Bezuvjetna optimizacija

13. siječnja 2020.

Problem:
za danu funkciju $f : D \subset \mathbb{R}^n \to \mathbb{R}$ odrediti $\boldsymbol{x}_0 \in \mathbb{R}^n$ tako da

$$f(\boldsymbol{x}_0) = \min_{\boldsymbol{x} \in D} f(\boldsymbol{x}).$$

Podsjetnik:

- neka je D kompaktan podskup. Tada neprekidna funkcija $f : D \to \mathbb{R}$ postiže globalni minimum/maksimum na D.
- (lokalni ekstrem funkcije $f : D \to \mathbb{R}$)
  Kažemo da je $\boldsymbol{x}_0$ lokalni minimum ako postoji $\varepsilon > 0$ takav da $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$, za svaki $\boldsymbol{x} \in K(\boldsymbol{x}_0, \varepsilon) \cap D$.

Koercitivnost

Definicija 1

Kažemo da je funkcija $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *koercitivna ako za svaki niz* $(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset D$ *takav da* $\|\boldsymbol{x}_n\| \to +\infty$ *vrijedi* $f(\boldsymbol{x}_n) \to +\infty$.

Nap:

- Koercitivnost funkcije f ekvivalentna je:

$$(\forall M > 0)(\exists R > 0)(\forall \boldsymbol{x} \in D) \quad \|\boldsymbol{x}\| > R \implies f(\boldsymbol{x}) > M.$$

- Nivo skup $l_\alpha(f) = \{x \in D | f(x) \leq \alpha\}$ je ograničen skup.
- **Tvrdnja:**
  Neka je $f : D \to R$ neprekidna, koercitivna funkcija i $D \subset \mathbb{R}^n$ zatvoren. Tada funkcija f postiže globalni minimum na D.

Nužan uvjet lokalnog ekstrema

Teorem 1 (Nužan uvjet lokalnog ekstrema)

Neka je $f : D \to \mathbb{R}$ funkcija, D otvoren podskup u $\mathbb{R}^n$. Neka je funkcija diferencijablina u točki $\boldsymbol{x} \in D$. Ako funkcija f postiže lokalni ekstrem u $\boldsymbol{x}$ tada $\nabla f(\boldsymbol{x}) = 0$.

Označimo s $\Omega := \{\boldsymbol{x} \in D \; : \; \nabla f(\boldsymbol{x}) = 0\}$ skup kritičnih (stacionarnih) točaka funkcije f.

Dovoljni uvjeti lokalnih max / min

$$A \in \mathbb{R}^{n\times n}_{\mathrm{Sym}} \text{ je pozitivno (negativno) } \begin{cases} \text{semidefinitna } \forall \xi \in \mathbb{R}^n,\ \langle A\xi, \xi\rangle \overset{(\leq)}{\geq} 0 \\ \text{definitna } \forall \xi \in \mathbb{R}^n \setminus \{0\},\ \langle A\xi, \xi\rangle \underset{(<)}{>} 0 \end{cases}$$

Ako nije ništa od gore navedenog kažemo da je indefinitna.
Sylvesterov kriterij: hermitska matrica je pozitivno definitna ako i samo ako su vodeće minore matrice strogo pozitivne.

Teorem 2 (Dovoljni uvjeti)

$U \subset \mathbb{R}^n$ otvoren, $f : U \to \mathbb{R}$ klase $\mathcal{C}^2$. Neka je x_0 kritična točka:

1. *ako je $D^2 f(x_0)$ negativno definitna u x_0 se poprima lokalni maksimum,*
2. *ako je $D^2 f(x_0)$ pozitivno definitna u x_0 se poprima lokalni minimum,*
3. *ako je $D^2 f(x_0)$ indefinitna, x_0 je sedlasta točka.*

Konveksne funkcije

Definicija 2

Kažemo da je funkcija $f : D \subset \mathbb{R}^n \to \mathbb{R}$ konveksna ako

$$(\forall x \in D)(\forall y \in D)(\forall t \in [0,1]) \quad f(tx+(1-t)y) \leq tf(x)+(1-t)f(y).$$

Karakterizacije:

- f je konveksna $\iff$ $A = \{(x,y) \in \mathbb{R}^n \times \mathbb{R}, y \geq f(x)\}$ je konveksan skup.
- $f \in \mathcal{C}^1$ je konveksna $\iff$ $(\forall x \in D)(\forall y \in D) \quad f(x) - f(y) \geq \nabla f(y)^\tau (x-y)$

Teorem 3

Ako je $f : \mathbb{R}^n \to \mathbb{R}$ konveksna funkcija tada je lokalni minimum ujedno i globalni. Dodatno, ako je funkcija diferencijabilna tada se minimum postiže u svim stacionarnim točkama.

Numeričke metode

Problem: tražimo lokalne minimume funkcije $f : \mathbb{R}^n \to \mathbb{R}$.
Označimo s $x_0 \in \mathbb{R}^n$ početnu aproksimaciju. Iterativnom metodom generiramo niz $(x_k)_{k\in\mathbb{N}}$:

$$x_{k+1} = x_k + \lambda_k d_k, \quad k \in \mathbb{N}_0. \tag{1}$$

Notacija:

- d_k - vektor smjera,
- λ_k - faktor koji određuje veličinu koraka.

Izbor d_k?

Definicija 3

Neka je $f : \mathbb{R}^n \to \mathbb{R}$, $x_0 \in \mathbb{R}^n$. Za vektor $d \in \mathbb{R}^n$ kažemo da je vektor silaska funkcije f u točki x_0 ako postoji $\varepsilon > 0$ tako da

$$\forall \lambda \in \langle 0, \varepsilon \rangle \quad f(x_0 + \lambda d) < f(x_0).$$

Nap:

- ako je $f \in \mathcal{C}^1$ funkcija u točki x "najbrže" pada u smjeru $-\nabla f(x)$,
- ako je $d \in \mathbb{R}^n$ takav da je $\nabla f(x)^\tau d < 0$ onda je d vektor silaska funkcije f u točki x,
- ako koristimo za vektor smjera u (1) $d_k = -\nabla f(x_k)$ metodu nazivamo *gradijentnom metodom*.

Odabir veličine koraka λ_k je netrivijalan. Ako zahtijevamo da se vrijednost funkcije f najviše smanji tada rješavamo jednodimenzionalan nelinearan problem:

$$(2) \qquad \begin{cases} \varphi_k(\lambda) = f(x_k - \lambda \nabla f(x_k)) \to \min \\ \lambda \geq 0 \end{cases}$$

Takav odabir λ_k nazivamo egzaktnim. Metodu zovemo *metodom najbržeg silaska* (MNS).
U praksi se (2) riješava koristeći heuristike poput *Armijovog pravila*.

Zadatak 1

Napravite dvije iteracije metode najbržeg silaska funkcije

$$f(x, y) = x^2 + y^2 - 2x + 4y + 1.$$

uz početnu aproskimaciju $(x_0, y_0) = (2, 1)$.

Zadatak 2

Napravite dvije iteracije metode najbržeg silaska funkcije

$$f(x, y) = (x - 1)^2 + (xy - 2)^2$$

uz početnu aproksimaciju $\boldsymbol{x}_0 = (0, -1)$. *Konvergiraju li iteracije prema točki minimuma?*

Zadatak 3

Napravite dvije iteracije metode najbržeg uzlaska s problemom:

$$f(x, y, z) = -4x^2 + 4y + 4xy - 11y^2 + 12yz - 4z^2 \to \max,$$

uz početnu aproksimaciju $\boldsymbol{x} = (0, 0, 0)$. *Kolika je udaljenost 1. i 2. iteracije od rješenja?*

Lema 1

Ako vrijedi barem jedna od sljedećih pretpostavki:

- *funkcija $f \in \mathcal{C}^1$ koercitivna,*
- $(\exists m > 0)(\forall x \in \mathbb{R}^n) \quad Hf(x) - m\mathrm{I} \geq 0,$

niz $(x_k)_{k\in\mathbb{N}}$ generiran MNS je ograničen za svaku početnu aproksimaciju $x_0 \in \mathbb{R}^n$.

Teorem 4 (Dovoljni uvjeti konvergencije MNS)

Neka je $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ te niz $(x_k)_{k\in\mathbb{N}}$ generiran MNS ograničen. Tada svaki konvergentan podniz niza (x_k) ima limes u skupu stacionarnih točaka Ω. Ako je skup stacionarnih točaka jednočlan, tada je niz konvergentan.

Zadatak 4

Dokažite konvergenciju u Zadatku 1. Pokažite da funkcija u Zadatku 2 nije koercitivna.

Newtonova metoda - podsjetnik

Do sada smo upoznali klasičnu Newtonovu metodu za pronalazak stacionarnih točaka (korijena jednadžbe). Postupak se može generalizirati na više dimenzija.
Ideja: aproksimiramo nelinarnu funkciju afinom, te tražimo nultočku afine aproksimacije.

jednodimenzionalno:
$f \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

višedimenzionalno:
$g \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$

$$x_{k+1} = x_k - \nabla g(x_k)^{-1} g(x_k)$$

Newtonova metoda za traženje minimuma/maksimuma funkcije

Ako radimo s konveksnim (konkavnim) funkcijama svaka stacionarna točka je ujedno i točka minimuma (maksimuma). Drugim riječima moramo pronaći skup

$$\Omega = \{x \mid \nabla f(x) = 0\}$$

Stavimo li $g(x) = \nabla f(x)$ u Newtonovu metodu za pronalazak stacionarnih točaka dobivamo:

$$\boxed{x_{k+1} = x_k - Hf(x_k)^{-1}\nabla f(x_k)}$$

što zovemo **Newtonova metoda za ekstreme**.
Alternativni izvod Newtonove metode za ekstreme:
aproksimacija funkcije kvadratnim polinom (Taylorovim) te traženje ekstrema aproksimacije.

Zadatak 5

Napravite dvije iteracije Newtonove metode uz početnu aproksimaciju $(1/2, 2)$ *za problem*

$$f(x,y) = \ln(1+x^2) + y^2 \to \min .$$

Zadatak 6

Dokažite da funkcija $f(x,y) = x^2y^2 + x^2 + y^2$ *ima jedinstvenu točku lokalnog minimuma u kojoj se ujedno dostiže globalni minimum. Objasnite što se događa s Newtonovom metodom za traženje ekstrema funkcije* f *ako kao početnu aproksimaciju uzmemo* $(1,1)$.

Zadatak 7

Neka je $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^4 - x^2 + 2xy + y^2$. *Nađite točke minimuma funkcije* f *i odredite sve točke ravnine iz kojih Newtonova metoda u jednom koraku dostiže točku minimuma.*

Napravili smo metodu najbržeg silaska (gradijentnu metodu) i Newtonovu metodu. Spomenimo i neke druge poznate iterativne metode:

- kvazi-Newtonove metode: Broydenova metoda, David-Fletcher-Powell metoda (kasnije zamijenjeno boljom Broyden–Fletcher–Goldfarb–Shanno metodom)
- Metoda konjugiranih gradijenata