

Odjel za matematiku

Sveučilište J. J. Strossmayera u Osijeku

## Nediferencijabilna optimizacija

Poslijediplomski doktorski studij matematike

### 1 Minimizacija bez ograničenja

**Definicija 1.** (Ortega and Rheinboldt, 2000) Kažemo da funkcija  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  u točki  $x^* \in D$  postiže lokalni minimum ako postoji okolina  $\mathcal{O}(x^*)$ , takva da je  $f(x) \geq f(x^*)$  za sve  $x \in \mathcal{O}(x^*) \cap D$ . Točku  $x^*$  zovemo točka lokalnog minimuma ili lokalni minimizator funkcije  $f$ .

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

Kažemo da je  $x^*$  točka globalnog minimuma (globalni minimizator) funkcije  $f$  ako je  $f(x) \geq f(x^*)$  za sve  $x \in D$ .

**Definicija 2.** (Ortega and Rheinboldt, 2000) Kažemo da funkcija  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  u točki  $x \in \text{int } D$  Gâteaux–diferencijabilna (G–diferencijabilna) ako postoji  $a \in \mathbb{R}^n$ , takav da za svaki  $h \in \mathbb{R}^n$  vrijedi

$$\lim_{t \rightarrow 0} \frac{1}{t} |f(x + th) - f(x) - ta^T h| = 0. \quad (1)$$

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

$$\lim_{t \rightarrow 0} \frac{1}{t} |f(x + te^j) - f(x) - ta_j| = 0.$$

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

Derivaciju funkcije  $f \in C^1(D)$  u točki  $x = (x_1, \dots, x_n)^T \in D$  označavat ćemo s

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

a **gradijent** funkcije  $f \in C^1(D)$  u točki  $x$  s

$$\text{grad } f(x) = \nabla f(x) = f'(x)^T,$$

a drugu derivaciju **Hessian** funkcije  $f \in C^2(D)$  u točki  $x$  s

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

Ako je  $f \in C^p(D)$ ,  $x + \alpha h \in D$  za  $\alpha \in [0, 1]$ , onda vrijedi Taylorova formula

$$f(x + h) - f(x) = f'(x)^T h + o(|h|), \quad p = 1, \quad (2)$$

$$f(x + h) - f(x) = f'(x)^T h + \frac{1}{2}(f''(x)h, h) + o(h^2), \quad p = 2, \quad (3)$$

**Definicija 3.** Kažemo da je  $x^* \in D \subseteq \mathbb{R}^n$  stacionarna točka funkcije  $f : D \rightarrow \mathbb{R}$  ako je  $\text{grad } f(x^*) = 0$ .

**Teorem 1.** (Jare and Stoer, 2004) Neka je  $D \subseteq \mathbb{R}^n$  otvoren skup. Ako je  $x^* \in D$  lokalni minimizator funkcija  $f \in C^1(D)$ , onda je  $x^*$  stacionarna točka funkcije  $f$ .

Ako je dodatno  $f \in C^2(D)$ , onda je Hessian  $H_f(x^*) \geq 0$  pozitivno semidefinitan.

Ako je  $x^* \in D$  stacionarna točka funkcije  $f \in C^2(D)$ , a Hessian  $H_f(x^*) > 0$  pozitivno definitan, onda je  $x^*$  strogi lokalni minimumizator funkcije  $f$  na  $D$ .

## 1.1 Konjugirana (dualna) funkcija konveksne diferencijabilne funkcije

Neka je  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  konveksna diferencijabilna funkcija. Tada je  $x^*$  njezin globalni minimizator onda i samo onda ako je  $f'(x^*) = 0$ , odnosno  $\text{grad } f(x^*) = 0$  (Avriel, 2003; Demjanov and Vasilev, 1981).

Za konveksnu diferencijabilnu funkciju  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  za svaki  $\xi \in \mathbb{R}^n$  kao i ranije definiramo konjugiranu (dualnu) funkciju

$$f^*(\xi) = \sup_{x \in \mathbb{R}^n} \{\xi^T x - f(x)\}. \quad (4)$$

Za dani  $\xi \in \mathbb{R}^n$  gradijent, odnosno Hessian, funkcije  $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $\psi(x) = \xi^T x - f(x)$  zadan je s

$$\text{grad } \psi(x) = \xi - \text{grad } f(x), \quad H_\psi(x) = -H_f(x).$$

Zato za dani  $\xi \in \mathbb{R}^n$

$$x^* \in \mathbb{R}^n \text{ je globalni maksimizator funkcije } \psi \Leftrightarrow \xi = \text{grad } f(x^*).$$

Primijetite da je zbog konveksnosti funkcije  $f$ ,  $H_\psi(x^*) < 0$ . Za  $\xi = \text{grad } f(x^*)$ , dualna funkcija  $f^*$  prima vrijednost

$$f^*(\xi) = x^{*T} \text{grad } f(x^*) - f(x^*).$$

Ako postoji  $(\text{grad } f)^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , onda možemo pisati  $x^* = (\text{grad } f)^{-1}(\xi)$ , a dualna funkcija  $f^*$  u točki  $\xi \in \mathbb{R}^n$  prima vrijednost

$$f^*(\xi) = \xi^T (\text{grad } f)^{-1}(\xi) - f((\text{grad } f)^{-1}(\xi)) =: \mathcal{H}_f(\xi)$$

**(Legendreova transformacija).**

**Primjer 1. (Youngova nejednakost)** Za konveksnu zatvorenu funkciju  $f : \mathbb{R} \rightarrow \mathbb{R}$  iz definicije dualne funkcije  $f^*$  slijedi

$$\xi x_0 \leq f(x_0) + f^*(\xi), \quad \xi, x_0 \in \mathbb{R} \quad (5)$$

Kako je za  $f(x) = \frac{1}{p}|x|^p$ ,  $1 < p < +\infty$ , pripadna dualna funkcija  $f^*(\xi) = \frac{1}{q}|\xi|^q$ , gdje je  $\frac{1}{p} + \frac{1}{q} = 1$ , na osnovi (5) za  $a, b \geq 0$  vrijedi

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Specijalno za  $p = q = 2$  dobivamo

$$ab \leq \frac{a^2}{2} + \frac{b^2}{2}.$$

**Primjer 2. (Theory of large deviations in probability theory)** Za  $f(x) = e^x - 1$ , dualna funkcija je  $f^*(\xi) = 1 - \xi + \xi \ln \xi$ , pa na osnovi Youngove nejednakosti za  $a, b \geq 0$  dobivamo

$$ab \leq e^a - b + b \ln b.$$

**Zadatak 1.** Ustanovite i dokažite kada u nejednakostima iz Primjera 1 i Primjera 2 vrijedi jednakost.

**Zadatak 2.** Odredite Legendreovu transformaciju (dualnu funkciju) funkcije  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f(x_1, x_2) = x_1^2 + x_2^2$ . Koristeći dobiveni rezultat odredite također dualne funkcije sljedećih funkcija:

- |                                                       |                                                 |
|-------------------------------------------------------|-------------------------------------------------|
| (a) $\varphi(x) = f(x) + \beta, \beta \in \mathbb{R}$ | (b) $\varphi(x) = f(x + y), y \in \mathbb{R}^2$ |
| (c) $\varphi(x) = f(\gamma x), \gamma \neq 0$         | (d) $\varphi(x) = kf(x), k > 0$                 |

**Zadatak 3.** Odredite Legendreovu transformaciju (dualnu funkciju) funkcije  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{2}x^T A x$ , gdje je  $A$  simetrična nesesingularna matrica.

## 2 Metode spusta

Promatramo iterativne metode oblika

$$x^{k+1} = x^k + \alpha_k p^k, \quad (6)$$

gdje je  $p^k \in \mathbb{R}^n$  vektor smjera kretanja, a  $\alpha_k > 0$  duljina koraka u smjeru  $p^k$ .

Primijetite da ako vektor smjera kretanja normiramo  $\|p^k\| = 1$ , onda  $\alpha_k$  predstavlja udaljenost točaka  $x^{k+1}$  i  $x^k$ .

**Lema 1.** (*Ortega and Rheinboldt, 2000*) Neka je funkcional  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $G$ -diferencijabilan u nekoj točki  $x \in \text{int}(D)$  i neka je  $f'(x)p < 0$  za neki  $p \in \mathbb{R}^n$ . Tada postoji  $\delta > 0$ , takav da je

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

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

$$\varphi'(0) = \lim_{\alpha \rightarrow 0} \frac{\varphi(\alpha) - \varphi(0)}{\alpha} = \lim_{\alpha \rightarrow 0} \frac{f(x + \alpha p) - f(x)}{\alpha} = f'(x)p.$$

Kako je  $f'(x)p < 0$  postoji  $\delta > 0$ , takav da je  $x + \alpha p \in D$  i  $\frac{1}{\alpha}(f(x + \alpha p) - f(x)) < 0$  za sve  $\alpha \in (0, \delta)$ .  $\square$

(a) **Gradijentna metoda:**  $p^k := -\text{grad } f(x^k)$ . Vrijedi

$$f'(x^k)p^k = -f'(x^k)\text{grad } f(x^k) = -\|\text{grad } f(x^k)\|^2 < 0.$$

(b) **Pokoordinatna relaksacija:** izaberemo onaj bazni vektor  $e^i \in \{e^1, \dots, e^n\}$  za koji je

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

Primijetite da je  $f'(x^k)e^i = (f'(x^k))_i$ . Vektor smjera kretanja  $p^k$  izabrat ćemo između  $\{e^i, -e^i\}$  za koji je

$$f'(x^k)p^k < 0.$$

Umjesto ortonormiranih baznih vektora  $\{e^1, \dots, e^n\}$  možemo izabrati neke druge bazne vektore iz  $\mathbb{R}^n$ .

(c) **Newtonova metoda:**  $p^k := -[f''(x^k)]^{-1} \text{grad } f(x^k)$ . Vrijedi

$$f'(x_k)p^k = -f'(x^k)[f''(x^k)]^{-1}f'(x^k)^T,$$

što je negativno ako je  $f''(x^k)$  pozitivno definitna matrica (tada je i  $[f''(x^k)]^{(-1)}$  pozitivno definitna).

(c) **Kvazi-Newtonova metoda:**  $p^k := -B^{-1} \text{grad } f(x^k)$ , gdje je  $B > 0$  pozitivno definitna matrica koja “nalikuje” na  $f''(x^k)$

$$f'(x_k)p^k = -f'(x^k)B^{-1}f'(x^k)^T < 0.$$

(Broyden, Davidon–Fletcher–Powell (DFP), Broyden–Fletcher–Goldfarb–Shano (BFGS)).

### 3 Metode izbora duljine koraka

Prema *Lemi 1*, ako je vektor smjera kretanja iz točke  $x^k$  u točku  $x^k$  dobro definiran ( $f'(x^k)p^k < 0$ ), onda se u iterativnom procesu (6)

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

uvijek može odabrati duljina koraka  $\alpha_k > 0$ , tako da bude  $f(x^{k+1}) < f(x^k)$  (Bonnans et al., 2006; Ortega and Rheinboldt, 2000).

#### 3.1 Metoda najbržeg spusta

Optimalni izbor duljine koraka dobiven postupkom jednodimenzionalne minimizacije

$$f(x^k + \alpha_k p^k) = \min_{\alpha > 0} \{f(x^k + \alpha p^k) : x^k + \alpha p^k \in D\}, \quad (7)$$

potjeće još od Cauchyja, a nazivamo ga Metoda najbržeg spusta.

#### 3.2 Princip H. Curry–M. Altman

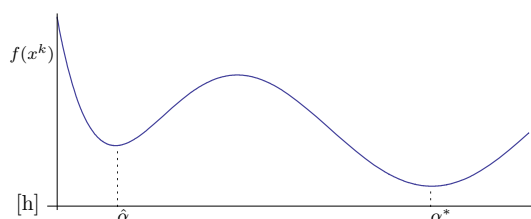
Definiramo funkciju  $\varphi : [0, +\infty) \rightarrow \mathbb{R}$ ,

$$\varphi(\alpha) = f(x^k + \alpha p^k), \quad (8)$$

i potražimo njene stacionarne točke rješavanjem jednadžbe po  $\alpha$

$$f'(x^k + \alpha p^k) p^k = 0. \quad (9)$$

Ako je  $f$  strogo konveksan na  $D$ , onda jednadžba (9) ima jedinstveno rješenje koje se podudara s onim dobivenim u (7). Primijetite da zbog uvjeta  $f'(x^k)p^k < 0$ , rješenje  $\alpha = 0$  nije prihvatljivo. Općenito jednadžba (9) može imati više rješenja.



Slika 1: Princip Curry–Altman

Prema Principu Curry ako je  $f'(x^k)p^k < 0$  za duljinu koraka  $\alpha_k$  treba izabrati najmanji pozitivni korijen  $\hat{\alpha}$  jednadžbe (9). Naime, za  $\alpha \in (0, \hat{\alpha})$  vrijedi

$$\varphi'(\alpha) = f'(x^k + \alpha p^k)p^k < 0 \quad \forall \alpha \in (0, \hat{\alpha}).$$

Nadalje, prema teoremu o srednjoj vrijednosti<sup>1</sup>

$$f(x^k + \hat{\alpha}p^k) - f(x^k) = \hat{\alpha} \int_0^1 f'(x^k + \hat{\alpha}tp^k)p^k dt < 0,$$

zbog  $f'(x^k + \alpha p^k)p^k < 0$  za sve  $\alpha \in [0, \hat{\alpha})$ .

Prema Principu M. Altman za neki fiksni  $\mu \in [0, 1]$  za  $\alpha_k$  treba izabrati najmanji pozitivni korijen jednadžbe

$$[f'(x^k + \alpha p^k) - \mu f'(x^k)]p^k = 0. \quad (10)$$

Ako je  $\mu = 0$ , princip se svodi na princip Curry. Ako je  $\mu > 0$ , duljina koraka  $\alpha_k$  je prvi pozitivni broj  $\alpha > 0$  za kojeg je  $f'(x^k + \alpha p^k)p^k$  jednak  $\mu$ -tom dijelu svoje početne veličine  $f'(x^k)p^k$ . Princip Altmana ima praktično značenje, ali i važno teorijsko značenje jer se može iskoristiti kod ispitivanja drugih metoda izbora duljine koraka.

### 3.3 Aproksimativna minimizacija i traženje nultočka

I kod metode najbržeg spusta i prilikom implementacije principa Curry–Altmana nemamo točno, već aproksimativno rješenje dobiveno nekim jednodimenzionalnim iterativnim postupkom.

Prema (3) za svaki  $x \in D$  vrijedi

$$f(x) = f(x^k) + f'(x^k)(x - x^k) + \frac{1}{2}(x - x^k)^T H_f(x^k)(x - x^k) + o(\|x - x^k\|^2)$$

Zato sljedeći kvadratni funkcional dobro aproksimira funkcional  $f$  u okolini točke  $x^*$

$$f_k(x) = f(x^k) + f'(x^k)(x - x^k) + \frac{1}{2}(x - x^k)^T H_f(x^k)(x - x^k). \quad (11)$$

Primjenjujući ovu formulu na funkciju (8) dobivamo

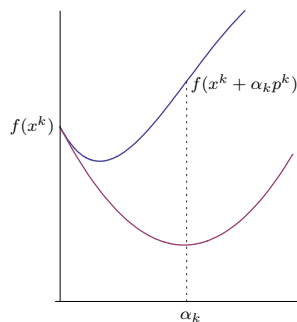
$$\varphi(\alpha) = f(x^k + \alpha p^k) \approx f(x^k) + \alpha f'(x^k)p^k + \frac{1}{2}\alpha^2(p^k)^T H_f(x^k)p^k. \quad (12)$$

Ako je  $(p^k)^T H_f(x^k)p^k > 0$ , onda je  $\alpha_k$  korektno određen iz formule

$$\alpha_k = \frac{-f'(x^k)p^k}{(p^k)^T H_f(x^k)p^k} > 0. \quad (13)$$

---

<sup>1</sup>Ako je  $F'(x + t(y - x))$  neprekidna po  $t$ , vrijedi  $F(y) - F(x) = \int_0^1 F'(x + t(y - x))(y - x)dt$ .



Slika 2: Aproksimacija minimuma

Budući da je veličine  $f'(x^k)$ ,  $H_f(x^k)$  često teško izračunati, umjesto minimzacije funkcije (12), možemo definirati interpolacijski polinom drugog reda, koji se s funkcijom (8) podudara u tri točke:  $u_1, u_2, u_3$  (obično se uzima  $u_1 = 0$ ) i  $\alpha_k$  određuje kao točka minimuma ovog interpolacijskog polinoma.

Kao što se vidi na *Slici 2*, vrijednost duljine koraka (13) ne mora osiguravati smjer spusta. Ali zato prema *Lemi 1* možemo izabrati parametar  $\omega_k > 0$ , tako da bude

$$f(x^k + \omega_k \alpha_k p^k) < f(x^k).$$

**Zadatak 4.** (*Ortega and Rheinboldt (2000), str. 254*) Pokažite da interpolacijski polinom drugog reda koji se s funkcijom  $\varphi(\alpha) = f(x + \alpha p)$  podudara u tri različite točke  $u_1, u_2, u_3$  ima oblik

$$\psi(\alpha) = \frac{(\alpha - u_2)(\alpha - u_3)}{(u_1 - u_2)(u_1 - u_3)} f_1 + \frac{(\alpha - u_1)(\alpha - u_3)}{(u_2 - u_1)(u_2 - u_3)} f_2 + \frac{(\alpha - u_1)(\alpha - u_2)}{(u_3 - u_1)(u_3 - u_2)} f_3,$$

gdje je  $f_i = f(x + u_i p)$ ,  $i = 1, 2, 3$ . Pokažite također da  $\psi$  postiže lokalni minimum u točki

$$\bar{\alpha} = \frac{1}{2} - \frac{(u_2^2 - u_3^2)f_1 + (u_3^2 - u_1^2)f_2 + (u_1^2 - u_2^2)f_3}{(u_2 - u_3)f_1 + (u_3 - u_1)f_2 + (u_1 - u_2)f_3}$$

ako je

$$\frac{2f_1}{(u_1 - u_2)(u_1 - u_3)} + \frac{2f_2}{(u_2 - u_1)(u_2 - u_3)} + \frac{2f_3}{(u_3 - u_1)(u_3 - u_2)} > 0.$$

Napišite izraze za  $\psi$  i  $\bar{\alpha}$  za specijalne slučajeve: (a)  $u_1 = 0, u_3 = 2u_2$ , (b)  $u_2 = 0, u_1 = -u_3$ . Testirajte ovaj princip na nekoliko primjera uz primjenu gradijentne metode.

### 3.4 Princip majoriziranja

Situacija prikazana na *Slici ??* motivacija je za definiranje princip majoriziranja (Ortega and Rheinboldt, 2000):

ako funkcija  $\psi : [0, \hat{\alpha}] \rightarrow \mathbb{R}$  zadovoljava

$$f(x^k + \alpha_k p^k) \leq \psi(\alpha) < f(x^k), \quad \forall \alpha \in (0, \hat{\alpha}), \quad (14)$$

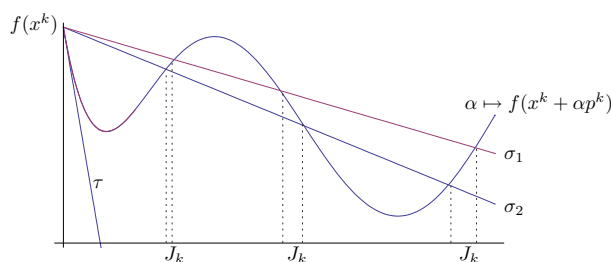
onda svaki izbor duljine koraka  $\alpha_k \in (0, \hat{\alpha})$  smanjuje vrijednost funkcionala  $f$ .

### 3.5 Princip A. Goldsteina

Neka je  $f$  diferencijabilan u točki  $x^k \in \text{int } D$  i neka je  $f'(x^k)p^k < 0$ . Tangenta na graf funkcije  $\varphi(\alpha) = f(x^k + \alpha p^k)$  u točki  $\alpha = 0$  zadana je s

$$\tau(\alpha) = f(x^k) + \alpha f'(x^k)p^k,$$

i ima negativni koeficijent smjera  $f'(x^k)p^k < 0$  (vidi *Sliku 3*).



Slika 3: Princip Goldsteina

Definirajmo dvije sekante tako da prolaze točkom  $(0, f(x^k))$ , a za  $\alpha > 0$  nalaze se iznad tangente. Ako izaberimo  $0 < \mu_1 \leq \mu_2 < 1$ , onda su moguće sekante zadane sa

$$\sigma_i(\alpha) = f(x^k) + \mu_i \alpha f'(x^k)p^k, \quad i = 1, 2$$

i za njih vrijedi

$$\tau(\alpha) < \sigma_2(\alpha) \leq \sigma_1(\alpha), \quad \forall \alpha > 0.$$

Naime zbog  $f'(x^k)p^k < 0$  i  $0 < \mu_1 \leq \mu_2 < 1$  za koeficijente smjerova vrijedi

$$\mu_1 f'(x^k)p^k \geq \mu_2 f'(x^k)p^k > f'(x^k)p^k.$$

Definirajmo sada skup

$$J_k = \{\alpha > 0 : [x^k, x^k + \alpha p^k] \subset D, \quad \sigma_2(\alpha) \leq f(x^k + \alpha p^k) \leq \sigma_1(\alpha)\}$$

Primijetite da za  $\alpha \in J_k$  zbog  $-\sigma_1(\alpha) \leq -f(x^k + \alpha p^k) \leq -\sigma_2(\alpha)$  vrijedi

$$0 < -\mu_1 \alpha f'(x^k)p^k \leq f(x^k) - f(x^k + \alpha p^k) \leq -\mu_2 \alpha f'(x^k)p^k.$$

Specijalno lijevi dio nejednakosti pokazuje da se za proizvoljni izbor  $\alpha \in J_k$  postiže sniženje vrijednosti minimizirajućeg funkcionala  $f$ .

Primijetite da se skup  $J_k$  smanjuje kako se razlika  $\mu_2 - \mu_1$  smanjuje, a specijalno za  $\mu_1 = \mu_2 =: \mu$  sekante se podudaraju, a  $\alpha_k$  zadovoljava jednakost

$$f(x^k + \alpha_k p^k) - \mu \alpha_k f'(x^k)p^k = f(x^k).$$

Prvobitno je Goldsteinov princip predložen u sljedećoj formi: Neka su  $0 < \mu_1 \leq \mu_2 < 1$  fiksne konstante. Ako je za neki  $\alpha_k > 0$ ,

$$f(x^k) - f(x^k + \alpha_k p^k) \geq -\mu_1 \alpha_k f'(x^k) p^k. \quad (15)$$

onda kao sljedeću aproksimaciju biramo točku  $x^{k+1} = x^k + \alpha_k p^k$ . Ako (15) nije ispunjeno, izaberemo parametar  $\omega_k$  tako da bude

$$0 < -\mu_1 \omega_k \alpha_k f'(x^k) p^k \leq f(x^k) - f(x^k + \omega_k \alpha_k p^k) \leq -\mu_2 \omega_k \alpha_k f'(x^k) p^k. \quad (16)$$

i kao sljedeću aproksimaciju biramo točku  $x^{k+1} = x^k + \omega_k \alpha_k p^k$ .

Ovakav parametar  $\omega_k$  uvijek se može izabrati. Naime, ako (15) nije ispunjeno, tj. ako za neki  $\mu_1 \in (0, 1)$  vrijedi

$$f(x^k) - f(x^k + \alpha_k p^k) < -\mu_1 \alpha_k f'(x^k) p^k, \quad (17)$$

definiramo funkciju  $\psi : [0, 1] \rightarrow \mathbb{R}$ ,  $\psi(0) = 1$ ,

$$\psi(\omega) = \frac{f(x^k) - f(x^k + \omega \alpha_k p^k)}{-\omega \alpha_k f'(x^k) p^k}, \quad \omega \in (0, 1].$$

Funkcija  $\psi$  je neprekidna na  $[0, 1]$  jer primjenom L'Hôpitalovog pravila dobivamo  $\lim_{\omega \rightarrow 0} \psi(\omega) = 1$ . Zbog (17),  $\psi(1) < \mu_1$ , a tada zbog neprekidnosti funkcije  $\psi$  ona na  $[0, 1]$  prima sve vrijednosti između  $\mu_1$  i 1. Kako je  $0 < \mu_1 \leq \mu_2 < 1$ , postoji  $\omega_k \in (0, 1)$ , takav da bude  $\mu_1 \leq \psi(\omega_k) \leq \mu_2$ , a to je nejednakost (16).

*Primjedba 1.* Brojeve  $\omega_k$  moguće je birati kao članove niza  $\left(\frac{1}{q^j}\right)$ ,  $q > 1$  i o tome postoji posebno razrađena analiza.

**Zadatak 5.** Neka je funkcional  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $G$ -diferencijabilan u nekoj točki  $x^k \in \text{int}(D)$  i neka je  $f'(x^k) p^k < 0$  za neki  $p^k \in \mathbb{R}^n$ .

Neka je nadalje  $\varphi(\alpha) = f(x^k + \alpha p^k)$  i  $t_1 > 0$ . Poznavajući

$$q_0 = \varphi(0), \quad q'_0 = \varphi'(0), \quad q_1 = \varphi(t_1),$$

(uočite da je  $q'_0 < 0$ ) odredite kvadratnu funkciju  $q(t) = at^2 + bt + c$ , tako da je

$$q(0) = q_0, \quad q'(0) = q'_0, \quad q(t_1) = q_1.$$

Uz koji uvjet se postiže minimum dobivene kvadratne funkcije  $q$ ? Odredite minimizator kvadratne funkcije  $q$ . Poznavajući ovu kvadratnu funkciju konstruirajte algoritam odabira duljine koraka  $\alpha_k$ .

## Literatura

- M. Alić, G. Nogo, *Optimizacija: Uvod u teoriju nužnih i dovoljnih uvjeta ekstrema*, Odjel za matematiku, Sveučilište u Osijeku, Osijek, 2004.
- M. Avriel, *Nonlinear Programming: Analysis and Methods 2ed.*, Dover Publications, Inc. Mineola, New York, 2003.
- M. S. Bazaraa, H. D. Sherali, C. M. Shetty, *Nonlinear Programming. Theory and Algorithms. 3<sup>rd</sup> Edition*, Wiley, New Jersey, 2006.
- J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, C. A. Sagastizábal, *Numerical Optimization. Theoretical and Practical Aspects*, Springer-Verlag, Berlin, 2006.
- S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, 2004.
- В. Ф. Дем'янов, Ф. П. Василев, Недифференцируемая оптимизация, Наука, Москва, 1981.
- J. E. Dennis, Jr, R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, 1996.
- J. E. Dennis Jr., J. J. More, *Quasi-Newton methods, motivation and theory*, SIAM Review, **19**(1977), 46-89.
- W. Fenchel, *On conjugate convex function*, Canadian Journal of Mathematics **1**(1949) 73-77
- P. E. Gill, W. Murray and M. H. Wright, *Practical Optimization*, Academic Press, 1981.
- J. B. Hiriart-Urruty, C. Lemaréchal, *Fundamentals of Convex Analysis*, Springer-Verlag, Berlin - Heidelberg, 2001.
- F. Jare, J. Stoer, *Optimierung*, Springer-Verlag, Berlin, 2004
- Ж. П. Обен, Нелинейный анализ и его экономические приложения, Мир, Москва, 1988.
- J. M. ORTEGA, W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, SIAM, Philadelphia, 2000. (postoji i ruski prijevod)
- A. Ruszczynski, *Nonlinear Optimization*, Princeton University Press, Princeton and Oxford, 2006

N.Z.Shor, *Nondifferentiable Optimization and Polynomial Problems*, Kluwer, London, 1998.

Р. Г. Стронгин, Численные методы в многоэкстремальных задачах, Наука, Москва, 1978.

Ф. П. Василев, Лекции по методам решения экстремальных задач, Издательство Московского университета, Москва, 1974.

S. Zlobec, J. Petrić, *Nelinearno programiranje*, Naučna knjiga, Beograd, 1989.