac{\beta}{2\alpha} = x_0 - \frac{f_0'}{f_0''},$$

odakle dobivamo rekurzivnu formulu na kojoj se temelji poznata *Newtonova metoda*:

$$x_{k+1} = x_k - \frac{f_k'}{f_k''}, \quad k = 0, 1, \ldots, \tag{22}$$

pri čemu su $f'_k := f'(x_k)$ te $f''_k := f''(x_k)$, $k = 0, 1, \ldots$.

U udžbenicima iz numeričke analize mogu se pronaći teoremi o konvergenciji *Newtonove metode* te njezinoj brzini konvergencije. S obzirom na to da je riječ o standardnim rezultatima, sljedeći teorem navodimo bez dokaza. Iskaz i dokaz ovog teorema s nešto blažim uvjetima na funkciju f možemo naći u [17].

Teorem 2. *Neka je $f\colon \mathbb{R} \to \mathbb{R}$ tri puta neprekidno diferencijabilna funkcija. Ako je $x^\star \in \mathbb{R}$ stacionarna točka takva da je $f''(x^\star) \neq 0$, onda postoji $\delta > 0$ takav da za svaki $x_0 \in \langle x^\star - \delta, x^\star + \delta\rangle$ niz (x_k) zadan rekurzivno sa (22) konvergira prema $x^\star$ kvadratnom brzinom konvergencije.*

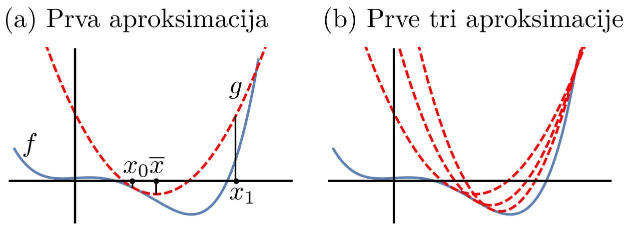
Primijetimo da je za primjenu rekurzivne formule (22) dovoljno da funkcija ima drugu derivaciju. Međutim, kako bi se dokazao prethodno navedeni teorem o konvergenciji te brzini konvergencije, postavljen je zahtjev da je funkcija tri puta neprekidno derivabilna. Situacija u kojoj su za primjenu metode potrebni blaži uvjeti, dok su za bilo kakav teorijski rezultat potrebni puno jači uvjeti, vrlo je česta u numeričkoj analizi.

U sljedećim pododjeljcima analizirat ćemo metode za koje nije potrebno izračunavanje druge derivacije, već samo vrijednost prve derivacije ili samo vrijednost funkcije. Smanjenje zahtjeva na funkciju očekivano će rezultirati smanjenjem brzine konvergencije odgovarajuće metode.

5.4.2 Metoda dvije točke

Metoda I

Neka su $x_0, x_1 \in \mathbb{R}$, $x_0 \neq x_1$ dvije točke za koje su poznate vrijednosti $f(x_0)$, $f'(x_0)$ i $f'(x_1)$, pri čemu dodatno pretpostavljamo da je $f'(x_0) \neq f'(x_1)$.



Slika 15: Graf funkcije f i kvadratnih interpolacijskih funkcija

U tom slučaju možemo potražiti kvadratnu interpolacijsku funkciju $g\colon \mathbb{R} \to \mathbb{R}$ oblika $g(x) = \alpha x^2 + \beta x + \gamma$, $\alpha, \beta, \gamma \in \mathbb{R}$, sa sljedećim svojstvima (vidi Sliku 15):

$$\begin{aligned} g(x_0) &= \alpha x_0^2 + \beta x_0 + \gamma = f(x_0) =: f_0, \\ g'(x_0) &= 2\alpha x_0 + \beta = f'(x_0) =: f'_0, \\ g'(x_1) &= 2\alpha x_1 + \beta = f'(x_1) =: f'_1. \end{aligned} \tag{23}$$

Rješavanjem sustava linearnih jednadžbi (23) dobivamo

$$\alpha = \frac{f_1' - f_0'}{2(x_1 - x_0)}, \quad \beta = \frac{x_1 f_0' - x_0 f_1'}{x_1 - x_0}.$$

Kvadratna funkcija g postiže globalni minimum u točki

$$\overline{x} = -\frac{\beta}{2\alpha} = x_0 - \frac{x_1 - x_0}{f_1' - f_0'} f_0'.$$

Općenito, dobivamo rekurzivnu formulu

$$x_{k+1} = x_{k-1} - \frac{x_k - x_{k-1}}{f_k' - f_{k-1}'} f_{k-1}', \quad k = 1, 2, \ldots, \tag{24}$$

pri čemu su $f_k' := f'(x_k)$, $k = 0, 1, \ldots$.

Primjedba 8. Primijetite da je $\frac{f_k' - f_{k-1}'}{x_k - x_{k-1}}$ aproksimacija funkcije f'' u točki x_{k-1} pa rekurzivnu formulu (24) zovemo *Metoda sekanti* čija generalizacija na funkcije s n nezavisnih varijabli vodi prema već spomenutoj važnojn klasi Kvazi-Newtonovih metoda (vidi t.**??**, str.**??**).

Algorithm 5 **(Metoda dvije točke)**

Require: f, x_0, x_1, ε
1: $f_0' = f'(x_0)$, $f_1' = f'(x_1)$
2: $\bar{x} = x_0 - \frac{x_0 - x_1}{f_0' - f_1'} f_0'$
3: **while** $|f'(\bar{x})| < \varepsilon$ (ili dok nije zadovoljen neki drugi kriterij zaustavljanja) **do**
4: $x_0 = x_1$, $x_1 = \bar{x}$
5: $f_0' = f'(x_0)$, $f_1' = f'(x_1)$
6: $\bar{x} = x_0 - \frac{x_0 - x_1}{f_0' - f_1'} f_0'$
7: **end while**
Ensure: $\bar{x}$

U sljedećem teoremu odgovaramo na pitanje konvergencije i brzine konvergencije iterativnog postupka Algoritma 5 (vidi [20, 23]).

Teorem 3. *Neka je $f\colon \mathbb{R} \to \mathbb{R}$ tri puta neprekidno diferencijabilna funkcija. Ako je $x^\star \in \mathbb{R}$ stacionarna točka takva da je $f''(x^\star) \neq 0$ i $f'''(x^\star) \neq 0$, onda postoji $\delta > 0$ takav da za svaki $x_0, x_1 \in \langle x^\star - \delta, x^\star + \delta\rangle$, $x_0 \neq x_1$, niz (x_k) zadan rekurzivno s (24) konvergira prema $x^\star$, pri čemu je brzina konvergencije metode $r = \frac{1+\sqrt{5}}{2} \approx 1.618$.*

Iz iterativne procedure (24) proizlazi Algoritam 5. Uočimo da smo u Algoritmu 5 naveli kriterij zaustavljanja, prema kojemu algoritam staje kada je derivacija funkcije f u trenutnoj aproksimaciji manja od unaprijed zadanog broja $\varepsilon > 0$. U literaturi postoje i drugi kriteriji zaustavljanja, ali budući da zahtijevaju posebnu analizu i obrazloženja, ovdje ih nećemo razmatrati.

Metoda II

U ovom pododjeljku kratko ćemo opisati još jednu *Metodu dvije točke*. Slično kao i u prvoj metodi, pretpostavimo da su $x_0, x_1 \in \mathbb{R}$, $x_0 < x_1$ dvije različite točke za koje su poznate vrijednosti $f(x_0)$, $f(x_1)$ te $f'(x_0)$, pri čemu je $f(x_0) \neq f(x_1)$. U tom slučaju tražimo kvadratnu interpolacijsku funkciju $g\colon \mathbb{R} \to \mathbb{R}$, $g(x) = \alpha x^2 + \beta x + \gamma$ sa sljedećim svojstvima:

$$\begin{aligned} g(x_0) &= \alpha x_0^2 + \beta x_0 + \gamma = f(x_0) =: f_0, \\ g(x_1) &= \alpha x_1^2 + \beta x_1 + \gamma = f(x_1) =: f_1, \\ g'(x_0) &= 2\alpha x_0 + \beta = f'(x_0) =: f_0'. \end{aligned} \tag{25}$$

Rješavanjem sustava (25) dobivamo

$$\begin{aligned} \alpha &= \tfrac{f_0 - f_1 - f_0'(x_0 - x_1)}{-(x_0 - x_1)^2}, \\ \beta &= f_0' + 2x_0 \tfrac{f_0 - f_1 - f_0'(x_0 - x_1)}{(x_0 - x_1)^2}, \end{aligned}$$

odakle dobivamo sljedeću aproksimaciju

$$x_2 = -\tfrac{\beta}{2\alpha} = x_0 - \tfrac{1}{2} \frac{(x_0 - x_1) f_0'}{f_0' - \frac{f_1 - f_0}{x_1 - x_0}}. \tag{26}$$

i općenitu rekurzivnu formulu:

$$x_{k+1} = x_{k-1} - \tfrac{1}{2} \frac{(x_{k-1} - x_k) f_{k-1}'}{f_{k-1}' - \frac{f_k - f_{k-1}}{x_k - x_{k-1}}}, \tag{27}$$

pri čemu su $f_k := f(x_k)$, $f_k' := f'(x_k)$, $k = 0, 1, \ldots$.

Za metodu koja se zasniva na iterativnoj proceduri (27) po uzoru na dokaz Teorema 3 može se pokazati analogna tvrdnja.

Zadatak 36. Metodom dvije točke *odredite aproksimaciju minimuma funkcije* $f(x) = x^3 - 4x + 5$ *korištenjem kriterija zaustavljanja iz Algoritma 5 uz* $\varepsilon = 0.005$. *Upotrijebite obje varijante metode.*

5.4.3 Metoda tri točke

Neka je $f\colon \mathbb{R} \to \mathbb{R}$ neprekidna funkcija koja općenito ne mora biti derivabilna. Kao i do sada, s $x^\star$ označimo nepoznatu točku u kojoj se postiže lokalni minimum funkcije f. Odaberimo tri međusobno različite početne aproksimacije x_0, x_1 i x_2 i to tako da su točke $(x_0, f(x_0)), (x_1, f(x_1))$ te $(x_2, f(x_2))$ nekolinearne te da vrijedi

$$\min_{i=0,1,2} x_i \leq x^\star \leq \max_{i=0,1,2} x_i.$$

U tom slučaju postoji jedinstveni interpolacijski polinom g drugog stupnja koji prolazi kroz te tri točke te glasi [17]:

$$g(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} f_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} f_1 + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} f_2,$$

gdje su $f_k := f(x_k), k = 0, 1, 2.$

Funkcija g je kvadratna funkcija, koja u standardnom zapisu glasi $g(x) = \alpha x^2 + \beta x + \gamma$, a njena točka globalnog minimuma je $\overline{x} := \operatorname*{argmin}_{x\in\mathbb{R}} g(x) = -\frac{\beta}{2\alpha}$. Nije teško vidjeti da je

$$\overline{x} = \tfrac{1}{2}(x_0 + x_1) + \tfrac{1}{2}\frac{(f_0 - f_1)(x_1 - x_2)(x_2 - x_0)}{(x_1 - x_2)f_0 + (x_2 - x_0)f_1 + (x_0 - x_1)f_2}. \tag{28}$$

Iz formule (28) proizlazi metoda kod koje polazeći od trojke točaka x_0, x_1 i x_2 dobivamo novu točku $\overline{x}$. U svakoj od tako dobivenih četvorki točaka izračunamo vrijednost funkcije f te među njima biramo one tri točke $\tilde{x}_0, \tilde{x}_1, \tilde{x}_2$ u kojima je vrijednost funkcije najmanja.

Algorithm 6 **(Metoda tri točke)**

Require: f, ε, x_0, x_1, x_2 (uz uvjet $\min\{x_0, x_1, x_2\} \le x^\star \le \max\{x_0, x_1, x_2\}$)

1: $f_0 = f(x_0)$, $f_1 = f(x_1)$, $f_2 = f(x_2)$
2: $\bar{x} = \frac{1}{2}(x_0 + x_1) + \frac{1}{2}\frac{(f_0-f_1)(x_1-x_2)(x_2-x_0)}{(x_1-x_2)f_0+(x_2-x_0)f_1+(x_0-x_1)f_2}$
3: **if** $(\overline{x} - x_0)(\overline{x} - x_2) \ge 0$ **then**
4:     **goto** Korak 8
5: **else**
6:     **goto** Korak 9
7: **end if**
8: Konstruirati novu trojku $\{x_0, x_1, x_2\}$ iz x_0, x_1, x_2 i $\overline{x}$ za koje je vrijednost funkcije f najmanja i **goto** Korak 1
9: **if** $|\overline{x} - x_1| < \varepsilon$ **then**
10:     **goto** Korak 13
11: **else**
12:     **goto** Korak 8
13: **end if**

Ensure: $\bar{x}$

Postupak nastavljamo sve dok se ne zadovolji neka unaprijed zadana točnost. Cijeli postupak opisan je u Algoritmu 6.

Vrijedi sljedeći teorem, čiji iskaz i dokaz možemo naći primjerice u [23].

Teorem 4. *Neka je $f : \mathbb{R} \to \mathbb{R}$ četiri puta neprekidno derivabilna funkcija. Ako je $x^\star \in \mathbb{R}$ stacionarna točka takva da je $f''(x^\star) \neq 0$, $f'''(x^\star) \neq 0$ i $f^{(iv)}(x^\star) \neq 0$, onda postoji $\delta > 0$ takav da za svaki $x_0, x_1, x_2 \in \langle x^\star - \delta, x^\star + \delta\rangle$, $x_i \neq x_j$, $i \neq j$, Algoritam 6 konvergira prema $x^\star$, pri čemu je brzina konvergencije metode $r = 1.32$.*

Analogno kao i do sada u svrhu dokaza teorema potrebno je postaviti znatno jači zahtjev na funkciju f od onoga koji je potreban za provođenje Algoritma 6.

Zadatak 37. *Traženu funkciju $g(x) = \alpha x^2 + \beta x + \gamma$ čiji graf prolazi točkama $(x_0, f(x_0))$, $(x_1, f(x_1))$, $(x_2, f(x_2))$ odredite kao Newtonov interpolacijski polinom drugog reda i tako odredite točku $\overline{x} := \operatorname*{argmin}_{x\in\mathbb{R}} g(x) = -\frac{\beta}{2\alpha}$.*

Zadatak 38. *Zadana je konveksna funkcija* $f \in C^1([a,b])$ *koja u točki* $\xi \in [a,b]$ *postiže jedinstveni globalni minimum. Neka su* $x_0, x_1, x_2 \in [a,b]$ *tri različite točke za koje su poznate vrijednosti* $f(x_0)$*,* $f(x_1)$ *te* $f'(x_2)$*.*

(a) Odredite kvadratnu interpolacijsku funkciju $g : [a,b] \to \mathbb{R}$*,* $g(x) = \alpha x^2 + \beta x + \gamma$ *sa sljedećim svojstvima:*

$$g(x_0) = f(x_0) =: f_0, \quad g(x_1) = f(x_1) =: f_1, \quad g'(x_2) = f'(x_2) =: f_2'.$$

(b) Odredite točku u kojoj funkcija g *postiže globalni minimum.*

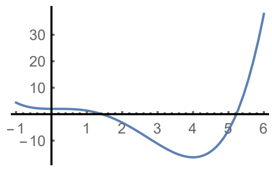
(c) Na osnovi (b) predložite iterativnu formulu koja će konvergirati jedinstvenom globalnom minimumu funkcije f*.*

5.4.4 Numerički primjer

U ovom ćemo odjeljku pomoću jednog jednostavnog numeričkog primjera ilustrirati efikasnost prethodno navedenih metoda te potvrditi teorijske rezultate koji se odnose na brzinu konvergencije pojedine metode. Promatrajmo u tu svrhu funkciju $f\colon \mathbb{R} \to \mathbb{R}$ zadanu formulom

$$f(x) = 2 - \tfrac{2}{25}x + \tfrac{61}{100}x^2 - \tfrac{43}{30}x^3 + \tfrac{1}{4}x^4.$$

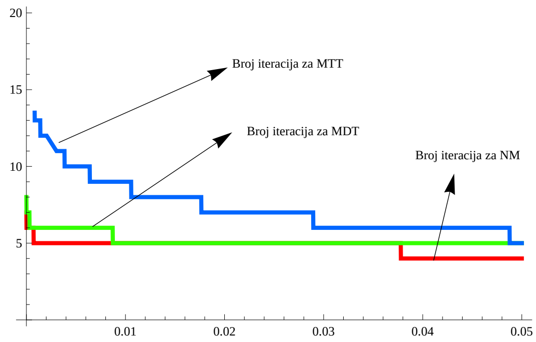
Graf funkcije f prikazan je na Slici 16.



Slika 16: Graf funkcije f

Nije teško vidjeti da funkcija f u točki $x^\star = 4$ postiže lokalni minimum, koji je ujedno i njezin globalni minimum. Svaku od prethodno opisanih metoda primjenjujemo na funkciju f uz početne aproksimacije $x_0 = 1.5, x_1 = 3$ i $x_2 = 6$. Kriterij zaustavljanja definiramo tako da algoritam staje ako je $|x_k - 4| < \varepsilon$, gdje je $\varepsilon > 0$ unaprijed zadana točnost. Pri tome za različite $\varepsilon \in [10^{-8}, 0.01]$ određujemo broj iteracija koji je potreban da bi se odgovarajući algoritam zaustavio. Početne aproksimacije odabrane su tako da sve metode konvergiraju.

Na Slici 17 prikazan je broj iteracija za *Metodu jedne točke-Newtonovu metodu* (`NM`), *Metodu dvije točke* (`MDT`) i *Metodu tri točke* (`MTT`) u ovisnosti o parametru ε. Kako je i očekivano, na Slici 17 vidimo da `NM` treba najmanje iteracija, `MDT` nešto više, dok `MTT` treba značajno najviše iteracija kako bi se zadovoljila zadana točnost.



Slika 17: Broj iteracija za *Newtonovu metodu* (NM), *Metodu dvije točke* (MDT) i *Metodu tri točke* (MTT)

Literatura

[1] M. Avriel, *Nonlinear Programming: Analysis and Methods.*, Dover Publications, Inc. Mineola, New York, 2006.

[2] A. Beck, L. Tetruashvili, *On the convergence of block coordinate descent type methods*, SIAM Journal on Optimization, (2013).

[3] J. F. Bonnans, J. C. Gilbert, C. Lemarechal, C. A. Sagastizabal, *Numerical Optimization. Theoretical and Practical Aspects*, Springer-Verlag, Berlin, 2006.

[4] D. L. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

[5] J. J. Dennis, R. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, 1996.

[6] E. M. T. Hendrix, B. G. Tóth, *Introduction to Nonlinear and Global Optimization*, Springer, 2010.

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

[8] D. Jukić, *Konveksni skupovi*, Odjel za matematiku, Sveučilište u Osijeku, 2015.

[9] J. C. Lagarias, J. A. Reeds, M. H. Wright, M. H. P. Wright, *Convergence properties of the Nelder-Mead simplex algorithm in low dimensions*, SIAM J. Optim., **9**(1998) 112–147.

[10] J. A. Nelder, R. Mead, *A simplex method for function minimization*, The Computer Journal, **7**(1965) 308–313.

[11] Y. Nestorov, *Efficiency of coordinate descent methods on huge-scale optimization problems*, SIAM Journal on Optimization, **22.2**(2012) 341–362.

[12] J. Nutini, M. Schmidt, I. H. Laradji, M. Friedlander, H. Koepke, *Coordinate descent converges faster with the Gauss-Southwell rule than random selection*, In: *Proceedings of the 32nd International Conference on Machine Learning (ICML-15).*, 2015.

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

[14] K. V. Price, R. M. Storn, J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*, Springer Verlag, Berlin, Heidelberg, 2005.

[15] P. Richtarik, M. Takac, *Iteration complexity of randomized block-coordinate descent methods for minimizing a composite function*, Mathematical Programming, **144.1-2**(2014) 1–38.

[16] R. Scitovski, *Problemi najmanjih kvadrata. Financijska matematika*, Ekonomski fakultet, Elektrotehnički fakultet, Sveučilište u Osijeku, 1993.

[17] 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`.

[18] R. Scitovski, N. Truhar, Z. Tomljanović, *Metode optimizacije*, Odjel za matematiku, Sveučilište u Osijeku, 2014.

[19] S. Singer, J. Nelder, *Nelder-Mead algorithm*, Scholarpedia, **4**(2009) 2928.

[20] G. W. Stewart, *Afternotes on Numerical Analysis*, SIAM, Philadelphia, 1996.

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

[22] E. Süli, D. Mayers, *An Introduction to Numerical Analysis, 2nd Ed.*, Cambridge University Press, Second printing, 2006.

[23] W. Sun, Y. Yuan, *Optimization theory and methods. Nonlinear programming*, Springer-Verlag, New York, 2006.