Sveučilište Josipa Jurja Strossmayera u Osijeku
Odjel za matematiku

Metode optimizacije

Rudolf Scitovski Ninoslav Truhar Zoran Tomljanović

Metode optimizacije

Istina je previše složena, stoga ju jedino možemo aproksimirati.

John von Neumann

Rudolf Scitovski

Ninoslav Truhar

Zoran Tomljanović

Sveučilište Josipa Jurja Strossmayera u Osijeku
Odjel za matematiku, 2014.

Dr.sc. Rudolf Scitovski, redoviti profesor u trajnom zvanju
e-mail: scitowsk@mathos.hr

Dr.sc. Ninoslav Truhar, redoviti profesor
e-mail: ntruhar@mathos.hr

Dr.sc. Zoran Tomljanović, docent
e-mail: ztomljan@mathos.hr

Sveučilište Josipa Jurja Strossmayera u Osijeku - Odjel za matematiku, Trg Ljudevita Gaja 6, HR-31000 Osijek

Izdavač:

Sveučilište Josipa Jurja Strossmayera u Osijeku - Odjel za matematiku

Recenzenti:

dr.sc. Miljenko Marušić, izvanredni profesor,
PMF - Matematički odsjek, Sveučilište u Zagrebu
dr.sc. Marko Vrdoljak, izvanredni profesor,
PMF - Matematički odsjek, Sveučilište u Zagrebu

Lektor:

Ivanka Ferčec, Elektrotehnički fakultet, Osijek

CIP zapis dostupan u računalnom katalogu Gradske i sveučilišne knjižnice Osijek pod brojem 131006006.

ISBN 978-953-6931-65-1

Ovaj udžbenik objavljuje se uz suglasnost Senata Sveučilišta Josipa Jurja Strossmayera u Osijeku pod brojem 7/14.
Ovaj udžbenik objavljuje se uz financijsku pomoć Ministarstva znanosti, obrazovanja i sporta Republike Hrvatske.

Tisak:
Grafika d.o.o.
Osijek

Sadržaj

POGLAVLJE 1

Uvod

Ovaj je tekst prvenstveno namijenjen studentima Sveučilišta Josipa Jurja Strossmayera u Osijeku - Odjela za matematiku za potrebe predmeta Metode optimizacije M028 koji se sluša u 4. semestru diplomskog studija. Namjena mu je upoznati studente s glavnim metodama jednodimenzionalne i višedimenzionalne minimizacije sa i bez ograničenja. Posebna je pozornost posvećena metodama minimizacije nediferencijabilnih funkcija. Pri tome se izbjegavalo "strogo" dokazivanje teorema, osim u slučajevima konstruktivnih dokaza koji sami po sebi upućuju na izgradnju ideja ili metoda.

Pretpostavimo da je zadana funkcija $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$. U okviru ovog teksta promatramo sljedeće probleme:

1. Odrediti $f^\star = \inf\limits_{x\in D} f(x)$;
2. Ako funkcija f postiže svoj infimum na D, odrediti $\operatorname*{argmin}\limits_{x\in D} f(x)$, tj. odrediti točku $x^\star \in D$ u kojoj funkcija f postiže svoj infimum;
3. Ako se $\inf\limits_{x\in D} f(x)$ ne postiže na D, odrediti niz $(x_n) \subset D$ za koji je $\lim\limits_{n\to\infty} f(x_n) = f^\star$.

Napomenimo da oznaka $\operatorname*{argmin}\limits_{x\in D} f(x)$ predstavlja skup točaka u kojima dana funkcija postiže svoju minimalnu vrijednost.

Primijetite da u općem slučaju $\operatorname*{argmin}\limits_{x\in D} f(x)$ ne mora biti jedna točka, već više točaka ili čitav podskup u D.

Primijetite također da je dovoljno razmatrati problem traženja lokalnog odnosno globalnog minimuma neke funkcije $f\colon D \to \mathbb{R}$, jer vrijedi

$$\operatorname*{argmin}_{x\in D} f(x) = \operatorname*{argmax}_{x\in D}(-f(x)).$$

Navedeni optimizacijski problemi imaju veliku primjenu u raznim dijelovima života. Samo kao ilustraciju navest ćemo nekoliko disciplina u kojima se pojavljuju takvi problemi. Na primjer, problem oblikovanja određenih mehaničkih objekata (oblikovanje dijelova automobilskih motora, stupova za dalekovode ili metalnih konstrukcija u visokogradnji, itd.) može se svesti na optimizacijski problem koji nazivamo optimizacija oblikovanja ili engleski *design optimization*.

Kao drugi primjer mogli bismo navesti ekonomiju koja uključuje tradicionalne optimizacijske teorije, ali također i neke aspekte teorije igara te proučavanje ekonomskih ravnotežnih stanja. 70-ih godina prošlog stoljeća ekonomisti su za modeliranje tržišta ili drugih ekonomskih pojmova počeli koristiti teoriju kontrole koja opisuje ponašanje dinamičkih sustava. Na primjer, u mikroekonomiji se u svrhu modeliranja ponašanja tržišta rada koriste modeli dinamičkog pretraživanja. Još jedan primjer upotrebe optimizacijskih metoda može se naći u okviru operacijskih istraživanja.

Iako smo na početku uvoda naveli da je ovaj tekst namijenjen studentima Sveučilišta Josipa Jurja Strossmayera u Osijeku - Odjela za matematiku, upravo činjenica da se razni problemi optimizicije pojavljuju u velikom broju u raznim dijelovima ljudske djelatnosti osigurava ovom tekstu širu primjenu.

Knjiga se sastoji od sljedećih šest poglavlja. U poglavlju 1. *Konveksni skupovi i konveksne funkcije* obrađuju se osnovni pojmovi vezani za konveksne skupove i funkcije, lokalni minimumi i općenito problemi ekstrema. U 2. poglavlju *Jednodimenzionalna minimizacija* proučavaju se metode jednodimenzionalne minimizacije kao što su metoda polovljenja, metoda zlatnog reza, metoda parabole, Brentova i Newtonova metoda. U poglavlju 3. *Problemi višedimenzionalne minimizacije* proučavaju se neke od metoda višedimenzionalne minimizacije, npr. metode izbora duljine koraka, i metoda najbržeg spusta, te se posebna pažnja posvećuje izboru smjerova optimizacije. Stoga se posebno proučava načelo H. Curry–M. Altmana, aproksimativna minimizacija i traženje nultočaka te načelo A. Goldsteina. Nadalje, u okviru ovog poglavlja posebno se proučavaju gradijentna metoda, metoda najbržeg spusta i minimizacija pomoću Newtonove metode. Poglavlje 5. posvećeno je problemu uvjetne minimizacije u okviru kojega se bavi nužnim i dovoljnim uvjetima za optimalnost, gradijentnom metodom i Newtonovom metodom s projekcijom. U poglavlju 6. proučavaju se neke osnovne metode direktnog traženja kao Nelder-Meadov i MDS algoritam te Powellova metoda u više dimenzija.

POGLAVLJE 2

Konveksni skupovi i konveksne funkcije

Konveksnost igra vrlo značajnu ulogu pri numeričkoj optimizaciji. Stoga ćemo ovaj tekst započeti proučavanjem pojma konveksnosti skupova.

2.1 Konveksni skupovi

Definicija 1. Kažemo da je skup $D \subseteq \mathbb{R}^n$ *konveksan* ako on za bilo koje dvije točke $x_1, x_2 \in D$ također sadrži i segment određen tim točkama, tj.

$$x_1, x_2 \in D \Rightarrow \lambda x_1 + (1-\lambda)x_2 \in D \quad \forall \lambda \in [0,1].$$

Ako je $D \subseteq \mathbb{R}^n$ neki skup, onda najmanji konveksni skup koji sadržava skup D zovemo *konveksnom ljuskom*[1] skupa D. Primjerice, ako se skup D sastoji od tri različite točke iz $\mathbb{R}^n$, onda je njegova konveksna ljuska trokut s vrhovima u tim točkama. Može se pokazati da vrijedi:

- Konveksan skup $D \subseteq \mathbb{R}^n$ sadrži svaku konveksnu kombinaciju od konačno svojih točaka, tj.

$$(x_1, \ldots, x_m \in D) \wedge \left(\lambda_1, \ldots, \lambda_m \geq 0, \quad \sum_{i=1}^{m} \lambda_i = 1\right) \Longrightarrow \sum_{i=1}^{m} \lambda_i x_i \in D;$$

[1] eng.: convex hull

- Skup svih konveksnih kombinacija točaka nekog skupa $D \subseteq \mathbb{R}^n$ je konveksna ljuska tog skupa.

Definicija 2. Kažemo da je skup $D \subseteq \mathbb{R}^n$ *strogo konveksan* ako

$$x_1, x_2 \in D, \quad x_1 \neq x_2 \Rightarrow \lambda x_1 + (1-\lambda)x_2 \in \mathrm{I}nt\, D \quad \forall \lambda \in (0,1).$$

Zadatak 1. Dokažite sljedeće tvrdnje:

1. Presjek konveksnih skupova je konveksan skup. Mora li unija dva konveksna skupa biti konveksan skup ?
2. Ako su A, B konveksni skupovi, onda su i $\lambda\, A$, $A+B$, $A-B$ konveksni skupovi.

Zadatak 2. Provjerite konveksnost sljedećih skupova:

a) $\overline{D(x_0, r)} = \{x \in \mathbb{R}^n \colon \|x - x_0\| \leq r\}$ (zatvorena kugla)

b) $\Gamma(a, l) = \{x \in \mathbb{R}^n \colon (a, x) = l\}$ (hiperravnina)

c) $D = \{(x,y,z) \in \mathbb{R}^3 \colon x + y \geq 1,\ 2x + 3y - 2z < 5\}$

d) $D = \{(x,y) \in \mathbb{R}^2 \colon y - |x^2 - x - 6| > 2\}$

Zadatak 3. Neka je $f \colon \mathbb{R}^n \to \mathbb{R}^m$ afina funkcija $f(x) = Ax + b$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ i D konveksan skup. Pokažite da je tada i $f(D) \subseteq \mathbb{R}^m$ također konveksan skup.

2.2 Konveksne funkcije

Definicija 3. Kažemo da je funkcija $f \colon \mathbb{R}^n \to \mathbb{R}$ *konveksna* ako vrijedi

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \tag{2.1}$$

$\forall x, y \in \mathbb{R}^n$ i $\forall \lambda \in [0,1]$.

Primjer 1. Navedimo neka elementarna svojstva konveksnih funkcija [2, 3]:

- ako su $f_1, \ldots, f_m$ konveksne funkcije i $\lambda_1, \ldots, \lambda_m \geq 0$, onda je $\sum\limits_{i=1}^{m} \lambda_i f_i$ konveksna funkcija;
- ako je f konveksna na $\mathbb{R}^n$, Ψ neopadajuća i konveksna na $\mathbb{R}$, onda je $\Psi \circ f$ konveksna na $\mathbb{R}^n$;

- ako su $f_1, \ldots, f_m$ konveksne funkcije, onda je $\sup_i f_i(x)$ konveksna funkcija;
- ako je $f \colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ konveksna funkcija na konveksnom skupu D, onda je ona neprekidna na $\operatorname{Int} D$.

Definicija 4. Kažemo da je funkcija $f \colon \mathbb{R}^n \to \mathbb{R}$ *strogo konveksna* ako $\forall x, y \in \mathbb{R}^n, \quad x \neq y$ vrijedi

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y), \quad \forall \lambda \in (0,1). \tag{2.2}$$

Definicija 5. Kažemo da je funkcija $f \colon \mathbb{R}^n \to \mathbb{R}$ *jako konveksna*[2] ako postoji takav realni broj $\gamma > 0$ da vrijedi

$$f\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(f(x) + f(y)) - \gamma \|x-y\|^2, \quad \forall x, y \in \mathbb{R}^n. \tag{2.3}$$

Primjedba 1. Može se pokazati da je funkcija $f \colon \mathbb{R}^n \to \mathbb{R}$ jako konveksna ako i samo ako postoji realan broj $\kappa > 0$ tako da bude

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) - \kappa \lambda (1-\lambda) \|x-y\|^2, \tag{2.4}$$
$$\forall x, y \in \mathbb{R}^n, \forall \lambda \in [0,1].$$

Naime, jedan smjer je jednostavan jer ako u (2.4) stavimo $\lambda = \frac{1}{2}$ i $\gamma = \frac{1}{4}\kappa$, dobivamo (2.3).

Primjedba 2. Pojam konveksnosti može se na isti način definirati ako je domena funkcije konveksan podskup od $\mathbb{R}^n$.

Primjer 2. Neka je $\mathcal{A} \colon \mathbb{R}^n \to \mathbb{R}^n$ simetrični linearni operator. Tada je kvadratna forma $f \colon \mathbb{R}^n \to \mathbb{R}_+, f(x) := \frac{1}{2}(\mathcal{A}x, x)$ konveksna funkcija onda i samo onda ako je operator $\mathcal{A}$ semidefinitan ($\mathcal{A} \geq 0$), tj. ako su sve njegove svojstvene vrijednosti nenegativne: označimo ih s $\lambda_1 \geq \cdots \lambda_n \geq 0$. Kako je (vidi primjerice [14], [15])

$$\lambda_n \|x\|^2 \leq (\mathcal{A}x, x) \leq \lambda_1 \|x\|^2 \text{ za sve } x \in \mathbb{R}^n, \tag{2.5}$$

može se pokazati da vrijedi

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) - \frac{1}{2}\lambda_n \lambda (1-\lambda) \|x-y\|^2. \tag{2.6}$$

Dakle, ako je $\mathcal{A}$ pozitivno definitni linearni operator ($\lambda_n > 0$), funkcija f je jako konveksna funkcija.

[2] eng.: strictly convex function = strogo konveksna funkcija ; strongly convex function = jako konveksna funkcija

Pokažimo da iz (2.5) slijedi (2.6). Iz definicije kvadratne forme $f(x)$ slijedi da je

$$\begin{aligned} f(\lambda x+(1-\lambda)y) &= \tfrac{1}{2}(\mathcal{A}(\lambda x+(1-\lambda)y),\lambda x+(1-\lambda)y) \\ &= \tfrac{1}{2}\left(\lambda^2(\mathcal{A}x,x)+2\lambda(1-\lambda)(\mathcal{A}y,x)+(1-\lambda)^2(\mathcal{A}y,y)\right) \\ &= \lambda f(x)+(1-\lambda)f(y)+\tfrac{1}{2}\left((\lambda^2-\lambda)(\mathcal{A}x,x)+\right. \\ &\quad \left. 2\lambda(1-\lambda)(\mathcal{A}y,x)+((1-\lambda)^2-(1-\lambda))(\mathcal{A}y,y)\right). \end{aligned}$$

Dakle, vidimo da vrijedi

$$f(\lambda x+(1-\lambda)y)=\lambda f(x)+(1-\lambda)f(y)-\tfrac{\lambda}{2}(1-\lambda)\left((\mathcal{A}(x-y),(x-y))\right),$$

što zajedno sa (2.5) povlači (2.6).

Navedimo još neke primjere konveksnih odnosno jako konveksnih funkcija.

Primjer 3. Funkcija $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x_1,x_2)=x_1^2+x_2^2$ je jako konveksna funkcija, dok je funkcija $g\colon \mathbb{R}^2 \to \mathbb{R}$, $g(x_1,x_2)=x_1^2$ konveksna, ali nije jako konveksna. Napišite matrice kvadratnih formi f i g i definirajte odgovarajuće linearne operatore. Što su njihove svojstvene vrijednosti? Kako izgledaju nivo plohe ovih funkcija?

Zadatak 4. Pokažite da je svaka jako konveksna funkcija ujedno i konveksna, ali da obrat ne vrijedi.

Zadatak 5. Provjerite jesu li sljedeće funkcije konveksne, strogo konveksne ili jako konveksne funkcije:

a) $f(x)=3x-4$,

b) $f(x)=6-5x-2x^2+x^3$,

c) $f(x)=3x^2+2x-3$,

d) $f(x)=e^{-x+1}+5$.

Rješenje: konveksne su funkcije pod $a), c), d)$, strogo konveksne su pod $c), d)$, a jako je konveksna funkcija pod $c)$.

Zadatak 6. Neka je f konveksna funkcija, $f\colon \mathcal{D} \to \mathbb{R}$, gdje je $\mathcal{D} \subseteq \mathbb{R}^n$ konveksan skup. Tada za svaki izbor točaka $x_1,x_2,\ldots,x_m \in \mathcal{D}$ i skalara $\alpha_1,\ldots,\alpha_m \geq 0$, $\sum\limits_{i=1}^{m}\alpha_i=1$ vrijedi

$$f\left(\sum_{i=1}^{m}\alpha_i x_i\right) \leq \sum_{i=1}^{m}\alpha_i f(x_i) \quad \text{(Jensenova nejednakost).}$$

Nadalje, koristeći Jensenovu nejednakost dokažite da vrijedi

$$\left(\sum_{i=1}^{m} \alpha_i x_i\right) \left(\sum_{i=1}^{m} \frac{\alpha_i}{x_i}\right) \geq 1 \quad \text{za} \quad x_1, \ldots, x_m > 0.$$

Zadatak 7. Neka su $g, h \colon \mathcal{D} \to \mathbb{R}$ konveksne funkcije, $\mathcal{D}$ konveksan te neka su α, β pozitivni realni brojevi. Dokažite da je tada i $f = \alpha g + \beta h$ konveksna funkcija.

Zadatak 8. Provjerite konveksnost sljedećih funkcija:

a) $f(x_1, x_2) = 4x_1^2 - 5x_2^2$,

b) $g(x_1, x_2, x_3) = 8x_1^2 - 16x_1x_2 + 8x_2^2 + 10x_3$.

Rješenje: funkcija g je konveksna dok funkcija f nije konveksna.

Zadatak 9. Neka je $f \colon D \subseteq \mathbb{R}^n \to \mathbb{R}$. Skup

$$P(f) = \{(x, \alpha) \in \mathbb{R}^{n+1} \colon x \in D \subseteq \mathbb{R}^n, \alpha \in \mathbb{R}, f(x) \leq \alpha\}$$

nazivamo *epigrafom*, a skup

$$Q(g) = \{(x, \alpha) \in \mathbb{R}^{n+1} \colon x \in D \subseteq \mathbb{R}^n, \alpha \in \mathbb{R}, g(x) \geq \alpha\}$$

podgrafom funkcije f. Odredite epigraf i podgraf funkcije $f \colon [0, 2\pi] \to \mathbb{R}$, $f(x) = \sin x + 2$. Pokažite da za proizvoljnu funkciju $f \colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ vrijedi:

1. ako je funkcija f konveksna, onda je $P(f)$ konveksan skup;
2. ako je funkcija g konkavna, onda je $Q(g)$ konveksan skup;
3. ako je funkcija g konkavna, onda je $(-g)$ konveksna funkcija.

Definicija 6. [20] Kažemo da je funkcija $f \colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ u točki $x \in \operatorname{Int} D$ Gâteaux–diferencijabilna (*G–diferencijabilna*) ako postoji $a \in \mathbb{R}^n$ takav da za svaki $h \in \mathbb{R}^n$ vrijedi

$$\lim_{t \to 0} \frac{1}{t} |f(x + th) - f(x) - ta^T h| = 0. \tag{2.7}$$

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

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

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

Jacobijevu matricu diferencijabilne funkcije f u točki $x = (x_1, \ldots, x_n)^T \in D$ označivat ćemo s

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

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

$$\nabla f(x) = f'(x)^T,$$

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

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

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

$$f(x+h) - f(x) = f'(x)h + o(|h|), \qquad p = 1, \tag{2.8}$$
$$f(x+h) - f(x) = f'(x)h + \tfrac{1}{2}(f''(x)h, h) + o(h^2), \qquad p = 2. \tag{2.9}$$

Lema 1. *Ako je $f \in C^1(D)$ konveksna funkcija na konveksnom skupu D, onda vrijedi*

$$(\nabla f(y), x-y) \le f(x) - f(y) \le (\nabla f(x), x-y), \qquad \forall x, y \in D. \tag{2.10}$$

Dokaz. Oduzimanjem $f(y)$ na obje strane nejednakosti (2.1) dobivamo

$$f(y + \lambda(x-y)) - f(y) \le \lambda(f(x) - f(y)).$$

Primjenjujući na lijevoj strani Lagrangeov teorem o srednjoj vrijednosti dobivamo

$$f(y + \lambda(x-y)) - f(y) = (\nabla f(y + \vartheta\lambda(x-y)), \lambda(x-y)), \qquad \vartheta \in (0,1).$$

Nakon dijeljenja nejednakosti s λ i stavljanja $\lambda \to +0$ dobivamo lijevu stranu tražene nejednakosti.

Desna strana nejednakosti dokazuje se na sličan način. Zaista, lijeva strana nejednakosti (2.10) glasi

$$f(x) - f(y) \geq (\nabla f(y), (x - y)) \,,$$

što nakon zamjene $x \leftrightarrow y$ daje

$$f(y) - f(x) \geq (\nabla f(x), (y - x)) \,,$$

a nakon dijeljenja s (-1) dobivamo desnu stranu nejednakosti (2.10). □

Lema 2. *Neka je $D \subseteq \mathbb{R}^n$ konveksan skup. Funkcija $f \in C^1(D)$ jako je konveksna onda i samo onda ako postoji $m > 0$ takav da je*

$$(\nabla f(y) - \nabla f(x), y - x) \geq m\|y - x\|^2 \quad \text{za sve } x, y \in D. \tag{2.11}$$

Dokaz. Pokazat ćemo jedan smjer leme. Naime, prema [1, Lema 1.1. 2°] funkcija $f \in C^1(D)$ jako je konveksna ako i samo ako postoji $m > 0$ takvo da je

$$f(x) - f(y) \geq (\nabla f(y), (x - y)) + \tfrac{m}{2}\|y - x\|^2 \quad \text{za sve } x, y \in D. \tag{2.12}$$

S druge strane, ako u prethodnoj jednakosti zamijenimo $x \leftrightarrow y$, slijedi da je f konveksna funkcija ako i samo ako

$$f(y) - f(x) \geq (\nabla f(x), (y - x)) + \tfrac{m}{2}\|y - x\|^2 \quad \text{za sve } x, y \in D, \tag{2.13}$$

odnosno

$$f(x) - f(y) \leq (\nabla f(x), (x - y)) - \tfrac{m}{2}\|y - x\|^2 \quad \text{za sve } x, y \in D. \tag{2.14}$$

Sada iz (2.12) i (2.14) slijedi

$$(\nabla f(y), (x - y)) + \tfrac{m}{2}\|y - x\|^2 \leq (\nabla f(x), (x - y)) - \tfrac{m}{2}\|y - x\|^2,$$

za sve $x, y \in D$, što povlači (2.11). □

Teorem 1. *Neka je $D \subseteq \mathbb{R}^n$ konveksan skup takav da je $\operatorname{Int} D \neq \emptyset$. Funkcija $f \in C^2(D)$ jako je konveksna onda i samo onda ako postoji $m > 0$ takav da je*

$$(f''(x)y, y) \geq m\|y\|^2 \quad \text{za sve } x \in D \text{ i za sve } y \in \mathbb{R}^n. \tag{2.15}$$

Dokaz. Budući da je $f \in C^2(D)$ jako konveksna, zamjenom $y \leftrightarrow x + \lambda y$ iz (2.13) slijedi da za sve $x, y \in D$ vrijedi

$$f(x + \lambda y) \geq f(x) + \lambda(\nabla f(x), y) + \tfrac{m}{2}\lambda^2\|y\|^2 \,. \tag{2.16}$$

S druge strane, Taylorova formula u okolini točke x daje

$$f(x + \lambda y) = f(x) + \lambda(\nabla f(x), y) + \tfrac{1}{2}\lambda^2 \left((f''(x + \vartheta\lambda y)y, y)\right), \qquad \vartheta \in (0, 1)\,,$$

što zajedno sa (2.16) povlači (2.15).

Obrat tvrdnje teorema dokažite sami. □

2.3 Lokalni minimum

Razmatrat ćemo neke nužne i dovoljne uvjete da (jako) konveksna funkcija postigne svoj lokalni odnosno globalni minimum.

Definicija 7. Neka je $f\colon D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ proizvoljna funkcija. Kažemo da je $x^\star \in D$ *točka lokalnog minimuma* funkcije f na skupu D ako postoji okolina $\mathcal{O}$ točke $x^\star$, tako da je $f(x) \geq f(x^\star)$ za sve $x \in \mathcal{O} \cap D$. Točku $x^\star$ zovemo *točkom strogog lokalnog minimuma* ako postoji okolina $\mathcal{O}$ točke $x^\star$, tako da vrijedi $f(x) > f(x^\star)$ za sve $x \in \mathcal{O} \cap D \backslash \{x^\star\}$. Točku $x^\star$ zovemo *točkom globalnog minimuma* funkcije f ako je $f(x) \geq f(x^\star)$ za sve $x \in D$.

Definicija 8. Kažemo da je $x^\star \in D \subseteq \mathbb{R}^n$ *stacionarna točka* funkcije $f : D \to \mathbb{R}$ ako je $\nabla f(x^\star) = 0$.

Teorem 2. *Ako je $f\colon D \to \mathbb{R}, \quad D \subseteq \mathbb{R}^n$ jako konveksna neprekidna funkcija na zatvorenom konveksnom skupu D, onda vrijedi:*

(i) funkcija f ograničena je odozdo na D, tj. $f^\star := \min f(x) > -\infty$;

(ii) postoji jedinstvena točka $x^\star \in D$, takva da bude $f(x^\star) = f^\star$;

(iii) skup $M(y) = \{x \in D\colon f(x) \leq f(y)\}$ ograničen je za svaki $y \in D$.

Dokaz. Za dokaz ekvivalencije (i) i (ii) vidi npr. [1, Teorem 2.1]. Ekvivalencija (i) i (iii) slijedi iz činjenica da je f neprekidna na D pa je stoga i svaki skup $M(y) = \{x \in D\colon f(x) \leq f(y)\}$ neprazan i ograničen (vidi primjerice [1, Korolar 1.1]). □

Lema 3. *Kako bi konveksna funkcija $f \in C^1(D)$ postigla svoj infimum na konveksnom skupu $D \subseteq \mathbb{R}^n$ u točki $x^\star \in D$ nužno je i dovoljno da bude:*

$$(\nabla f(x^\star), x - x^\star) \geq 0 \quad \text{za sve } x \in D. \tag{2.17}$$

Ako je $x^\star \in \operatorname{Int} D$, onda je ovaj uvjet ekvivalentan jednakosti $\nabla f(x^\star) = 0$.

Dokaz. Pretpostavimo da konveksna funkcija f postiže svoj infimum u točki $x^\star \in D$. Tada za svaki $x \in D$ i svaki $t \in [0,1]$ vrijedi $x^\star + t(x - x^\star) = tx + (1-t)x^\star \in D$, pa je $f(x^\star) \leq f(x^\star + t(x - x^\star))$, odnosno

$$f(x^\star + t(x - x^\star)) - f(x^\star) \geq 0, \qquad t \in [0,1]\,.$$

Korištenjem svojstva derivacije složene funkcije slijedi

$$(\nabla f(x^\star), x - x^\star) = \lim_{t \to 0^+} \tfrac{1}{t} \left[f(x^\star + t(x - x^\star)) - f(x^\star) \right] \geq 0.$$

Drugi smjer i ekvivalentnost uvjeta (2.17) i jednakosti $\nabla f(x^\star) = 0$ za $x^\star \in \operatorname{Int} D$ dokažite sami. □

Primjer 4. Razmotrimo kvadratnu funkciju $f\colon \mathbb{R}^2 \to \mathbb{R}$,

$$f(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2.$$

Ovu funkciju možemo promatrati kao kvadratnu formu kojoj je pridružena simetrična matrica $A = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}$. Ovoj matrici u bazi (e_1, e_2) pridružen je jedinstveni simetrični linearni operator $\mathcal{A}$ sa svojstvenim vrijednostima $\lambda_1 = \frac{1}{2}, \lambda_2 = \frac{3}{2}$. Prema *primjeru* 2. ova funkcija jako je konveksna na čitavom području definicije, a prema *lemi* 3. i *teoremu* 2. postoji jedinstvena točka minimuma $x^\star$ koja se može pronaći tako da riješimo jednadžbu $\nabla f(x) = 0$. Matrica drugih derivacija (Hessijan) funkcije f u točki $x^\star$ mora biti pozitivno definitna. Lako se može vidjeti da je točka minimuma ove funkcije $x^\star = (0, 0)$.

Nadalje, ako pretpostavimo da je $f\colon D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ dovoljno puta diferencijabilna funkcija definirana na otvorenom skupu $D \subset \mathbb{R}^n$, onda možemo koristiti jednostavne nužne i dovoljne uvjete egzistencije lokalnog minimuma koji su dobro razrađeni u elementarnoj analizi. Sljedeći teorem daje nužne uvjete lokalnog minimuma u tom slučaju.

Teorem 3. *Neka je $f \in C^1(D)$ neprekidno diferencijabilna funkcija definirana na otvorenom skupu $D \subset \mathbb{R}^n$ i neka je $x^\star$ točka lokalnog minimuma od f. Tada vrijedi:*

(i) $x^\star$ je stacionarna točka funkcije f,

(ii) ako je $f \in C^2(D)$, njezin Hessijan $\nabla^2 f(x^\star)$ u točki $x^\star$ je simetrična pozitivno semidefinitna matrica.

Dokaz. (i) Pretpostavimo da je $f \in C^1(D)$ i da $x^\star$ nije stacionarna točka funkcije f, tj. $\nabla f(x^\star) \neq 0$, i definirajmo funkciju

$$\varphi(t) := f\left(x^\star - t\nabla f(x^\star)\right), \quad t \in \mathbb{R}.$$

Za maleni $|t|, t \in \mathbb{R}$ funkcija φ je neprekidno diferencijabilna funkcija za koju vrijedi

$$\varphi'(0) = -\nabla f(x^\star)^T \nabla f(x^\star) = -\|\nabla f(x^\star)\|_2^2$$

Kako je $\nabla f(x^\star) \neq 0$, mora biti $\varphi'(0) < 0$. Zato postoji $\epsilon > 0$, takav da bude $\varphi(\epsilon) < \varphi(0)$. To znači da $f\left(x^\star - \epsilon\nabla f(x^\star)\right) < f(x^\star)$, odnosno $x^\star$ nije točka lokalnog minimuma od f, što je u suprotnosti s polaznom pretpostavkom.

(ii) Pretpostavimo da je $f \in C^2(D)$ i da $\nabla^2 f(x^\star)$ nije pozitivno semidefinitna matrica. To znači da postoji vektor $d \in \mathbb{R}^n, \quad d \neq 0$, takav da je $d^T \nabla^2 f(x^\star) d < 0$. Prema Taylorovom teoremu, a zbog činjenice da je $\nabla f(x^\star) = 0$, za maleni $t > 0$ postoji $\tau \in (0, t)$, takav da bude

$$f(x^\star - td) = f(x^\star) + \tfrac{1}{2} t^2 d^T \nabla^2 f(x^\star - \tau d) d.$$

Odavde za dovoljno maleni $t > 0$ zbog neprekidnosti od $\nabla^2 f$ vrijedi $f(x^\star - td) < f(x^\star)$. To bi značilo da $x^\star$ nije točka lokalnog minimuma od f, što je u suprotnosti s polaznom pretpostavkom. $\square$

Sljedeći teorem daje dovoljne uvjete lokalnog minimuma.

Teorem 4. *Neka je $f \in C^2(D)$, $D \subset \mathbb{R}^n$ otvoren skup, a $x^\star \in D$ stacionarna točka od f u kojoj je Hessijan $\nabla^2 f(x^\star)$ pozitivno definitan. Tada je $x^\star$ točka strogog lokalnog minimuma funkcije f.*

Dokaz. Kako je $x^\star \in \operatorname{Int} D$ i $\nabla f(x^\star) = 0$, prema Taylorovom teoremu za svaki dovoljno maleni vektor $d \in \mathbb{R}^n$ vrijedi

$$f(x^\star + d) = f(x^\star) + \tfrac{1}{2} d^T \nabla^2 f(x^\star + \vartheta d) d, \qquad \vartheta \in (0, 1).$$

Kako je $\nabla^2 f(x^\star)$ pozitivno definitan, postoji $\lambda > 0$, takav da za sve $d \in \mathbb{R}^n$ vrijedi

$$d^T \nabla^2 f(x^\star) d \geq \lambda d^T d.$$

Zbog toga je

$$\begin{aligned} f(x^\star + d) &= f(x^\star) + \tfrac{1}{2} d^T \nabla^2 f(x^\star) d - \tfrac{1}{2} d^T \left(\nabla^2 f(x^\star) - \nabla^2 f(x^\star + \vartheta d) \right) d \\ &\geq f(x^\star) + \tfrac{1}{2} \left(\lambda - \| \nabla^2 f(x^\star) - \nabla^2 f(x^\star + \vartheta d) \| \right) d^T d, \end{aligned}$$

pri čemu je korištena Cauchyjeva nejednakost i svojstvo kompatibilnosti vektorskih i matričnih normi. Zbog $\lambda > 0$ i neprekidnosti od $\nabla^2 f$, za dovoljno maleni vektor d takvi da $\lambda - \| \nabla^2 f(x^\star) - \nabla^2 f(x^\star + \vartheta d) \| > 0$ imamo da je $f(x^\star + d) > f(x^\star)$ za dovoljno malene vektore $d \neq 0$. Po *definiciji* 7. to znači da je $x^\star$ točka strogog lokalnog minimuma funkcije f. $\square$

Zadatak 10. Odredite točke lokalnog minimuma funkcije

a) $f(x, y) = 2 - 2x + x^2 + 12y + 2y^2$,

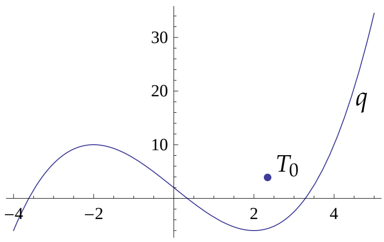
b) $f(x, y) = \frac{16}{x} + \frac{x}{2y} + y$.

Rješenje: a) minimum je u točki $(1, -3)$, b) minimum je u točki $(8, 2)$.

2.4 Primjeri koji vode na probleme ekstrema

Navest ćemo nekoliko primjera u kojima se prirodno pojavljuje problem optimizacije funkcije jedne ili više varijabli. Ovi primjeri kasnije će nam poslužiti kao test-primjeri kod raznih metoda minimizacije.

Primjer 5. Treba izračunati l_2 udaljenost točke $T_0 = (3,4)$ do kubične parabole $q(x) = 0.5x^3 - 6x + 2$ (vidi *sliku* 2.1.).



Slika 2.1. Udaljenost točke T_0 do grafa kubne parabole

l_2 udaljenost točke $T_0 = (x_0, y_0)$ do neke točke $T = (x, f(x))$ na grafu funkcije q zadana je s

$$d_2(x) := d_2(T_0(x_0,y_0), T(x,q(x)) = \sqrt{(x-x_0)^2 + (q(x)-y_0)^2},$$

pa je određivanje udaljenosti točke $T_0(x_0,y_0)$ do grafa funkcije q zadane na intervalu $[a,b]$ problem određivanja globalnog minimuma funkcije d_2 na segmentu $[a,b]$ (vidi *sliku* 2.2*a*).

Budući da je $x \mapsto \sqrt{x}$ monotono rastuća funkcija, onda naš problem možemo svesti na određivanje globalnog minimuma jednostavnije funkcije

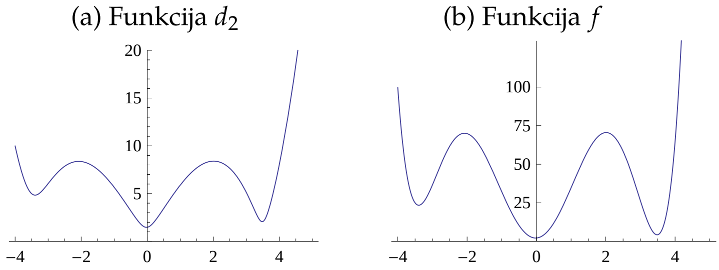
$$f(x) = (x-x_0)^2 + (q(x)-y_0)^2. \tag{2.18}$$

Kako je $q(x) = 0.5x^3 - 6x + 2$ iz (2.18) dobivamo (vidi *sliku* 6.2.*b)*

$$f(x) = (x-3)^2 + \left(\tfrac{1}{2}x^3 - 6x + 2 - 4\right)^2.$$

Dakle, minimizacija funkcije f je problem minimizacije "glatke" (beskonačno derivabilne) funkcije jedne varijable (polinoma 6. stupnja), pa govorimo o *problemu jednodimenzionalne minimizacije*. Na slici se može uočiti da ova funkcija ima tri lokalna minimuma od kojih je jedan ujedno i globalni minimum.

U sljedećem primjeru promatrat ćemo isti problem kao u primjeru 5., ali kao mjeru udaljenosti koristit ćemo drugu razdaljinsku funkciju.



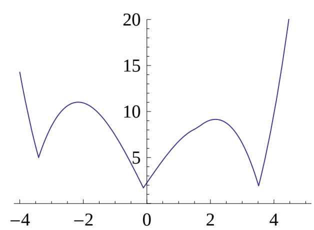
Slika 2.2. l_2 i LS udaljenost točke T_0 do grafa kubične parabole

Primjer 6. Treba izračunati l_1 udaljenost točke $T_0 = (3, 4)$ do kubične parabole $q(x) = 0.5x^3 - 6x + 2$ (vidi *sliku 2.1.*).

l_1 udaljenost točke $T_0 = (x_0, y_0)$ do neke točke $T = (x, f(x))$ na grafu funkcije q zadana je s

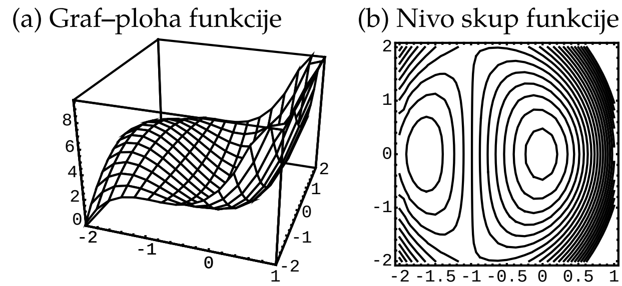
$$d_1(x) = d_1(T_0(x_0, y_0), T(x, q(x)) = |x - x_0| + |q(x) - y_0|,$$

pa je određivanje l_1 udaljenosti točke T_0 do grafa funkcije q zadane na intervalu $[a, b]$ problem određivanja globalnog minimuma funkcije d_1 na segmentu $[a, b]$ (vidi *sliku* 2.3.). Dakle, i u ovom primjeru radi se također o problemu jednodimenzionalne minimizacije, ali ovaj puta minimizirajuća funkcija je nediferencijabilna (graf ima "vrhove"). Također na slici se može uočiti da ova funkcija ima tri lokalna minimuma od kojih je jedan ujedno i globalni minimum.



Slika 2.3. l_1 udaljenost točke T_0 do grafa kubične parabole

Primjer 7. Funkcija $f(x_1, x_2) = 2x_1^3 + x_1x_2^2 + 5x_1^2 + x_2^2$ ima četiri stacionarne točke: $T_1 = (0, 0)$, $T_2 = (-\frac{5}{3}, 0)$, $T_3 = (-1, 2)$, $T_4 = (-1, -2)$. Može se pokazati da je T_1 točka minimuma, T_2 točka maksimuma, a da su T_3 i T_4 sedlaste točke. Na *slici 2.4.* prikazana je ploha i nivo skupovi ove funkcije.

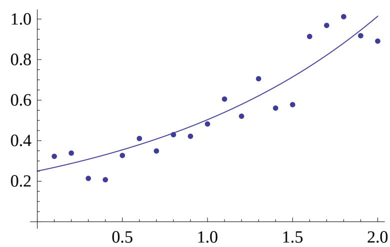


Slika 2.4. Funkcija $f(x_1,x_2) = 2x_1^3 + x_1x_2^2 + 5x_1^2 + x_2^2$

Primjer 8. Zadani su podaci mjerenja (t_i, y_i) $i = 1,\ldots,m$, $y_i > 0$. Treba odrediti optimalne parametre $b^\star, c^\star$ eksponencijalne model-funkcije

$$f(t;b,c) = b\,e^{ct},$$

tako da suma kvadrata izmjerenih od teoretskih vrijednosti bude minimalna (vidi *sliku* 2.5.).



Slika 2.5. Podaci i eksponencijalna model-funkcija

Optimalne parametre $b^\star, c^\star$ eksponencijalne model-funkcije $f(t;b,c) = b\,e^{ct}$ potražit ćemo tako da minimiziramo funkcional

$$F(b,c) = \frac{1}{2}\sum_{i=1}^{m}\left(b\,e^{cx_i} - y_i\right)^2.$$

Može se pokazati da uvijek postoji točka $(b^\star, c^\star) \in \mathbb{R}^2$ u kojoj se postiže globalni minimum ove funkcije. Primijetite da se u ovom slučaju radi o problemu minimizacije "glatke" (višestruko derivabilne) funkcije dviju varijabli, tako da govorimo o *problemu višedimenzionalne minimizacije*.

POGLAVLJE 3

Jednodimenzionalna minimizacija

Zadana je funkcija $f\colon [a,b] \to \mathbb{R}$, koja na intervalu $[a,b]$ postiže $\inf\limits_{x\in[a,b]} f(x)$. Promatramo sljedeće probleme:

1. Odrediti barem jedan x^* tako da je $x^\star \in \operatorname*{argmin}\limits_{x\in[a,b]} f(x)$ i $f^\star = f(x^\star)$;

2. Odrediti $U = \{u \in [a,b] \colon f(u) = f^\star\}$, tj. skup svih točaka iz segmenta $[a,b]$ u kojima se postiže minimum funkcije f.

3.1 Strogo kvazikonveksne funkcije

Promatrat ćemo jednu važnu široku klasu funkcija koje svoj infimum postižu na segmentu $[a,b]$. Od tih funkcija neće se zahtijevati derivabilnost, pa čak ni neprekidnost (vidi primjerice [20, 29]).

Definicija 9. Kažemo da je funkcija $f\colon [a,b] \to \mathbb{R}$ *strogo kvazikonveksna* na intervalu $[a,b]$ ako postiže svoj infimum $f^\star = \inf\limits_{x\in[a,b]} f(x)$ u nekoj točki $x^\star \in [a,b]$ i ako postoje $\alpha,\beta \in [a,b]$, $a \le \alpha \le \beta \le b$ takvi da je

(i) funkcija f strogo monotono padajuća na $[a,\alpha)$ (ako je $a < \alpha$);

(ii) funkcija f strogo monotono rastuća na $(\beta,b]$ (ako je $\beta < b$);

(iii) $f(u) = f^\star$ za svaki $u \in (\alpha,\beta)$ (ako je $\alpha < \beta$).

Termin "strogo kvazikonveksna funkcija" može se pronaći u literaturi i pod nazivom "strogo unimodalna funkcija" (vidi primjerice [15]).

Pojam kvazikonveksne odnosno strogo kvazikonveksne funkcije može se definirati i na drugi način (vidi primjerice [20, 29]), pri čemu taj pojam možemo definirati za realnu funkciju definiranu na nekom skupu $D \subseteq \mathbb{R}^n$.

Definicija 10. Funkcija $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ *kvazikonveksna* je na D ako vrijedi

$$f(\alpha x + (1-\alpha)y) \leq \max\{f(x), f(y)\} \quad \forall x, y \in D, \quad \forall \alpha \in [0,1]. \tag{3.1}$$

Ako u (3.1) vrijedi stroga nekednakost za $\alpha \in (0,1)$ i $x \neq y$, onda kažemo da je f *strogo kvazikonveksna* na D.

Primjedba 3. Svaka konveksna funkcija je kvazikonveksna, ali obrat ne vrijedi. Ako su funkcije f_i, $i = 1, \ldots, m$ kvazikonveksne na D, onda je i funkcija $f(x) = \max\limits_{i=1,\ldots,m} f_i(x)$ također kvazikonveksna na D.

Lema 4. *([20, 9]) Funkcija $f : D \to \mathbb{R}$ kvazikonveksna je na D onda i samo onda ako je za svaki $\lambda \in \mathbb{R}$ nivo skup $D_\lambda = \{x \in D : f(x) \leq \lambda\}$ konveksan.*

Dokaz. *(Nužnost)* Pretpostavimo da je funkcija $f : D \to \mathbb{R}$ kvazikonveksna i dokažimo da je za proizvoljni $\lambda \in \mathbb{R}$ skup D_λ konveksan. Neka su $x, y \in D_\lambda$. To znači da je $f(x) \leq \lambda$ i $f(y) \leq \lambda$. Kako je f kvazikonveksna, za proizvoljni $\alpha \in [0,1]$ vrijedi

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

iz čega zaključujemo da je $\alpha\, x + (1-\alpha)y \in D_\lambda$.

(Dovoljnost) Pretpostavimo da je za proizvoljni $\lambda \in \mathbb{R}$ skup D_λ konveksan i dokažimo da je tada funkcija $f : D \to \mathbb{R}$ kvazikonveksna. Za proizvoljne $x, y \in D$ definirajmo $\lambda_0 = \max\{f(x), f(y)\}$. Tada su $x, y \in D_{\lambda_0}$, a kako je D_{λ_0} konveksan, onda je $\alpha\, x + (1-\alpha)y \in D_{\lambda_0}, \quad \forall \alpha \in [0,1]$, iz čega slijedi

$$f(\alpha\, x + (1-\alpha)y) \leq \lambda_0 = \max\{f(x), f(y)\}.$$

□

Primjer 9. Navedimo nekoliko primjera strogo kvazikonveksnih funkcija:

$f_1 : [a,b] \to \mathbb{R}$, $f_1(x) = x^2$, $\qquad f_2 : [a,b] \to \mathbb{R}$, $f_2(x) = |x| + x + \operatorname{sign} x$,
$f_3 : [1,2] \to \mathbb{R}$, $f_3(x) = \ln x$,

$$f_4 : [0,2] \to \mathbb{R}, \quad f_4(x) = \begin{cases} x^2, & x \in [0,1] \\ \sqrt{x}, & x \in [1,2] \end{cases}$$

Primjer 10. Funkcija $g_1 : [-1,1] \to \mathbb{R}$, $\ g_1(x) = \begin{cases} |x|, & x \neq 0 \\ -1, & x = 0 \end{cases}$ strogo je kvazikonveksna funkcija, dok funkcija $g_2 : [-1,1] \to \mathbb{R}$, $\ g_2(x) = \begin{cases} |x|, & x \neq 0 \\ 1, & x = 0 \end{cases}$ nije strogo kvazikonveksna funkcija.

Primjer 11. Funkcija $f : [0,3] \to \mathbb{R}$, $\ f(x) = \begin{cases} (x-1)^2 + 1, & x \leq 1 \\ 1, & 1 \leq x \leq 2 \\ \sqrt{x-2} + 1, & x \geq 2 \end{cases}$
strogo je kvazikonveksna funkcija, koja svoj infimum postiže u svim točkama intervala $[1,2]$.

Zadatak 11. Kako treba proširiti funkciju $f(x) = \frac{1}{e^{1/u^2}+1}$ u točki $x_0 = 0$ da bi bila strogo kvazikonveksna na (a) $[0,1]$, (b) $[-1,1]$, (c) $[-1,0]$?

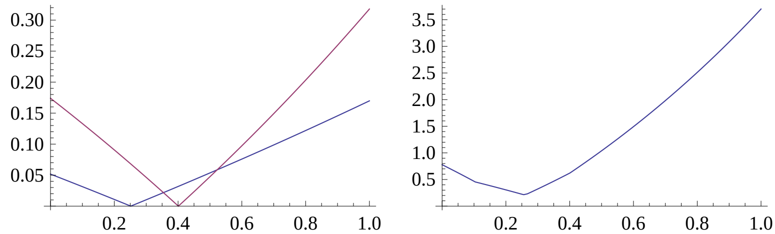
Zadatak 12. Kako treba proširiti funkciju $f(x) = \begin{cases} x, & x > 0 \\ -x + A, & x < 0 \end{cases}$ u točki $x_0 = 0$ da bi bila strogo kvazikonveksna na (a) $[0,1]$, (b) $[-1,0]$, (c) $[-1,1]$, (d) $[-a,b]$ na $\mathbb{R}$? Razmotrite slučajeve: $A = 0, +1, -1$.

Zadatak 13. Ako su f_1, f_2 kvazikonveksne funkcije na $[a,b]$, jesu li takve i funkcije: $f_1 + f_2$, $f_1 - f_2$, $f_1 \cdot f_2$, $\frac{f_1}{f_2}$?

Zadatak 14. Pokažite da zbroj dvije strogo kvazikonveksne funkcije ne mora biti strogo kvazikonveksna funkcija. Primjerice, za podatke (t_i, y_i), $i = 1, \ldots, m$, funkcija $c \mapsto |e^{ct_i} - y_i|$ je kvazikonveksna, ali funkcija

$$F(c) = \sum_{i=1}^{m} |e^{ct_i} - y_i| \ ?$$

ne mora biti strogo kvazikonveksna funkcija. Pokušajte pronaći podatke koji potvrđuju ovu tvrdnju.



Slika 3.1. Eksponencijalna regresija

Zadatak 15. 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 sljedeća funkcija kvazikonveksna

$$F(c) = \sum_{i=1}^{m} |e^{ct_i} - y_i| \text{ ?}$$

3.2 Metoda jednodimenzionalne minimizacije

Navest ćemo nekoliko najpoznatijih metoda jednodimenzionalne minimizacije i to najprije metode kod kojih nije potrebno poznavanje derivacije minimizirajuće funkcije (Metoda polovljenja, Metoda zlatnog reza, Metoda parabole i Brentova metoda), a nakon toga Newtonovu metodu koja pretpostavlja dovoljnu glatkost minimizirajuće funkcije.

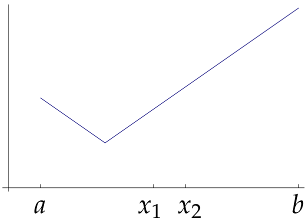
Metoda polovljenja

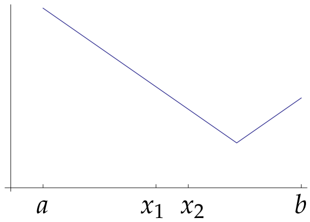
Zadana je strogo kvazikonveksna funkcija $f\colon [a,b] \to \mathbb{R}$. Treba odrediti točku $x^\star = \operatorname*{argmin}_{x\in[a,b]} f(x) \in [a,b]$ s točnošću $\delta > 0$. U nekoliko koraka opisat ćemo ovu vrlo popularnu metodu za koju nije potrebno poznavanje derivacije funkcije.

Korak 1. Definirati:

$$x_1 := \tfrac{1}{2}(a+b) - \tfrac{\delta}{2},$$
$$x_2 := \tfrac{1}{2}(a+b) + \tfrac{\delta}{2}, \quad \delta > 0;$$
$$[a_1, b_1] := \begin{cases} [a, x_2], & f(x_1) \le f(x_2) \\ [x_1, b], & f(x_1) > f(x_2) \end{cases} .$$

Svojstva:





Slika 3.2. Metoda polovljenja

$$
\begin{aligned}
&x^\star \in [a_1, b_1],\\
&b_1 - a_1 = \tfrac{1}{2}(b-a) + \tfrac{\delta}{2}\\
&x_1^\star := \tfrac{1}{2}(a_1 + b_1),\\
&|x^\star - x_1^\star| \le \tfrac{1}{2}(b_1 - a_1) = \tfrac{1}{2^2}(b-a) + \tfrac{1}{2^2}\delta < \tfrac{1}{2^2}(b-a) + \delta.
\end{aligned}
$$

Korak 2. Definirati:

$$
\begin{aligned}
&x_3 := \tfrac{1}{2}(a_1 + b_1) - \tfrac{\delta}{2},\\
&x_4 := \tfrac{1}{2}(a_1 + b_1) + \tfrac{\delta}{2}, \quad \delta > 0;\\
&[a_2, b_2] := \begin{cases} [a_1, x_4], & f(x_3) \le f(x_4) \\ [x_3, b_1], & f(x_3) > f(x_4) \end{cases}.
\end{aligned}
$$

Svojstva:

$$
\begin{aligned}
&x^\star \in [a_2, b_2],\\
&b_2 - a_2 = \tfrac{1}{2}(b_1 - a_1) + \tfrac{\delta}{2} \quad = \tfrac{1}{2^2}(b-a) + \tfrac{3}{4}\delta < \tfrac{1}{2^2}(b-a) + \delta,\\
&x_2^\star := \tfrac{1}{2}(a_2 + b_2),\\
&|x^\star - x_2^\star| \le \tfrac{1}{2}(b_2 - a_2) < \tfrac{1}{2^3}(b-a) + \tfrac{1}{2}\delta < \tfrac{1}{2^3}(b-a) + \delta
\end{aligned}
$$

Korak k. Neka je poznat interval $[a_{k-1}, b_{k-1}]$ koji sadržava $x^\star$ i neka je

$$b_{k-1} - a_{k-1} < \tfrac{1}{2^{k-1}}(b-a) + \delta, \quad k \ge 2.$$

Definirati:

$$
\begin{aligned}
&x_{2k-1} := \tfrac{1}{2}(a_{k-1} + b_{k-1}) - \tfrac{\delta}{2},\\
&x_{2k} := \tfrac{1}{2}(a_{k-1} + b_{k-1}) + \tfrac{\delta}{2}, \quad \delta > 0;\\
&[a_k, b_k] := \begin{cases} [a_{k-1}, x_{2k}], & f(x_{2k-1}) \le f(x_{2k}) \\ [x_{2k-1}, b_{k-1}], & f(x_{2k-1}) > f(x_{2k}) \end{cases}.
\end{aligned}
$$

Svojstva:

$$
\begin{aligned}
& b_k - a_k = \tfrac{1}{2}(b_{k-1} - a_{k-1}) + \tfrac{\delta}{2} < \tfrac{1}{2^k}(b-a) + \tfrac{\delta}{2} \\
& x_k^\star := \tfrac{1}{2}(a_k + b_k), \\
& |x^\star - x_k^\star| \le \tfrac{1}{2}(b_k - a_k) < \tfrac{1}{2^{k+1}}(b-a) + \tfrac{\delta}{2} < \tfrac{1}{2^{k+1}}(b-a) + \delta.
\end{aligned}
$$

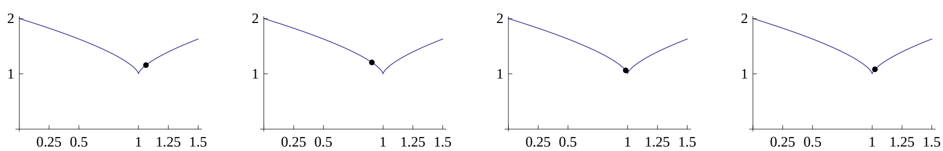
Metodu nazivamo *Metodom polovljenja* jer u svakom koraku interval dijelimo na gotovo dvije polovine –ako je δ manji, time je dijeljenje intervala bliže raspolavljanju. Najbolja točnost koja se može postići je $\delta > 0$. Više o metodi vidi u [15, 20, 29].

Primjer 12. Metodu polovljenja ilustrirat ćemo na primjeru minimizacije funkcije $f : [0, 1.5] \to \mathbb{R}$, $f(x) = 1 + \sqrt[3]{(x-1)^2}$. U ovom slučaju je $x^\star = 1$.

Tijek iterativnog procesa prikazan je u tablici 3.1. i na slici 3.3.

k	x_k	x_{k+1}	$x_k^\star$	$\lvert x_k^\star - x^\star\rvert$
1	0.625	0.875	1.0625	0.0625
2	0.9375	1.1875	0.9063	0.0938
3	0.7813	1.0313	0.9844	0.0156
4	0.8594	1.1094	1.0234	0.0234

Tablica 3.1: Metoda polovljenja za funkciju $f(x) = 1 + \sqrt[3]{(x-1)^2}$



Slika 3.3. Metoda polovljenja za funkciju $f(x) = 1 + \sqrt[3]{(x-1)^2}$

Niže je naveden kratki *Mathematica*–program za Metodu polovljenja.

```
In[1]:=f[x_] := 1 + ((x - 1)^2)^(1/3)
In[2]:=a = 0; b = 1.5; xz = 1; d = .25; n = 5; xa = Table[0, {i, n}];
In[3]:= sl1 = Plot[f[x], {x, a, b}];
In[4]:=Do[
          x1 = (a + b - d)/2;
          x2 = (a + b + d)/2;
          If[f[x1] <= f[x2], b = x2, a = x1];
          xa[[i]] = (a + b)/2;
```

```
    Print[i, " ", {x1, x2}, " ", u = xa[[i]], " ", Abs[u - xz]];
    Print[sl[i] = Show[{sl1,
       Graphics[{PointSize[.03], Point[{{xa[[i]], f[xa[[i]]]}}]}]}]];
   , {i, n}]
```

Metoda zlatnog reza

Jedna korekcija prethodno pokazane metode polovljenja sastoji se u tome da točke oko polovišta intervala biramo u smislu omjera "zlatnog reza":

$$x_1 = a + \tfrac{3-\sqrt{5}}{2}(b-a), \tag{3.2}$$

$$x_2 = a + b - x_1 = a + \tfrac{\sqrt{5}-1}{2}(b-a). \tag{3.3}$$

Vrijedi

$$x_1 < x_2,$$

$$\tfrac{b-a}{b-x_1} = \tfrac{b-x_1}{x_1-a} = \tfrac{b-a}{x_2-a} = \tfrac{x_2-a}{b-x_2} = \tfrac{1+\sqrt{5}}{2} \approx 1.618.$$

Primijetite također da x_1 čini zlatni rez na $[a, x_2]$:

$$x_2 - x_1 < x_1 - a = b - x_2; \qquad \tfrac{x_2-a}{x_1-a} = \tfrac{x_1-a}{x_2-x_1}.$$

Metodu zlatnog reza opisat ćemo u nekoliko koraka.

1. $[a_1, b_1] = [a, b]$;

   Definirati x_1, x_2 kao u (3.2.-3.3.);

   $$\begin{array}{lcll} f(x_1) \le f(x_2) & \Longrightarrow & [a_2, b_2] = [a_1, x_2]; & x_2^\star = x_1, \\ f(x_1) > f(x_2) & \Longrightarrow & [a_2, b_2] = [x_1, b_1]; & x_2^\star = x_2; \end{array}$$

   $b_2 - a_2 = x_2 - a_1 = b_1 - x_1 = \tfrac{\sqrt{5}-1}{2}(b-a)$.

2. Simetrično s $x_2^\star$ izabrati točku $x_3 = a_2 + b_2 - x_2^\star$, koja također čini zlatni rez intervala $[a_2, b_2]$.
   Usporedbom $f(x_3)$ i $f(x_2^\star)$ dobivamo novi interval $[a_3, b_3]$;

   $b_3 - a_3 = x_3 - a_2 = b_2 - x_2^\star = \left(\tfrac{\sqrt{5}-1}{2}\right)^2 (b-a)$.

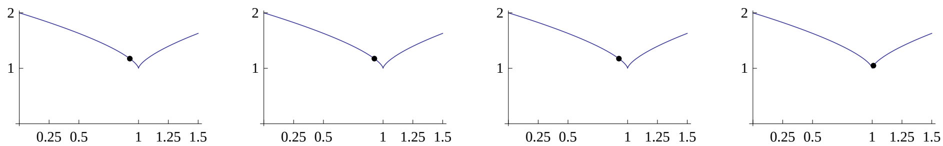
Više o metodi vidi u [15, 20, 29].

Primjer 13. Metodu zlatnog reza ilustrirat ćemo također na primjeru minimizacije funkcije iz *primjera* 12.: $f : [-1, 2] \to \mathbb{R}, f(x) = 1 + \sqrt[3]{(x-1)^2}$.

Tijek iterativnog procesa prikazan je u tablici 3.2. i na slici 3.4.
Niže je naveden kratki *Mathematica*–program za Metodu zlatnog reza.

k	x_k	x_{k+1}	$x_k^\star$	$\lvert x_k^\star - x^\star\rvert$
1	0.5729	0.9271	0.9271	0.0729
2	0.9270	1.1459	0.9271	0.0729
3	0.7918	0.9271	0.9271	0.0729
4	0.9270	1.0106	1.0106	0.0106

Tablica 3.2: Metoda zlatnog reza za funkciju $f(x) = 1 + \sqrt[3]{(x-1)^2}$



Slika 3.4. Metoda zlatnog reza za funkciju $f(x) = 1 + \sqrt[3]{(x-1)^2}$

```
In[1]:=f[x_] := 1 + ((x - 1)^2)^(1/3)
In[2]:=a = 0; b = 1.5; xz = 1; n = 5; xa = Table[0, {i, n}];
In[3]:= sl1 = Plot[f[x], {x, a, b}];
In[4]:=Do[
          x1 = a + (3 - Sqrt[5.]) (b - a)/2;
          x2 = a + (Sqrt[5.] - 1) (b - a)/2;
          If[f[x1] <= f[x2], b = x2; xa[[i]] = x1, a = x1; xa[[i]] = x2];
          Print[i, " ", {x1, x2}, " ", u = xa[[i]], " ", Abs[u - xz]];
          Print[sl[i] = Show[{sl1,
             Graphics[{PointSize[.03], Point[{{xa[[i]], f[xa[[i]]]}}]}]}]];
         , {i, n}]
```

Primjer 14. Zadana je funkcija $f : [-2, 2] \to \mathbb{R}$, $f(x) = 5|x+1| - |3x+2| + |5x-2|$ (vidi *sliku 3.5a)*. Metodom polovljenja i metodom zlatnog reza odredite njene točke lokalnih minimuma i maksimuma.

Metoda parabole

Razmotrit ćemo još jednu jednostavnu metodu za određivanje lokalnog minimuma funkcije $f : [a, b] \to \mathbb{R}$, kod koje također neće biti potrebno poznavati derivaciju funkcije. U intervalu $[a, b]$ izabrat ćemo takvu točku $c \in [a, b]$, u kojoj funkcija f prima manju vrijednost nego na rubovima intervala, tj.

$$f(c) < \min\{f(a), f(b)\}. \tag{3.4}$$

Pomoću tri točke $(a, f(a))$, $(c, f(c))$, $(b, f(b))$ odredit ćemo interpolacijski polinom primjenom Lagrangeove formule.

$$P_2(x) = f(a)p_a(x) + f(c)p_c(x) + f(b)p_b(x), \qquad \text{gdje je} \tag{3.5}$$
$$p_a(x) = \frac{(x-c)}{(a-c)}\frac{(x-b)}{(a-b)}, \quad p_c(x) = \frac{(x-a)}{(c-a)}\frac{(x-b)}{(c-b)}, \quad p_b(x) = \frac{(x-a)}{(b-a)}\frac{(x-c)}{(b-c)}$$

Polinom P_2 aproksimira funkciju f na intervalu $[a, b]$ (vidi *sliku* 3.5*a*). Zato ćemo umjesto traženja minimuma funkcije f potražiti minimum interpolacijskog polinoma drugog stupnja. Lako dobivamo stacionarnu točku kvadratne funkcije (3.5) (provjerite!)

$$u_0 = \frac{1}{2}\frac{a^2(f(b)-f(c))+c^2(f(a)-f(b))+b^2(f(c)-f(a))}{a(f(b)-f(c))+c(f(a)-f(b))+b(f(c)-f(a))} \tag{3.6}$$

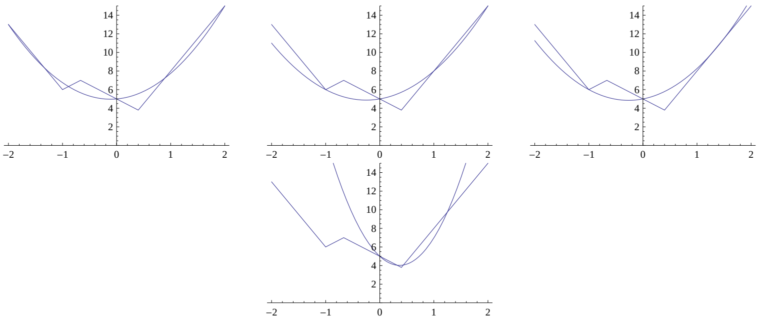
Pretpostavka (3.4) povlači da se radi o minimumu. Točku u_0 uzimamo kao početnu aproksimaciju točke minimuma $u^\star$. Zadamo li točnost $\varepsilon > 0$, iterativni postupak nastavljamo na sljedeći način:

- Ako je $|f(u_0) - P_2(u_0)| < \varepsilon$, uzimamo $u^\star = u_0$ i postupak je završen.
- U suprotnom sortiramo brojeve a, c, b, u_0 i potražimo onaj u kojem funkcija f prima najmanju vrijednost (primijetite da to mogu biti samo u_0 ili c). Taj broj označimo s c, a s a i b označimo neposredne susjede.

Postupak dalje ponavljamo. Niže je naveden program koji podržava opisani postupak. Izradite odgovarajući modul.

```
In[1]:=While[k = k + 1;
       x = {a, c, b}; y = f[x];
       u0 = (c^2 (f[a]-f[b]) + a^2 (f[b]-f[c]) + b^2 (f[c]-f[a]))/
            (2 (c(f[a]-f[b]) +  a  (f[b]-f[c]) +  b  (f[c]-f[a])));
       Print[k, {a, c, b},{u0, f[u0], Abs[f[u0]-P[u0, x, y]]}];
       Abs[f[u0] - P[u0, x, y]] > eps,
       nodes = Sort[{a, c, u0, b}];
        fnodes = f[nodes];
         min = Min[fnodes];
          imin = Position[fnodes, min][[1, 1]];
       c = nodes[[imin]];
        a = nodes[[imin - 1]];
         b = nodes[[imin + 1]]]
```

Primjer 15. Za funkciju $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 5|x+1| - |3x+2| + |5x-2|$ iz *primjera* 14. primjenom metode parabola odredit ćemo točku minimuma funkcije f na intervalu $[-2, 2]$ uz točnost $\varepsilon = .005$.



Slika 3.5. Metoda parabole za funkciju $f(x) = 5|x+1| - |3x+2| + |5x-2|$

Tijek iterativnog postupka prikazan je u tablici 3.3. i na slici 3.5. Ulazne podatke zadat ćemo na sljedeći način

```
ln[2]:=a=-2.; b=2.; eps=.00005; k=0;
       c =(a+b)/2; If[f[c] = Min[f[a], f[b]], Print["Biraj novi c"]];
       f[x_] := 5 Abs[x + 1] - Abs[3 x + 2] + Abs[5 x - 2];
       slf = Plot[f[x], {x, a, b}, AxesOrigin -> {-2, 0},
                ImageSize -> 200]
       x = {a, c, b}; y = f[x];
       slP = Plot[P[u, x, y], {u, -2, 2}, PlotRange -> {0, 20}];
       Show[slP, slf]
```

k	a	c	b	u_0	$f(u_0)$	$\lvert f(u_0) - P_2(u_0)\rvert$
1	-2.	0.	2.	-1.	6.	0.75
2	-1.	0.	2.	1.25	9.75	0.375
3	-1.	0.	1.25	0.3203	4.0391	1.5428
4	0.	0.3203	1.25	0.6726	5.7084	0.9930
5	0.	0.3203	0.6726	0.4478	4.13432	0.1789
6	0.	0.3203	0.4478	0.3782	3.8654	0.1833
7	0.3203	0.3782	0.4478	0.3771	3.8687	0.0034

Tablica 3.3: Iterativni proces metode parabole

Zadatak 16. Primjenom metode parabole odredite aproksimaciju minimuma funkcije uz točnost $\varepsilon = 0.005$ ako je:

a) $f\colon [-2,1] \to \mathbb{R}$, $f(x) = -x - 3\frac{x-1}{x-2}$.

b) $f\colon [-2,2] \to \mathbb{R}$, $f(x) = (x+2)^3 - 4(x) - \ln(x+5)$.

Rješenje: a) $u^\star \approx \frac{41}{216} = 0.189815$ b) $u^\star \approx -0.827565$.

Brentova metoda

U primjenama se često koristi tzv. *Brentova metoda* za jednodimenzionalnu minimizaciju koja se zasniva na kombinaciji metode zlatnog reza i metode parabole. Metoda zlatnog reza pogodna je i za minimizaciju funkcije u najgorim mogućim slučajevima pri čemu se u minimizacijskom procesu metoda ponaša kao da "bježi" od minimuma. Međutim, zašto očekivati najgore? Ako je funkcija paraboličnog oblika u okolini minimuma, što je i slučaj za dovoljno glatke funkcije, tada će metoda parabole konvergirati minimumu u svega nekoliko koraka. S druge strane, ako je funkcija "nezgodna" tada je metoda zlatnog reza pouzdana metoda. Kompromis između dviju navedenih metoda sastoji se u tome da se obje koriste ovisno o tome koja se metoda u danom koraku učini pogodnijom. Međutim, prilikom kombiniranja tih dviju metoda mora se paziti na više stvari, izmešđu pstalog i na sljedeće:

(i) prilikom prelaska s jedne metode na drugu treba pripaziti na nepotrebno dodatno izračunavanje vrijednosti funkcije;

(ii) posebna pažnja mora se posvetiti zaustavnom kriteriju;

(iii) pravilo prema kojemu ćemo odlučivati koju metodu treba primijeniti mora biti robusno.

Metoda koja uzima u obzir sve navedno zove se Brentova metoda [6, 23]. Brentova metoda prilikom optimizacijskog procesa prati 6 točaka (ne nužno različitih): a, b, x, u, v i w. Minimum se nalazi unutar intervala (a, b); x je posljednja točka gdje je vrijednost funkcije bila najmanja; w je sljedeća točka iza x koja odgovara drugoj po veličini funkcijskoj vrijednosti; v je prethodna vrijednost od w, u je točka u kojoj je vrijednost funkcije izračunata zadnji puta; obično se u algoritmu prati i točka koja je u sredini intervala (a, b), ali se u njoj vrijednost funkcije ne mora računati. Navest ćemo pojedina načela koja su korisna. Interpolacija parabolom radi se kroz točke x, v i w, a kako bi korak dobiven metodom parabole bio prihvatljiv moraju se zadovoljiti:

(i) Sljedeća aproksimacija mora biti unutar intervala (a, b);

(ii) Udaljenost potencijalne sljedeće aproksimacije do x mora biti manja od polovine udaljenosti zadnje dvije aproksimacije;

Pravilo (ii) osigurava da metoda parabole konvergira ka nekoj točki umjesto da je ušla u neki nekonvergentni ciklički krug aproksimacija. U najgorem mogućem slučaju, gdje su aproksimacije dobivene metodom parabole prihvatljive ali beskorisne, metoda će približno alternirati s metode parabole na metodu zlatnog reza, a konvergirat će zbog metode zlatnog reza. Uspoređivanje udaljenosti između zadnjih aproksimacija čini se nužnim jer iskustvo pokazuje da je bolje ne "kažnjavati algoritam" za jedan loš korak ako se u sljedećem koraku situacija može popraviti. Nadalje, ako je u danoj točki vrijednost funkcije poznata, tada se u točkama, koje su unutar okoline radijusa $\epsilon > 0$, vrijednost funkcije ne računa ponovno jer će u vrijednosti funkcije tada dominirati pogreška pri njezinom računanju. U tom slučaju uvode se dodatne modifikacije za dobivanje potencijalne nove točke. Tipični zaustavni kriterij za Brentovu metodu je da su a i b udaljeni $2 \times x \times \epsilon$.

Detaljniji opis Brentove metode može se pronaći u [6, 13]. Također, u [23] može se naći cijeli algoritam Brentove metode.

Newtonova metoda

Za razliku od ranije razmatranih metoda, u ovom poglavlju kratko ćemo opisati jednu od najčešće korištenih metoda za određivanje lokalnog minimuma derivabilne funkcije. Kasnije ćemo detaljnije govoriti o ovoj metodi prilikom traženja lokalnog minimuma derivabilne funkcije više varijabli.

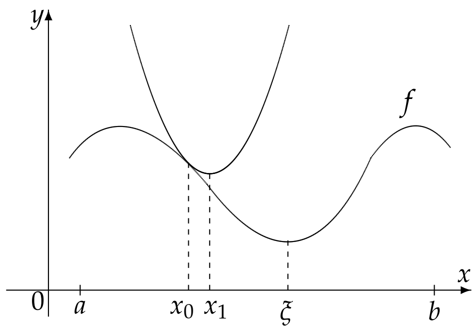
Pretpostavimo da je $f \in C^2[a,b]$ i da u točki $\xi \in (a,b)$ postiže lokalni minimum. To znači da je $f'(\xi) = 0$ i $f''(\xi) \geq 0$. Obratno, ako je za neki $\xi \in (a,b)$, $f'(\xi) = 0$ i $f''(\xi) > 0$, onda funkcija f u točki $\xi \in (a,b)$ postiže lokalni minimum. Točka $\xi \in (a,b)$ za koju vrijedi $f(\xi) = 0$ naziva se *stacionarnom točkom* funkcije f.

Iz navedenog slijedi da se traženje lokalnog minimuma derivabilne funkcije f može svesti na traženje stacionarne točke ili rješavanje jednadžbe $f'(x) = 0$. U tu svrhu možemo primijeniti poznatu Newtonovu metodu tangenti (vidi [26]).

Ipak, lokalni minimum derivabilne funkcije $f \in C[a,b]^2$ potražit ćemo direktnom minimizacijom. Ako postoji $\xi \in [a,b]$, prethodnom primjenom metode bisekcije možemo postići da na nekom suženom intervalu funkcija f postiže jedinstveni lokalni minimum. Zato, bez smanjenja općenitosti, možemo pretpostaviti da je $\xi \in [a,b]$ jedinstvena točka lokalnog minimuma funkcije f.

Izaberimo $x_0 \in (a,b)$ i na osnovi Taylorove formule funkciju f u okolini točke x_0 aproksimirajmo kvadratnom funkcijom

$$k_1(x) = f(x_0) + f'(x_0)(x - x_0) + \tfrac{1}{2}f''(x_0)(x - x_0)^2$$



Slika 6.5. Newtonova metoda

Sljedeću aproksimaciju x_1 tražene točke ξ birat ćemo tako da odredimo minimum kvadratne funkcije k_1

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)}.$$

Ponavljajući postupak dobivamo niz $x_0, x_1, ..., x_n, ...$ zadan rekurzivnom formulom

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

koji uz neke uvjete (vidi primjerice [10]) konvergira prema ξ. Ovo je tzv. *obična Newtonova metoda minimizacije*. Primijetite da se ona formalno podudara s Newtonovom metodom tangenti za rješavanje jednadžbe $f'(x) = 0$.

Niže navodimo jednostavni *Mathematica*-modul za Newtonovu metodu minimizacije. Iterativni proces zaustavit ćemo kada $|f'(x)| < \varepsilon$.

```
In[1]:= NewtonMin1[f_, x0_, eps_, it_] := Module[{x = N[x0], k = 0},
          While[Print[k, ": ", {x, f[x], f'[x]}];
                Abs[f'[x]] > eps && k < it,
                    x = x - f'[x]/f''[x];
                    k = k + 1
               ];
    If[f''[x] <= 0, Print["Pronadjena je stacionarna tocka"]];
                {x,f[x]} ];
```

Primjer 16. Razmotrimo problem određivanja euklidske udaljenosti točke $T_0 = (3, 4)$ do kubne parabole $q(x) = .5x^3 - 6x + 2$ (vidi *Primjer* 5.). Problem se svodi na minimizaciju derivabilne funkcije $f(x) = (x-3)^2 + (q(x)-4)^2$ na intervalu $[3, 4]$.

k	x_k	$f(x_k)$	$f'(x_k)$
0	4.	37.	218.
1	3.7254	2.7759	45.9025
2	3.6287	0.4094	4.5238
3	3.6169	0.3824	0.0625
4	3.6168	0.3825	0.00001

Tablica 3.4: Newtonova metoda

Uz početnu aproksimaciju $x_0 = 4$ (za $x_0 = 3$ proces ne konvergira!) i točnost $\varepsilon = 0.005$ tijek iterativnog procesa prikazan je u tablici 3.4. To znači da je udaljenost točke T_0 do grafa kubne parabole $q(x) = .5x^3 - 6x + 2$ jednaka $d(T_0, q) = \sqrt{f(x^\star)} = 0.618416$, pri čemu se taj minimum postiže u točki $T^\star(x^\star, q(x^\star)) = (3.61676, 3.95472)$.

Pokušajte odgovarajućim izborom početne aproksimacije pronaći druge lokalne minimume.

Lokalni minimum iste funkcije možemo uspješno potražiti direktnom primjenom *Mathematica*-modula: `FindMinimum` ili `NMinimize`. Primjena *Mathematica*-modula: `NMinimize` ne daje globalni minimum funkcije f.

Zadatak 17. Odredite l_2 udaljenost točke $T_0 = (2, 1)$ do funkcije $q(x) = x^3 + x^2 - 2x$ pomoću Newtonove metode tangenti uz toleranciju $\epsilon = 0.05$.

Rješenje: $x^\star \approx 1.27278$, aproksimacija l_2 udaljenosti je 0.739872.

Zadatak 18. Prethodni zadatak riješite tako da metodom zlatnog reza odredite l_1 udaljenost uz $\epsilon = 0.05$.

Rješenje: $x^\star \approx 1.23607$, aproksimacija l_1 udaljenosti je 0.81966.

POGLAVLJE 4

Problem višedimenzionalne minimizacije

Iterativni proces za traženje točke lokalnog minimuma dovoljno puta neprekidno diferencijabilne funkcije $f : \mathbb{R}^n \to \mathbb{R}$ općenito je oblika

$$x_{k+1} = x_k + \alpha_k p_k, \quad k = 0, 1, \ldots,$$

gdje je x_0 *početna aproksimacija*, p_k *vektor smjera kretanja* od točke x_k u točku x_{k+1}, a $\alpha_k > 0$ *duljina koraka* u smjeru vektora p_k.

Kretanje od točke x_k u točku x_{k+1} treba ostvariti tako da se postigne smanjenje vrijednosti funkcije, tj. tako da bude $f(x_{k+1}) < f(x_k)$. To se može postići tako da vektor p_k izaberemo tako da bude

$$(\nabla f(x_k), p_k) < 0, \tag{4.1}$$

gdje je $\nabla f(x_k)$ gradijent funkcije f u točki x_k. Štoviše, sljedeća lema daje dokaz navedene tvrdnje.

Lema 5. *[20] Neka je funkcional $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, G–diferencijabilan u nekoj točki $x \in \operatorname{Int}(D)$ 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. Definiramo pomoćnu funkciju $\varphi : [0, +\infty\rangle \to \mathbb{R}$, $\varphi(\alpha) = f(x + \alpha p)$, za koju je

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

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 $\alpha \in (0, \delta)$. $\square$

Navedimo nekoliko primjera.

(a) Gradijentna metoda: $p_k := -\nabla f(x_k)$. Vrijedi

$$f'(x_k)p_k = -f'(x_k)\nabla f(x_k) = -\|\nabla f(x_k)\|^2 < 0.$$

(b) Pokoordinatna relaksacija: izaberemo onaj bazni vektor $e^i \in \{e^1, \ldots, e^n\}$ 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 $\{e^1, \ldots, e^n\}$ možemo izabrati neke druge bazne vektore iz $\mathbb{R}^n$.

(c) Newtonova metoda: $p_k := -[f''(x_k)]^{-1}\nabla f(x_k)$. 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).

(d) Kvazi-Newtonova metoda: $p_k := -B^{-1}\nabla f(x_k)$, gdje je $B > 0$ pozitivno definitna matrica koja "nalikuje" $f''(x_k)$

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

4.1 Metode izbora duljine koraka

Prema *lemi* 5., 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

$$x_{k+1} = x_k + \alpha_k p_k$$

uvijek može odabrati duljina koraka $\alpha_k > 0$, tako da bude $f(x_{k+1}) < f(x_k)$ [4, 20]. Navest ćemo nekoliko najpoznatijih metoda izbora duljine koraka.

Metoda najbržeg spusta

Optimalni izbor duljine koraka dobiven postupkom jednodimenzionalne minimizacije

$$f(x_k + \alpha_k p_k) = \min_{\alpha>0}\{f(x_k + \alpha p_k) : x_k + \alpha p_k \in D\}, \tag{4.2}$$

potječe još od Cauchyja, a nazivamo ga *Metodom najbržeg spusta*.

Načelo H. Curry–M. Altmana

Definirajmo funkciju $\varphi : [0, +\infty\rangle \to \mathbb{R}$,

$$\varphi(\alpha) = f(x_k + \alpha p_k), \tag{4.3}$$

i potražimo njezine stacionarne točke rješavanjem jednadžbe po α

$$f'(x_k + \alpha p_k)\, p_k = 0. \tag{4.4}$$

Ako je f strogo konveksna na D, onda jednadžba (4.4) ima jedinstveno rješenje koje se podudara s korakom dobivenim u (4.2). Primijetite da zbog uvjeta $f'(x_k)\, p_k < 0$, nije prihvatljivo rješenje $\alpha = 0$. Jednadžba (4.4) može općenito imati više rješenja.

Slika 4.1. Načelo Curry–Altmana

Prema *načelu Curryja*, ako je $f'(x_k)p_k < 0$, za duljinu koraka α_k *treba izabrati najmanji pozitivni korijen* $\hat{\alpha}$ jednadžbe (4.4). 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[1]

$$f(x_k + \hat{\alpha} p_k) - f(x_k) = \hat{\alpha} \int_0^1 f'(x_k + \hat{\alpha} t p_k) p_k dt < 0,$$

[1] Ako je $F'(x + t(y-x))$ neprekidna po t, vrijedi $F(y) - F(x) = \int_0^1 F'(x + t(y-x))(y-x)dt$.

zbog $f'(x_k + \alpha p_k)\, p_k < 0$ za sve $\alpha \in [0, \hat{\alpha})$.

Prema *načelu M. Altmana*, za neki fiksni $\mu \in [0,1]$ za α_k *treba izabrati najmanji pozitivni korijen* jednadžbe

$$[f'(x_k + \alpha p_k) - \mu f'(x_k)]\, p_k = 0. \tag{4.5}$$

Ako je $\mu = 0$, načelo se svodi na načelo Curry. Ako je $\mu > 0$, duljina koraka α_k prvi je pozitivni broj $\alpha > 0$ za kojeg je $f'(x_k + \alpha p_k)\, p_k$ jednak μ-tom dijelu svoje početne veličine $f'(x_k)\, p_k$. Načelo Altmana ima praktično značenje, ali i važno teorijsko značenje jer se može iskoristiti kod ispitivanja drugih metoda izbora duljine koraka.

Aproksimativna minimizacija i traženje nultočaka

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

Prema (2.9), za svaki $x \in D$ vrijedi

$$f(x) = f(x_k) + f'(x_k)(x - x_k) + \tfrac{1}{2}(x - x_k)^T f''(x_k)(x - x_k) + o(\|x - x_k\|^2).$$

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

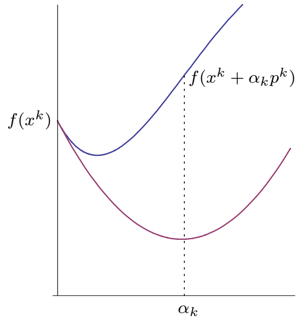
$$f_k(x) = f(x_k) + f'(x_k)(x - x_k) + \tfrac{1}{2}(x - x_k)^T f''(x_k)(x - x_k). \tag{4.6}$$

Primjenjujući ovu formulu na funkciju (4.3) dobivamo

$$\varphi(\alpha) = f(x_k + \alpha p_k) \approx f(x_k) + \alpha f'(x_k) p_k + \tfrac{1}{2}\alpha^2 (p_k)^T f''(x_k) p_k. \tag{4.7}$$

Ako je $(p_k)^T f''(x_k) p_k > 0$, onda je α_k korektno određen iz formule

$$\alpha_k = \tfrac{-f'(x_k) p_k}{(p_k)^T f''(x_k) p_k} > 0. \tag{4.8}$$



Slika 4.2. Aproksimacija minimuma

Budući da je često teško izračunati veličine $f'(x_k), f''(x_k)$, umjesto minimizacije funkcije (4.7) možemo definirati interpolacijski polinom drugog reda, koji se s funkcijom (4.3) podudara u tri točke: u_1, u_2, u_3 (obično se uzima $u_1 = 0$) i α_k određuje kao točku minimuma ovog interpolacijskog polinoma.

Kao što se vidi na *slici* 4.2., vrijednost duljine koraka (4.8) ne mora osiguravati smanjenje funkcijske vrijednosti. Ali zato prema *lemi* 5. možemo izabrati parametar $\omega_k > 0$, tako da bude

$$f(x_k + \omega_k \alpha_k p_k) < f(x_k).$$

Zadatak 19. ([20], 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 ψ 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

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

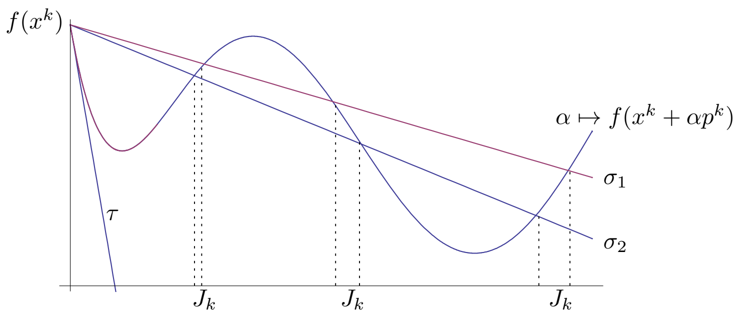
Napišite izraze za ψ i $\bar{\alpha}$ za specijalne slučajeve: (a) $u_1 = 0, u_3 = 2u_2$, (b) $u_2 = 0, u_1 = -u_3$. Opišite ovo načelo na nekoliko primjera uz primjenu gradijentne metode.

Načelo A. Goldsteina

Neka je f diferencijabilna u točki $x_k \in \operatorname{Int} D$ 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* 4.3.).



Slika 4.3. Načelo Goldsteina

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

$$\sigma_i(\alpha) = f(x_k) + \mu_i \alpha f'(x_k)p_k, \qquad i = 1, 2$$

i za njih vrijedi

$$\tau(\alpha) < \sigma_2(\alpha) \le \sigma_1(\alpha), \qquad \forall\, \alpha > 0.$$

Naime, zbog $f'(x_k)p_k < 0$ i $0 < \mu_1 \le \mu_2 < 1$ za koeficijente smjerova vrijedi

$$\mu_1 f'(x_k)p_k \ge \mu_2 f'(x_k)p_k > f'(x_k)p_k.$$

Definirajmo sada skup

$$J_k = \{\alpha > 0 : [x_k, x_k + \alpha p_k] \subset D, \quad \sigma_2(\alpha) \le f(x_k + \alpha p_k) \le \sigma_1(\alpha)\}.$$

Primijetite da za $\alpha \in J_k$ zbog $-\sigma_1(\alpha) \le -f(x_k + \alpha p_k) \le -\sigma_2(\alpha)$ vrijedi

$$0 < -\mu_1 \alpha f'(x_k)p_k \le f(x_k) - f(x_k + \alpha p_k) \le -\mu_2 \alpha f'(x_k)p_k.$$

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 α_k zadovoljava jednakost

$$f(x_k + \alpha_k p_k) - \mu \alpha_k f'(x_k) p_k = f(x_k).$$

Prvobitno je Goldsteinovo načelo predloženo u sljedećem obliku: Neka su $0 < \mu_1 \leq \mu_2 < 1$ fiksne konstante. Ako je za neki $\alpha_k > 0$,

$$f(x_k) - f(x_k + \alpha_k p_k) \geq -\mu_1 \alpha_k f'(x_k) p_k. \tag{4.9}$$

onda kao sljedeću aproksimaciju biramo točku $x_{k+1} = x_k + \alpha_k p_k$. Ako nije ispunjeno (4.9), izaberemo parametar ω_k tako da bude

$$0 < -\mu_1 \omega_k \alpha_k f'(x_k) p_k \leq f(x_k) - f(x_k + \omega_k \alpha_k p_k) \leq -\omega_k \alpha_k f'(x_k) p_k. \tag{4.10}$$

i kao sljedeću aproksimaciju biramo točku $x_{k+1} = x_k + \omega_k \alpha_k p_k$.

Ovakav parametar ω_k može se uvijek izabrati. Naime, ako nije ispunjeno (4.9), tj. ako za neki $\mu_1 \in (0,1)$ vrijedi

$$f(x_k) - f(x_k + \alpha_k p_k) < -\mu_1 \alpha_k f'(x_k) p_k, \tag{4.11}$$

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

$$\psi(\omega) = \tfrac{f(x_k) - f(x_k + \omega \alpha_k p_k)}{-\omega \alpha_k f'(x_k) p_k}, \qquad \omega \in (0,1].$$

Funkcija ψ neprekidna je na $[0,1]$ jer primjenom L'Hôpitalovog pravila dobivamo $\lim\limits_{\omega \to 0} \psi(\omega) = 1$. Zbog (4.11), $\psi(1) < \mu_1$, a tada zbog neprekidnosti funkcije ψ ona na $[0,1]$ prima sve vrijednosti između μ_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 (4.10).

Primjedba 4. Brojeve ω_k moguće je birati kao članove niza $\left(\frac{1}{q^j}\right)$, $q > 1$ i o tome postoji posebno razrađena analiza.

Zadatak 20. Neka je funkcional $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, G–diferencijabilan u nekoj točki $x_k \in \operatorname{Int}(D)$ 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 se uvjet postiže minimum dobivene kvadratne funkcije q? Odredite minimizator kvadratne funkcije q. Poznavajući ovu kvadratnu funkciju konstruirajte algoritam odabira duljine koraka α_k.

4.2 Gradijentna metoda

Još je 1845. godine A. Cauchy uočio da je smjer najbržeg pada neprekidno derivabilne funkcije $f : \mathbb{R}^n \to \mathbb{R}$ u točki x u smjeru antigradijenta $(-\nabla f(x))$. Naime, ako za neki vektor $d \in \mathbb{R}^n$ definiramo diferencijabilnu funkciju $\varphi : \mathbb{R} \to \mathbb{R}$ formulom

$$\varphi(t) := f(x + td),$$

onda zbog $\varphi'(0) = (\nabla f(x), d) \leq \|\nabla f(x)\| \|d\|$ iz čega slijedi da:

- za $d := \frac{\nabla f(x)}{\|\nabla f(x)\|}$, $\varphi'(0) = \|\nabla f(x)\|$ je najveće, pa smjer gradijenta pokazuje smjer najbržeg porasta funkcije f;
- za $d := -\frac{\nabla f(x)}{\|\nabla f(x)\|}$, vrijedi $\varphi'(0) = -\|\nabla f(x)\|$ je najmanje, pa smjer antigradijenta pokazuje smjer najbržeg pada funkcije f.

Na toj ideji zasniva se *obična gradijentna metoda* za minimizaciju funkcije $f : \mathbb{R}^n \to \mathbb{R}$

$$x_{k+1} = x_k + \alpha_k p_k, \quad k = 0, 1, \ldots, \tag{4.12}$$

gdje je $\alpha_k > 0$ *duljina koraka* u smjeru *vektora smjera kretanja* $p_k = -\nabla f(x_k)$, koji očigledno zadovoljava uvjet (4.1). Duljinu koraka α_k računamo sljedećim algoritmom (koji će biti opravdan u dokazu *Leme* 6. i *teorema* 5.):

Algoritam 1 . Algoritam za izračunavanje duljine koraka

Korak 0. Izabrati $0 < \varepsilon < 1$, $h \in (0, 1)$ i staviti $\alpha = 1$;
Korak 1. Izračunati $x = x_k + \alpha p_k$ i provjeriti uvjet

$$f(x) - f(x_k) \leq \varepsilon\alpha(\nabla f(x_k), p_k)) \tag{4.13}$$

Korak 2. Ako je uvjet (4.13) ispunjen, staviti $\alpha_k = \alpha$; inače staviti $\alpha = h\alpha$ i prijeći na *Korak 1.*

Najprije ćemo sagraditi *Mathematica*-modul `Korak[n,f,d,pk,xk]`, koji će izračunavati duljinu koraka α_k iz točke x_k) u smjeru vektora p_k). Pri tome je n broj varijabli, f ime minimizirajuće funkcije, a d njezin gradijent.

```
In[1]:= Korak[n_, f_, pk_, xk_] :=
              Module[{epsilon = .5, h = .5, f0, f1, fd},
              al = 1;   w = Table[x[i] -> xk[[i]], {i, n}];
              d[x_] := Table[D[f[x], x[i]], {i, n}];
              f0 = f[x] /. w;
              fd = d[x] /. w;
    While[f1 = f[x] /. Table[x[i] -> xk[[i]] + al pk[[i]], {i, n}];
               f1 - f0 > epsilon al Apply[Plus, fd pk], al = h al];
                                            al]
```

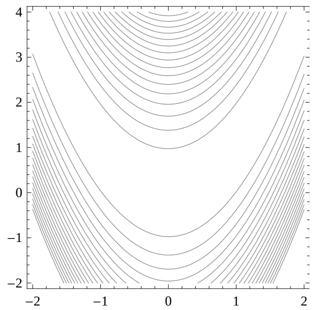
Gradijentnu metodu implementirat ćemo u *Mathematica*-modulu kao proceduru `MinGrad[n,f,d,x0,eps,it]`, gdje je x_0 početna aproksimacija, *eps* tražena točnost, a *it* maksimalno dozvoljen broj iteracija. Modul `MinGrad` poziva modul `Korak`, a iterativni postupak završava onda ako je $\|\nabla f(x_k)\|_\infty < eps$ za neki k ili ako je $k > it$.

```
In[2]:=
MinGrad[n_, f_, x0_, eps_, it_]:= Module[{k = 0, xs = x0, xn, gr, w},
    w = Table[x[i] -> x0[[i]], {i, n}];
    d[x_] := Table[D[f[x], x[i]], {i, n}];
While[ w = Table[x[i] -> xs[[i]], {i, n}]; grad = Max[Abs[d[x] /. w]];
        grad > eps && k < it,
        pk = -d[x] /. w;
        alpha = Korak[n, f, pk, xs];
Print["x", k, "=", xs, "   f=", f[x] /. Table[x[i] -> xs[[i]], {i, n}],
               "   |grad|=", grad, "   alpha=", alpha];
           xn = xs + alpha pk;
           xs = xn; k = k + 1;
              ];
                {xn, f[x] /. Table[x[i] -> xn[[i]], {i, n}]}]
```

Modul `MinGrad` testirat ćemo na poznatoj *Rosenbrockovoj test funkciji* koja postiže minimum u točki $x^\star = (1,1)$ i za koju ćemo odmah definirati gradijent i nacrtati nivo krivulje u okolini minimuma, a djelomične rezultate izvođenja programa za $x_0 = (0.5, 0.5)$ dati u tablici 4.1.

```
In[3]:= n = 2;
f[x_] := 100(x[2] - x[1]^2)^2 + (1 - x[1])^2
f[x] /. {x[1] -> x1, x[2] -> x2}
d[x_] := Table[D[f[x], x[i]], {i, n}]
d[x] /. {x[1] -> x1, x[2] -> x2}
ContourPlot[f[x] /. {x[1] -> x1, x[2] -> x2}, {x1, -2, 2}, {x2, -2, 4},
Contours -> 20, ContourShading -> None];

Out[3]:= (1 - x1)^2 + 100(-x1^2 + x2)^2
{-2(1 - x1) - 400 x1(-x1^2 + x2), 200(-x1^2 + x2)}
```



Slika 4.4. Rosenbrockova funkcija $f(x_1, x_2) = (1 - x_1)^2 + 100(-x_1^2 + x_2)^2$

k	(x_k)	$f(x_k)$	$\|\nabla f(x_k)\|_\infty$	α_k
0	(0.5,0.5)	6.5	51	0.00195313
1	(0.599609, 0.402344)	0.343602	11.0691	0.00195313
2	(0.621229, 0.38562)	0.143477	0.681796	0.00195313
10	(0.703733, 0.492322)	0.0886257	0.58356	0.00195313
100	(0.785646, 0.615674)	0.0461927	0.313171	0.00195313
500	(0.895571, 0.801784)	0.0109123	0.114267	0.00195313
1000	(0.984068, 0.968167)	0.000258777	0.0556747	0.03125

Tablica 4.1: Tijek gradijentne metode za Rosenbrockovu funkciju

Sljedeća lema pokazuje da gradijentna metoda (4.12) osigurava konvergenciju iterativnog procesa ili prema $\inf f$ ili prema nekoj stacionarnoj točki funkcije f.

Lema 6. *Pretpostavimo da*

(i) funkcija $f : \mathbb{R}^n \to \mathbb{R}$ ograničena je odozdo,

(ii) gradijent ∇f zadovoljava Lipschitzov uvjet, tj. postoji $L > 0$, tako da bude

$$\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|, \quad \text{za sve } x, y \in \mathbb{R}^n, \tag{4.14}$$

(iii) parametri α_k biraju se iz uvjeta (4.13).

Tada za gradijentnu metodu (4.12) vrijedi

$$\|\nabla f(x_k)\| \to 0, \text{ za } k \to \infty$$

neovisno o izboru početne aproksimacije x_0, a parametri α_k uvijek se mogu izabrati prethodno opisanim algoritmom za izračunavanje duljine koraka.

Dokaz. Prema Lagrangeovom teoremu srednje vrijednosti postoji $\vartheta \in (0,1)$, tako da bude

$$f(x_k + (x - x_k)) - f(x_k) = (\nabla f(x_k + \vartheta(x - x_k)), x - x_k).$$

Dodavanjem i oduzimanjem $f'(x_k)(x - x_k)$ na desnoj strani, uz oznake: $\xi := x_k + \vartheta(x - x_k)$, $f_k := f(x_k)$, $f'_k := \nabla f(x_k)$, $f'_\xi := \nabla f(\xi)$ ovu jednakost možemo zapisati na sljedeći način

$$f(x) - f_k = (f'_k, x - x_k) + (f'_\xi - f'_k, x - x_k).$$

Za $x = x_k - \alpha f'_k$ korištenjem Cauchy-Schwartzove nejednakosti i uvjeta (4.14) dobivamo

$$\begin{aligned} f(x) - f_k &= -\alpha(f'_k, f'_k) - \alpha(f'_\xi - f'_k, f'_k) \\ &\leq -\alpha\|f'_k\|^2 + \alpha\|f'_\xi - f'_k\|\|f'_k\| \leq -\alpha\|f'_k\|^2 + \alpha L\|\xi - x_k\|\|f'_k\| \\ &\leq -\alpha\|f'_k\|^2 + \alpha L\|x - x_k\|\|f'_k\| = -\alpha\|f'_k\|^2 + \alpha L\alpha\|f'_k\|\|f'_k\| \\ &= \alpha\|f'_k\|^2(\alpha L - 1). \end{aligned}$$

Kako je $0 < L < \infty$, za $\varepsilon \in (0,1)$ iz algoritma za izračunavanje duljine koraka parametar α možemo izabrati tako da bude $\alpha > 0$ i $\alpha L - 1 \leq -\varepsilon$. Zato iz prethodne nejednakosti dobivamo

$$f(x) - f_k \leq \alpha(f'_k, f'_k)(-\varepsilon) \Longrightarrow f(x) - f_k \leq \varepsilon\alpha(f'_k, -f'_k), \tag{4.15}$$

što je upravo uvjet (4.13). Drugim riječima, uvjet (4.13) bit će ispunjen za $\alpha \leq \frac{1-\varepsilon}{L}$.

Ako stavimo $x := x_{k+1}$, onda prethodna nejednakost pokazuje da je $f_{k+1} < f_k$ za svaki $k = 0, 1, \ldots$ uz uvjet da je $\|f'_k\| \neq 0$, tj. niz (f_n) je monotono padajuć. Kako je po pretpostavci funkcija f ograničena odozdo, onda to znači da je i niz (f_n) ograničen odozdo. Dakle, niz (f_n) je konvergentan, a onda i Cauchyjev, pa možemo zaključiti da vrijedi

$$f_{k+1} - f_k \to 0 \qquad \text{za } k \to \infty.$$

Iz nejednakosti (4.15) slijedi $f_{k+1} - f_k \leq -\varepsilon\alpha\|f'_k\|^2$, odnosno

$$\|f'_k\|^2 \leq \frac{f_k - f_{k+1}}{\varepsilon\alpha_k}.$$

Primijetimo da niti za jedan k parametar α_k nije jednak nuli. Naime, izbor parametra α_k jamči da će biti $\alpha_k \geq \bar{\alpha} > 0$, gdje za $\bar{\alpha}$ može poslužiti bilo koja konstanta manja ili jednaka od $\frac{1-\varepsilon}{L}$. □

Teorem 5. *Neka je $f \in C^2(\mathbb{R}^n)$ i*

$$m\|y\|^2 \leq (f''(x)y, y) \leq M\|y\|^2, \quad M > m > 0 \qquad \text{za sve } x, y \in \mathbb{R}^n \tag{4.16}$$

Ako se niz (x_n) definira pomoću iterativnog procesa (4.12) a parametri α_k biraju iz uvjeta (4.13), tada neovisno o izboru početne aproksimacije $x_0 \in \mathbb{R}^n$ vrijedi

$$x_k \to x^\star \quad i \quad f(x_k) \to f(x^\star),$$

gdje je $x^\star \in \mathbb{R}^n$ jedinstvena točka minimuma funkcije f. Pri tome vrijede ocjene:

$$\|x_k - x^\star\| \le \frac{\|\nabla f(x_k)\|}{m}, \qquad f_k - f^\star \le q^k (f_0 - f^\star), \qquad \|x_k - x^\star\| \le C\, q^{\frac{k}{2}}, \quad (4.17)$$

gdje je $C < \infty$, $\quad q \in (0,1)$.

Dokaz. Prema *teoremu 1.* lijevi dio nejednakosti (4.16) ekvivalentan je činjenici da je f jako konveksna funkcija. Nadalje, prema *teoremu 2.* to znači da je funkcija f ograničena odozdo i da postoji jedinstvena točka minimuma $x^\star \in \mathbb{R}^n$. Zato još treba dokazati da niz (x_n) konvergira prema $x^\star$ neovisno o izboru početne aproksimacije $x_0 \in \mathbb{R}^n$ te dokazati ocjene (4.17).

Pokažimo najprije da niz (x_n) konvergira prema $x^\star$. Primjenom Taylorove formule za funkciju f u okolini točke x, za $x^\star$ dobivamo

$$f(x^\star) = f(x) + (\nabla f(x), x^\star - x) + \tfrac{1}{2}\left(f''(\xi)(x^\star - x), x^\star - x\right)$$

a odavde korištenjem (4.16) dobivamo

$$\begin{aligned} f(x) - f(x^\star) =& (\nabla f(x), x - x^\star) - \tfrac{1}{2}\left(f''(\xi)(x^\star - x), x^\star - x\right) \\ \le& \ (\nabla f(x), x - x^\star) - \tfrac{m}{2}\|x - x^\star\|^2, \end{aligned}$$

što uz primjenu Cauchyjeve nejednakosti daje

$$f(x) - f(x^\star) \le \|\nabla f(x)\| \|x - x^\star\| - \tfrac{m}{2}\|x - x^\star\|^2. \qquad (4.18)$$

Primjenom Taylorove formule za funkciju f u okolini točke $x^\star$ i korištenjem činjenice da je $\nabla f(x^\star) = 0$, dobivamo

$$f(x) - f(x^\star) = \tfrac{1}{2}\left(f''(\zeta)(x - x^\star), x - x^\star\right).$$

Koristeći ovdje (4.16) dobivamo

$$\tfrac{m}{2}\|x - x^\star\|^2 \le f(x) - f(x^\star) \le \tfrac{M}{2}\|x - x^\star\|^2. \qquad (4.19)$$

Iz (4.18) i (4.19) slijedi

$$\begin{aligned} & \|\nabla f(x)\| \|x - x^\star\| - \tfrac{m}{2}\|x - x^\star\|^2 \ge f(x) - f(x^\star) \ge \tfrac{m}{2}\|x - x^\star\|^2, \\ \Longrightarrow \quad & m\|x - x^\star\|^2 \le \|\nabla f(x)\| \|x - x^\star\|, \end{aligned}$$

a odavde

$$\|x - x^\star\| \le \frac{\|\nabla f(x)\|}{m}. \qquad (4.20)$$

Desna strana ove nejednakosti za $x = x_k$ prema *lemi 6.* konvergira prema nuli, što znači da $\|x_k - x^\star\| \to 0$ za $k \to \infty$, tj. $x_k \to x^\star$.

Uvrštavanjem nejednakosti (4.19) i (4.20) u (4.18) dobivamo

$$f(x) - f(x^\star) \le \|\nabla f(x)\| \frac{\|\nabla f(x)\|}{m} - \frac{m}{2}\frac{2}{M}(f(x) - f(x^\star)),$$

iz čega slijedi

$$\frac{1}{m}\|\nabla f(x)\|^2 \geq (f(x) - f(x^\star)) + \frac{m}{M}(f(x) - f(x^\star)),$$

odnosno

$$\|\nabla f(x)\|^2 \geq m\left(1 + \frac{m}{M}\right)(f(x) - f(x^\star)). \tag{4.21}$$

Uvrštavanjem (4.21) u nejednakost (4.15) za $x = x_k - \alpha\nabla f_k$ dobivamo (prisjetimo se da (4.15) glasi $f(x) - f(x_k) \leq -\varepsilon\alpha\|\nabla f_k\|^2$)

$$f_{k+1} - f_k \leq -\varepsilon\alpha_k m\left(1 + \frac{m}{M}\right)(f(x_k) - f(x^\star)). \tag{4.22}$$

S druge strane, primjenom Taylorove formule za funkciju f u okolini točke x_k, za $x = x_k - \alpha\nabla f_k$ dobivamo

$$\begin{aligned} f(x) - f(x_k) &= (\nabla f_k, x - x_k) + \frac{1}{2}\left(f''(\xi)(x - x_k), x - x_k\right) \\ &= -\alpha\|\nabla f_k\|^2 + \frac{\alpha^2}{2}\left(f''(\xi)\nabla f_k, \nabla f_k\right) \end{aligned}$$

Primjenom (4.16) na gornju jednakost dobivamo:

$$f(x) - f(x_k) \quad\leq\quad -\alpha\|\nabla f_k\|^2 + \frac{\alpha^2}{2}M\|\nabla f_k\|^2 = -\alpha\left(1 - \frac{\alpha M}{2}\right)\|\nabla f_k\|^2.$$

Ako $\varepsilon > 0$ izaberemo tako da $1 - \frac{\alpha M}{2} > \varepsilon$, onda iz prethodne nejednakosti dobivamo uvjet (4.13) pomoću kojeg biramo duljinu koraka α_k

$$f(x) - f(x_k) \leq -\alpha\varepsilon\|\nabla f_k\|^2.$$

Dakle, uvjet (4.13) bit će ispunjen za $\alpha_k \leq \overline{\alpha} = \frac{2(1-\varepsilon)}{M}$.

Iz (4.22) slijedi $\left[f_{k+1} - f_k \leq -\varepsilon\alpha_k m\left(1 + \frac{m}{M}\right)(f(x_k) - f(x^\star))\right]$

$$\begin{aligned} f_{k+1} - f^\star &= (f_{k+1} - f_k) + (f_k - f^\star) \leq \left(1 - \varepsilon\alpha_k m\left(1 + \frac{m}{M}\right)\right)(f(x_k) - f(x^\star)) \\ &\leq q(f_k - f^\star), \end{aligned}$$

uz oznaku $q := 1 - \varepsilon\alpha_k m\left(1 + \frac{m}{M}\right)$, što primijenjeno više puta daje

$$f_k - f^\star \leq q^k(f_0 - f^\star). \tag{4.23}$$

Kako je $\alpha_k \leq \overline{\alpha} = \frac{2(1-\varepsilon)}{M}$, bit će $q \geq 1 - \frac{2\varepsilon(1-\varepsilon)m}{M}\left(1 + \frac{m}{M}\right)$. Ako sada q shvatimo kao realnu funkciju varijable ε i pronađemo njezin minimum, dobivamo $q_{min} = 1 - \frac{m}{2M}\left(1 + \frac{m}{M}\right)$ za $\varepsilon = \frac{1}{2}$. Sada se vidi da je u uvjet (4.13) najbolje staviti $\varepsilon = \frac{1}{2}$. To će osigurati najveću moguću brzinu konvergencije gradijentne metode.

Na kraju, iz (4.19) i (4.23) dobivamo ocjenu pogreške

$$\|x_k - x^\star\| \leq \sqrt{\frac{2}{m}}\sqrt{f_k - f^\star} \leq q^{\frac{k}{2}}\sqrt{\frac{2}{m}}\sqrt{f_0 - f^\star} \leq Cq^{\frac{k}{2}}.$$

$\square$

Zadatak 21. Uz primjenu gradijentne metode odredite aproksimaciju minimuma funkcije $f(x_1, x_2, x_3) = (x_1 - 4)^4 + (x_2 - 3)^2 + 4(x_3 + 5)^4$ uz početnu aproksimaciju $x_0 = [3,\ 2,\ -5]^T$ i točnost $\epsilon = 0.5$.

Rješenje: $x^\star \approx [3.7072,\ 2.78125,\ -5]^T$.

Zadatak 22. Koristeći Matlab uz primjenu gradijentne metode odredite aproksimaciju minimuma funkcije $f(x, y) = 2x^3 + xy^2 + 5x^2 + y^2$ uz početnu aproksimaciju $[-1,\ -1]^T$ i točnost $\epsilon = 0.005$.

Rješenje: $x^\star \approx [0.0003, -0.0018]^T$.

Zadatak 23. Pokažite da niz aproksimacija dobiven gradijentnom metodom za funkciju $f(x_1, x_2) = -19x_1 + 4x_1^2 + 16x_2 - 3x_1 x2 + 5x_2^2$ konvergira neovisno o izboru početne aproksimacije x_0.

Metoda najbržeg spusta

Iterativnu metodu

$$x_{k+1} = x_k + \alpha_k p_k, \quad k = 0, 1, \ldots, \tag{4.24}$$

gdje je duljina koraka $\alpha_k > 0$ u smjeru vektora smjera kretanja $p_k = -\nabla f(x_k)$ definiramo kao

$$\alpha_k = \underset{\alpha \geq 0}{\operatorname{argmin}} f\left(x_k + \alpha \nabla f(x_k)\right), \tag{4.25}$$

nazivamo *Metodom najbržeg spusta*[2]. Jednodimenzionalnu minimizaciju u (4.25) možemo provesti nekom od metoda za jednodimenzionalnu minimizaciju (metoda zlatnog reza, metoda parabole, Brentova metoda).

Metodu ćemo ispitati uz korištenje Rosenbrockove test funkcije. Djelomični rezultati vide se u tablici 4.2.

k	(x_k)	$f(x_k)$	$\|\nabla f(x_k)\|_\infty$	α_k
0	(0.5,0.5)	6.5	51	0.00195313
1	(0.501683, 0.49835)	6.33264	50.4955	0.00235693
2	(0.503364, 0.496708)	6.1677	49.9871	0.00235024
10	(0.516722, 0.483858)	4.9362	45.7882	0.00229504

Tablica 4.2: Metoda najbržeg spusta za Rosenbrockovu funkciju

Metoda najbržeg spusta, kao i gradijentna metoda, imaju isključivo teorijsko značenje i rijetko se koriste u praktičnim primjenama zbog niske brzine konvergencije, što potvrđuje i navedeni jednostavni primjer s *Rosenbrockovom funkcijom*.

[2]eng.: Steepest Descent Method

4.3 Newtonova metoda

Iterativni proces za traženje točke lokalnog minimuma dovoljno puta neprekidno diferencijabilne funkcije $f : \mathbb{R}^n \to \mathbb{R}$ općenito je oblika

$$x_{k+1} = x_k + \alpha_k p_k, \quad k = 0, 1, \ldots, \tag{4.26}$$

gdje je x_0 *početna aproksimacija*, p_k *vektor smjera kretanja* od točke x_k u točku x_{k+1}, a $\alpha_k > 0$ *duljina koraka* u smjeru vektora p_k.

Kretanje od točke x_k u točku x_{k+1} treba ostvariti tako da se postigne smanjenje vrijednosti funkcije, tj. tako da bude $f(x_{k+1}) < f(x_k)$. To se može postići tako da vektor p_k izaberemo tako da bude (vidi primjerice [15], [22])

$$(\nabla f(x_k), p_k) < 0, \tag{4.27}$$

gdje je $\nabla f(x_k)$ gradijent funkcije f u točki x_k. Naime, ako za funkciju f napišemo Taylorovu formulu u okolini točke x_k

$$f(x) = f(x_k) + (\nabla f_k, x - x_k) + \tfrac{1}{2}\left(f''\left(x_k + \vartheta(x - x_k)\right)(x - x_k), x - x_k\right), \ \vartheta \in (0,1)$$

i stavimo $x = x_k + \alpha_k p_k$, uz oznaku $\xi_k = x_k + \vartheta(x - x_k)$ dobivamo

$$f(x_k + \alpha_k p_k) = f(x_k) + \alpha_k(\nabla f_k, p_k) + \tfrac{\alpha_k^2}{2}\left(f''(\xi_k)p_k, p_k\right), \qquad \vartheta \in (0,1)$$

Ako je $(\nabla f_k, p_k) < 0$ i $f''(\xi_k)$ pozitivno definitna, onda za male vrijednosti parametara α_k vrijedi $f(x_{k+1}) := f(x_k + \alpha_k p_k) < f(x_k)$.

Pretpostavimo da je promatrana funkcija f strogo konveksna i dovoljno glatka i da je na neki način izabrana početna aproksimacija x_0. Analogno Newtonovoj metodi za jednodimenzionalnu minimizaciju u okolini točke x_0 funkciju f aproksimirat ćemo kvadratnom funkcijom dobivenom iz Taylorove formule

$$\psi(x) = f(x_0) + (\nabla f(x_0), x - x_0) + \tfrac{1}{2}(f''(x_0)(x - x_0), x - x_0)$$

Funkcija ψ strogo je konveksna ($f''(x_0)$ pozitivno je definitna) i njezin minimum postiže se u točki $x_1 = x_0 - [f''(x_0)]^{-1}\nabla f(x_0)$, koju ćemo smatrati sljedećom aproksimacijom. Primijetite da smo iz točke x_0 u točku x_1 "došli" preko *vektora smjera kretanja* definiranog s $p = -[f''(x_0)]^{-1}\nabla f(x_0)$. Ovaj vektor definira smjer spusta iz točke x_0 jer zbog pozitivne definitnosti Hessijana f ispunjava uvjet spusta (4.27)

$$(\nabla f(x_0), p) = (\nabla f(x_0), -[f''(x_0)]^{-1}\nabla f(x_0)) = -(f''(x_0)p, p) < 0.$$
$$\left[(\nabla f_0, p) = f_0' p = -f_0'\left([f_0'']^{-1} f_0''\right)[f_0'']^{-1}\nabla f_0 = -p^T f_0'' p = -(f_0'' p, p)\right]$$

Ponavljajući navedeni postupak dolazimo do iterativne metode

$$x_{k+1} = x_k - [f''(x_k)]^{-1}\nabla f(x_k), \qquad k = 0, 1, \ldots,$$

koja se u literaturi obično naziva *običnom Newtonovom* ili *Newton-Raphsonovom metodom* (vidi primjerice [10], [13], [15], [18], [20], [22], [28]). Pri tome se u implementaciji metode računanje inverzne matrice $[f''(x_k)]^{-1}$ izbjegava tako da umjesto toga u svakom koraku rješavamo jedan sustav linearnih jednadžbi. Iterativni postupak definiramo kao

$$x_{k+1} = x_k + p_k, \qquad \text{(gdje je } p_k \text{ rješenje sustava)}$$
$$f''(x_k)p_k = -\nabla f(x_k).$$

Uzevši u obzir da kretanje od točke x_k u smjeru vektora p_k donosi smanjenje vrijednosti funkcije, kao i u slučaju gradijentne metode, uvodimo *duljinu koraka* α_k i tako definiramo *Newtonovu metodu s regulacijom koraka*

$$x_{k+1} = x_k - \alpha_k[f''(x_k)]^{-1}\nabla f(x_k), \qquad k = 0, 1, \ldots, \tag{4.28}$$

koju u implementaciji definiramo kao

$$\begin{aligned} &x_{k+1} = x_k + \alpha_k p_k, \qquad \text{(gdje je } p_k \text{ rješenje sustava)} \\ &f''(x_k)p_k = -\nabla f(x_k). \end{aligned} \tag{4.29}$$

Kao i u slučaju gradijentne metode parametar duljine koraka α_k možemo definirati na više načina (vidi 4.1). Ovdje ćemo primijeniti specijalni slučaj Goldsteinovog principa opisanog u 4.1, str. 38, a posebno u primjedbi 4.

Algoritam 2 . Algoritam za izračunavanje duljine koraka

Korak 0. Staviti $\alpha = 1$, $h \in (0, 1)$ i izabrati $0 < \varepsilon < \frac{1}{2}$;
Korak 1. Izračunati $x = x_k + \alpha_k p_k$ i provjeriti uvjet

$$f(x) - f(x_k) \leq \varepsilon\alpha(\nabla f(x_k), p_k); \tag{4.30}$$

Korak 2. Ako je uvjet (4.30) ispunjen, staviti $\alpha_k = \alpha$; inače staviti $\alpha := h\alpha$ i prijeći na *Korak 1.*

Konvergencija Newtonove metode

Pretpostavit ćemo da je promatrana funkcija $f : \mathbb{R}^n \to \mathbb{R}$ dvostruko neprekidno diferencijabilna i jako konveksna, tj.

$$m\|y\|^2 \leq (f''(x)y, y) \leq M\|y\|^2, \quad M > m > 0, \qquad \forall x, y \in \mathbb{R}^n. \tag{4.31}$$

Prema *teoremu* 2., str.12., takva je funkcija ograničena odozdo i posjeduje jedinstvenu točku minimuma.

Može se pokazati da također vrijedi (vidi primjerice [15], [22])

$$\frac{m}{M^2}\|y\|^2 \leq ((f''(x)]^{-1}y,y) \leq \frac{1}{m}\|y\|^2, \quad M > m > 0 \qquad \text{za sve } x,y \in \mathbb{R}^n. \tag{4.32}$$

iz čega slijedi ocjena

$$(\nabla f_k, p_k) = -(\nabla f_k, [f''(x_k)]^{-1}\nabla f_k) \leq -\frac{m}{M^2}\|\nabla f_k\|^2. \tag{4.33}$$

Teorem 6. *Neka funkcija $f \in C^2(\mathbb{R}^n)$ zadovoljava uvjet (4.31) i neka se parametri duljine koraka α_k biraju iz uvjeta (4.30).*

Tada neovisno o izboru početne aproksimacije x_0 iterativni proces (4.28) konvergira prema jedinstvenoj točki minimuma $x^\star$ superlinearnom brzinom

$$\|x_{k+1} - x^\star\| \leq \lambda_k\|x_k - x^\star\|, \tag{4.34}$$

gdje $\lambda_k \to 0$ za $k \to \infty$.

Dokaz. Za $x = x_k + \alpha p_k$, gdje je $p_k = -[f_k'']^{-1}\nabla f_k$, korištenjem (4.31) u Taylorovoj formuli dobivamo

$$f(x) - f(x_k) = \alpha\,(\nabla f_k, p_k) + \frac{\alpha^2}{2}\,(f''(\xi)p_k, p_k) \leq \alpha\,(\nabla f_k, p_k)\left(1 + \frac{\alpha M\|p_k\|^2}{2(\nabla f_k, p_k)}\right)$$

Kako je s druge strane

$$\left[(\nabla f_k, p_k) = (\nabla f_k)^T p_k = (\nabla f_k)^T\,[f_k'']^{-1} f_k'' \quad p_k = -p_k^T f_k'' p_k\right]$$

$$(\nabla f_k, p_k) = -(f''(x_k)p_k, p_k) \leq -m\|p_k\|^2,$$

dobivamo

$$\left[-(\nabla f_k, p_k) \geq m\|p_k\|^2 \Rightarrow \left(1 + \frac{\alpha M\|p_k\|^2}{2(\nabla f_k, p_k)}\right) = \left(1 - \frac{\alpha M\|p_k\|^2}{-2(\nabla f_k, p_k)}\right) \leq \left(1 - \frac{\alpha M}{2m}\right)\right]$$

$$f(x) - f(x_k) \leq \alpha\,(\nabla f_k, p_k)\left(1 - \frac{\alpha M}{2m}\right).$$

Izaberimo $\varepsilon > 0$ tako da bude

$$1 - \frac{\alpha M}{2m} \geq \varepsilon. \tag{4.35}$$

Množenjem ove nejednakosti s negativnim brojem $\alpha(\nabla f_k, p_k) < 0$ dobivamo

$$\varepsilon\alpha(\nabla f_k, p_k) \geq \alpha(\nabla f_k, p_k)\left(1 - \frac{\alpha M}{2m}\right) \geq f(x) - f(x_k).$$

Dakle, uvjet (4.30) bit će ispunjen ako je ispunjen uvjet (4.35), odnosno ako je $\alpha \leq \bar{\alpha} = \frac{2(1-\varepsilon)m}{M}$. Time je opravdan način izbora duljine koraka α_k.

Kako je $(\nabla f_k, p_k) < 0$ za $\|\nabla f_k\| \neq 0$, iz uvjeta

$$f_{k+1} - f_k \leq \varepsilon\alpha_k(\nabla f_k, p_k) \tag{4.36}$$

slijedi $f_{k+1} < f_k$, tj. niz (f_k) monotono je padajuć. Zbog uvjeta jake konveksnosti funkcije f on je i ograničen odozdo, što znači da je konvergentan i da vrijedi $f_{k+1} - f_k \to 0$. Korištenjem uvjeta (4.32) u (4.36) i činjenice da je $f''(x_k)$ simetrična matrica dobivamo

$$f_{k+1} - f_k \leq \varepsilon\alpha_k(\nabla f_k, p_k) = -\varepsilon\alpha_k([f_k'']^{-1}\nabla f_k, \nabla f_k) \leq -\varepsilon\alpha_k \tfrac{m}{M^2}\|\nabla f_k\|^2, \quad (4.37)$$

a odavde

$$\|\nabla f_k\|^2 \leq \tfrac{M^2}{m\varepsilon\alpha_k}\left(f_k - f_{k+1}\right),$$

iz čega slijedi $\|\nabla f_k\| \to 0$. Time je pokazano da su rezultati analogni rezultatima *leme* 6. ostali sačuvani.

Nadalje, slično kao u dokazu *teorema* 5. pokazat ćemo da niz definiran s (4.28) konvergira prema točki minimuma $x^\star$. Budući da je f jako konveksna funkcija, postoji jedinstvena točka minimuma $x^\star \in \mathbb{R}^n$. Pokažimo još da niz (x_n) konvergira prema $x^\star$ neovisno o izboru početne aproksimacije $x_0 \in \mathbb{R}^n$.

Primjenom Taylorove formule za funkciju f u okolini točke x, za $x^\star$ dobivamo

$$f(x^\star) = f(x) + (\nabla f(x), x^\star - x) + \tfrac{1}{2}\left(f''(\xi)(x^\star - x), x^\star - x\right).$$

Nadalje, korištenjem uvjeta jake konveksnosti (4.31)

$$\left[(f''(\xi)(x^\star - x), (x^\star - x)) \geq m\|x^\star - x\|^2\right]$$

dobivamo

$$\begin{aligned} f(x) - f(x^\star) =& (\nabla f(x), x - x^\star) - \tfrac{1}{2}\left(f''(\xi)(x^\star - x), x^\star - x\right) \\ &\leq (\nabla f(x), x - x^\star) - \tfrac{m}{2}\|x - x^\star\|^2, \end{aligned}$$

što uz primjenu Cauchyjeve nejednakosti daje

$$f(x) - f(x^\star) \leq \|\nabla f(x)\|\|x - x^\star\| - \tfrac{m}{2}\|x - x^\star\|^2. \quad (4.38)$$

Primjenom Taylorove formule za funkciju f u okolini točke $x^\star$ i korištenjem činjenice da je $\nabla f(x^\star) = 0$, dobivamo

$$f(x) - f(x^\star) = \tfrac{1}{2}\left(f''(\zeta)(x - x^\star), x - x^\star\right).$$

Koristeći ovdje uvjet jake konveksnosti (4.31) dobivamo

$$\tfrac{m}{2}\|x - x^\star\|^2 \leq f(x) - f(x^\star) \leq \tfrac{M}{2}\|x - x^\star\|^2. \quad (4.39)$$

Iz (4.38) i (4.39) slijedi

$$\|\nabla f(x)\|\|x - x^\star\| - \tfrac{m}{2}\|x - x^\star\|^2 \geq f(x) - f(x^\star) \geq \tfrac{m}{2}\|x - x^\star\|^2,$$

a odavde $[m\|x - x^\star\|^2 \leq \|\nabla f(x)\|\|x - x^\star\|]$

$$\|x - x^\star\| \leq \tfrac{\|\nabla f(x)\|}{m}. \quad (4.40)$$

Kao što smo ranije utvrdili, desna strana ove nejednakosti za $x = x_k$ konvergira prema nuli, što znači da $\|x_k - x^\star\| \to 0$ za $k \to \infty$, tj. $x_k \to x^\star$.

Prijeđimo sada na dokaz ocjene brzine konvergencije (4.34). U tu svrhu najprije primijetimo da vrijedi

$$(\nabla f_k, p_k) = -(f''(x_k)p_k, p_k) \leq -m\|p_k\|^2$$

Iz (4.36) slijedi $(\nabla f_k, p_k) \to 0$, stoga koristeći gornju nejednakost imamo

$$\|p_k\|^2 \leq \frac{(\nabla f_k, p_k)}{-m} \to 0, \qquad k \to \infty \quad \Rightarrow \quad \|p_k\| \to 0, \qquad k \to \infty. \tag{4.41}$$

Pokažimo nadalje da su u općoj Newtonovoj metodi (4.31), počevši od neke iteracije, svi parametri α_k jednaki 1. Koristeći Taylorovu formulu te dodavanjem i oduzimanjem broja $\frac{\alpha_k^2}{2}(f_k'' p_k, p_k)$ dobivamo

$$\begin{aligned}
f_{k+1} - f_k &= \alpha_k(\nabla f_k, p_k) + \tfrac{\alpha_k^2}{2}(f_k'' p_k, p_k) + \tfrac{\alpha_k^2}{2}\left((f''(\xi) - f_k'')p_k, p_k)\right) \\
(4.41) &\leq \alpha_k(\nabla f_k, p_k)\left(1 - \tfrac{\alpha_k}{2} + \tfrac{\alpha_k(-1)}{2}\tfrac{\|(f''(\xi)-f_k'')p_k\|\cdot\|p_k\|}{(-1)(\nabla f_k, p_k)}\right) \\
(4.41) &\leq \alpha_k(\nabla f_k, p_k)\left(1 - \tfrac{\alpha_k}{2} - \tfrac{\alpha_k}{2}\tfrac{\|f''(\xi)-f_k''\|\cdot\|p_k\|^2}{m\|p_k\|^2}\right),
\end{aligned}$$

gdje je $\xi = x_k + \vartheta(x_{k+1} - x_k), \quad \vartheta \in [0,1]$.

Budući da

$$\|f''(\xi) - f''(x_k)\| \leq \|f''(\xi) - f''(x^\star)\| + \|f''(x^\star) - f''(x_k)\|,$$

i $\|x_k - x^\star\| \to 0$, odnosno $\|\xi - x^*\| \to 0$ za $k \to \infty$, također zbog neprekidnosti funkcije f'' vrijedi $\|f''(\xi) - f''(x_k)\| \to 0$ za $k \to \infty$. Zbog toga za proizvoljni $0 < \varepsilon < \frac{1}{2}$ postoji prirodan broj $N_0(\varepsilon) \in \mathbb{N}$ takav da za svaki $k \geq N_0(\varepsilon)$ vrijedi

$\left[\alpha_k \geq 1 \Rightarrow -\frac{\alpha_k}{2} \leq -\frac{1}{2} \Rightarrow 1 - \frac{\alpha_k}{2} \leq \frac{1}{2}\right]$

$$1 - \tfrac{1}{2} \geq \quad 1 - \tfrac{\alpha_k}{2} - \tfrac{\alpha_k}{2}\tfrac{\|f''(\xi)-f_k''\|}{m} \geq \varepsilon,$$

uz uvjet da je $\alpha_k \geq 1$. Budući da je α_k u uvjetu (4.30) biran tako da bude $\alpha_k \leq 1$, to znači da će u svim iteracijama čiji je redni broj veći ili jednak od $N_0(\varepsilon)$, biti $\alpha_k = 1$.

Dakle, za $k \geq N_0(\varepsilon)$ vrijedi $\quad [x_{k+1} = x_k - [f_k'']^{-1}\nabla f_k]$

$$(x_{k+1} - x^\star, x_{k+1} - x^\star) = \left(x_k - x^\star - [f_k'']^{-1}\nabla f_k, x_{k+1} - x^\star\right).$$

Ako na desnoj strani ove jednakosti iskoristimo Lagrangeovu formulu, uz $[\nabla f(x^\star) = 0]$ dobit ćemo

$$\begin{aligned}
([f_k'']^{-1}\nabla f_k, x_{k+1} - x^\star) &= ([f_k'']^{-1}(\nabla f_k - \nabla f(x^\star)), x_{k+1} - x^\star) \\
&= ([f_k'']^{-1}f''(\xi)(x_k - x^\star), x_{k+1} - x^\star),
\end{aligned}$$

gdje je $\xi = x^\star + \vartheta(x_k - x^\star), \quad \vartheta \in (0,1)$, dobivamo

$$\begin{aligned} \|x_{k+1} - x^\star\|^2 &= \left(\left(I - [f_k'']^{-1} f''(\xi)\right)(x_k - x^\star), x_{k+1} - x^\star\right) \\ &= \left([f_k'']^{-1}\left(f_k'' - f''(\xi)\right)(x_k - x^\star), x_{k+1} - x^\star\right) \\ \text{(4.32)} \quad &\leq \tfrac{1}{m}\|f_k'' - f''(\xi)\| \cdot \|x_k - x^\star\| \cdot \|x_{k+1} - x^\star\|, \end{aligned}$$

a odavde

$$\|x_{k+1} - x^\star\| \leq \lambda_k \|x_k - x^\star\|, \qquad \lambda_k = \tfrac{1}{m}\|f_k'' - f''(\xi)\|. \tag{4.42}$$

Budući da $\|f_k'' - f''(\xi)\| \to 0$ za $k \to \infty$, teorem je dokazan. □

Korolar 1. *Neka funkcija $f \in C^2(\mathbb{R}^n)$ zadovoljava uvjet (4.31) i Lipschitzov uvjet*

$$\|f''(x) - f''(y)\| \leq L\|x - y\|, \quad L > 0, \quad \forall x, y \in \mathbb{R}^n.$$

Tada niz (4.28)

$$x_{k+1} = x_k - \alpha_k [f''(x_k)]^{-1} \nabla f(x_k), \qquad k = 0, 1, \ldots,$$

gdje se parametri α_k biraju iz uvjeta (4.30), konvergira prema točki minimuma $x^\star$ neovisno o izboru početne aproksimacije x_0 kvadratičnom brzinom konvergencije

$$\|x_{k+1} - x^\star\| \leq \tfrac{L}{m}\|x_k - x^\star\|^2.$$

Dokaz. Koristeći Lipschitzov uvjet broj λ_k iz nejednakost (4.42) možemo pisati $[\xi = x^\star + \vartheta(x_k - x^\star), \quad 0 < \vartheta < 1]$

$$\begin{aligned} \lambda_k = \tfrac{1}{m}\|f''(x_k) - f''(\xi)\| &\leq \tfrac{L}{m}\|x_k - x^\star - \vartheta(x_k - x^\star)\| \\ &= \tfrac{L}{m}\|x_k(1 - \vartheta) - x^\star(1 - \vartheta)\| \\ &< \tfrac{L}{m}\|x_k - x^\star\|, \end{aligned}$$

što daje traženu ocjenu. □

Ako duljinu koraka α_k biramo iz uvjeta (4.2)

$$\alpha_k := \arg\min_{\alpha > 0} f(x_k + \alpha p_k),$$

tada za Newtonovu metodu (4.28)

$$x_{k+1} = x_k - \alpha_k [f''(x_k)]^{-1} \nabla f(x_k), \qquad k = 0, 1, \ldots,$$

vrijedi

Korolar 2. *Neka funkcija $f \in C^2(\mathbb{R}^n)$ zadovoljava uvjet (4.31) i neka se duljina koraka α_k bira iz uvjeta (4.2).*

Tada Newtonov iterativni proces (4.28) konvergira prema rješenju $x^\star$ superlinearnom brzinom bez obzira na izbor početne aproksimacije x_0. Ako pored toga funkcija f'' zadovoljava i Lipschitzov uvjet, onda je brzina kvadratična.

Dokaz. Neka je

$$\overline{x}_{k+1} = x_k - [f''(x_k)]^{-1}\nabla f(x_k)$$
$$x_{k+1} = x_k - \alpha_k[f''(x_k)]^{-1}\nabla f(x_k),$$

gdje je α_k izabran iz uvjeta (4.2). Budući da je $f(x_{k+1}) \le f(\overline{x}_{k+1})$, koristeći nejednakost (4.39), dobivamo $\left[\frac{m}{2}\|x - x^\star\|^2 \le f(x) - f(x^\star) \le \frac{M}{2}\|x - x^\star\|^2\right]$

$$\frac{m}{2}\|x_{k+1} - x^\star\|^2 \le f(x_{k+1}) - f(x^\star) \le f(\overline{x}_{k+1}) - f(x^\star) \le \frac{M}{2}\|\overline{x}_{k+1} - x^\star\|^2. \quad (4.43)$$

Za $\overline{x}_{k+1} = x_k - [f''(x_k)]^{-1}\nabla f(x_k)$ možemo primijeniti nejednakost (4.42)

$$\|\overline{x}_{k+1} - x^\star\| \le \lambda_k\|x_k - x^\star\|,$$

$$\lambda_k = \frac{1}{m}\|f_k'' - f''(\xi)\| = \frac{1}{m}\|f_k'' - f''(x^\star + \vartheta(x_k - x^\star))\|,$$

gdje $\lambda_k \to 0$ za $k \to \infty$. Ako ovu nejednakost iskoristimo u (4.43), dobivamo

$$\begin{aligned}\|x_{k+1} - x^\star\| &\le \sqrt{\frac{M}{m}}\|\overline{x}_{k+1} - x^\star\| \le \lambda_k\sqrt{\frac{M}{m}}\|x_k - x^\star\| \\ &\le \gamma_k\|x_k - x^\star\|,\end{aligned}$$

gdje $\gamma_k = \lambda_k\sqrt{\frac{M}{m}} \to 0$ za $k \to \infty$.

Ako još vrijedi i Lipschitzov uvjet, onda je $\lambda_k \le \frac{L}{m}\|x_k - x^\star\|$, pa vrijedi

$$\|x_{k+1} - x^\star\| \le C\|x_k - x^\star\|^2, \qquad C = \frac{L}{m}\sqrt{\frac{M}{m}}.$$

□

Zadatak 24. Newtonovom metodom odredite prve dvije aproksimacije minimuma funkcije $f(x,y) = 2x^3 + xy^2 + 5x^2 + y^2$ iz *primjera 7.* uz početnu aproksimaciju $x_0 = [-1, -1]^T$. Jeste li na taj način dobili aproksimaciju točke minimuma?

Rješenje: $x_2 = [-1, -2.05]^T$. To nije aproksimacija točke minimuma, već aproksimacija točke u kojoj ova funkcija ima sedlastu točku.

Zadatak 25. Prethodni zadatak riješite uz $x_0 = [1, 1]^T$. Dobivate li na taj način aproksimaciju točke minimuma?

Rješenje: $x_2 = [0.04237, 0.06759]^T$. To je aproksimacija točke minimuma.

Zadatak 26. Newtonovim metodom s regulacijom koraka (prema danom algoritmu za izračunavanje duljine koraka) odredite prve dvije aproksimacije minimuma funkcije $f(x,y) = x^4 + 3(x^2 + 2y^2) + 4xy + 5x + 6y + 7$, uz početnu aproksimaciju $x_0 = [0.5, -1]^T$ i točnost $\epsilon = 0.5$.

Rješenje: $x_2 = [-0.4169, -0.361]^T$.

Quasi–Newtonove metode

Budući da problem minimizacije diferencijabilne funkcije $f : \mathbb{R}^n \to \mathbb{R}$ možemo gledati kao problem rješavanja jednadžbe

$$F(x) = 0, \quad F : \mathbb{R}^n \to \mathbb{R}^n.$$

gdje je $F(x) = f'(x)$, gdje je F Jacobijeva matrica funkcije f. Ako je $f(x) = (f_1(x), \ldots, f_n(x))^T, \quad x = (x_1, \ldots, x_n)^T$, tada se Jacobijeva matrica sastoji od parcijalnih derivacija funkcija f_i, preciznije:

$$f' = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}.$$

Najprije ćemo razmotriti problem rješavanja sustava nelinearnih jednadžbi. Općenito se sustav nelinearnih jednadžbi kao i problem minimizacije funkcije više varijabli rješava iterativnim metodama. Smatra se da je klasu novih "kvazi-Newtonovih" metoda za rješavanje problema minimizacije funkcije više varijabli bez ograničenja uveo Davidon (1959.), a za rješavanje sustava nelinearnih jednadžbi Broyden (1965.). U literaturi se za klasu ovih metoda koriste razni nazivi: *quasi-Newton, variable metric, variance, secant, update, modification methods*.

Sustavi nelinearnih jednadžbi

Za funkciju $F : \mathbb{R}^n \to \mathbb{R}^n$ promatramo problem rješavanja jednadžbe $F(x) = 0$, odnosno sustava nelinearnih jednadžbi

$$f_i(x_1, \ldots, x_n) = 0, \qquad i = 1, \ldots, n,$$

gdje su funkcije f_i komponente vektorske funkcije $F(x) = (f_1(x), \ldots, f_n(x))^T, \quad x = (x_1, \ldots, x_n)^T$. Pretpostavimo da funkcija F ima sljedeća svojstva:

$$\begin{array}{ll} (i) & F \text{ je neprekidno diferencijabilna na otvorenom konveksnom} \\ & \text{skupu } \mathcal{D}, \text{ tj. } F \in C^1(\mathcal{D}), \\ (ii) & \text{postoji } x^\star \in \mathcal{D} \text{ tako da } F(x^\star) = 0 \text{ i da je pri tome} \\ & F'(x^\star) \text{ nesingularna matrica;} \end{array} \tag{4.44}$$

Dodatno, uz uvjete (4.44) možemo zahtijevati i ispunjenje jačeg *Lipschitzovog uvjeta*

$$(iii) \quad \|F'(x) - F'(x^\star)\| \leq L\|x - x^\star\|, \qquad x \in \mathcal{D}. \tag{4.45}$$

Primijetite da ako je $\mathcal{D}$ dovoljno malen, onda je ovaj uvjet ispunjen za dvostruko neprekidno diferencijabilnu funkciju u $x^\star$.

Uz pretpostavku da imamo k-tu aproksimaciju x_k rješenja $x^\star$, sljedeću $(k+1)$-u aproksimaciju kod obične Newtonove metode dobit ćemo tako da funkciju

F u okolini točke x_k pomoću Taylorove formule aproksimiramo linearnim aproksimandom

$$L(x) = F(x_k) + F'(x_k)(x - x_k)$$

i riješimo linearni sustav $L(x) = 0$. Tako dobivamo Newtonov iterativni postupak za rješavanje jednadžbe $F(x) = 0$

$$x_{k+1} = x_k - F'(x_k)^{-1} F(x_k).$$

Za praktičnu implementaciju obično koristimo sljedeći algoritam.

Korak 1. Za poznati x_k izračunati $F(x_k)$; ako je x_k prihvatljiv, stop. U protivnom izračunati $F'(x_k)$ i prijeći na Korak 2.

Korak 2. Riješiti sustav $F'(x_k)p = -F(x_k)$ za p, staviti $p_k := p$ i izračunati $x_{k+1} = x_k + p_k$.

Newtonova metoda poboljšava se uvođenjem parametra duljine koraka $0 < \alpha_k \leq 1$

$$x_{k+1} = x_k - \alpha_k F'(x_k)^{-1} F(x_k).$$

Karakteristike Newtonove metode dane su sljedećim teoremom.

Teorem 7. *Neka funkcija $F : \mathbb{R}^n \to \mathbb{R}^n$ zadovoljava pretpostavke (4.44). Nadalje pretpostavimo da je F' Lipschitz neprekidna, tj. neka vrijedi (4.45) pri čemu je $\mathcal{D}$ kugla oko $x^\star$ polumjera r. Također pretpostavimo da je $\|F'(x^\star)^{-1}\| \leq \beta$. Ako početnu iteraciju x_0 izaberemo tako da $\|x_0 - x^\star\| \leq \min\{r, 1/(2\beta L)\}$, tada je niz (x_n) definiran Newtonovim iterativnim postupkom dobro definiran i konvergira prema $x^\star$ kvadratičnom brzinom, tj.*

$$\|x_{k+1} - x^\star\| \leq \beta L \|x_k - x^\star\|^2, \qquad k = 0, 1, \ldots . \tag{4.46}$$

Dokaz. Prvo ćemo pokazati da pretpostavke da je F' Lipschitz neprekidna, da je $F'(x^\star)$ regularna i da je $\|F'(x^\star)^{-1}\| \leq \beta$ povlače da je $F'(x_0)$ regularno za sve x_0 za koje vrijedi da je $\|x_0 - x^\star\| \leq \min\{r, 1/(2\beta L)\}$.

Zaista,

$$\|F'(x^\star)^{-1} \left(F'(x_0) - F'(x^\star)\right)\| \leq \beta L \|x_0 - x^\star\| \leq \frac{1}{2}, \tag{4.47}$$

odakle slijedi

$$\begin{aligned} \|F'(x_0)^{-1}\| &= \|F'(x_0)^{-1} F'(x^\star) F'(x^\star)^{-1}\| \leq \|F'(x_0)^{-1} F'(x^\star)\| \, \|F'(x^\star)^{-1}\| \\ &\leq \frac{1}{1 - \frac{1}{2}} \|F'(x^\star)^{-1}\| \leq 2\beta . \end{aligned} \tag{4.48}$$

Pretposljednja nejednakost u gornjem izrazu jednostavna je posljedica sljedećeg računa. Iz (4.47) slijedi

$$\|F'(x^\star)^{-1} F'(x_0) - I\| \leq \frac{1}{2},$$

što znači da je $I - \big(F'(x^\star)^{-1}\,F'(x_0) - I\big) = F'(x^\star)^{-1}\,F'(x_0)$ regularno i vrijedi

$$\|F'(x_0)^{-1}\,F'(x^\star)\| \le \frac{1}{1-\frac{1}{2}}\,.$$

Ovim smo pokazali da je $F'(x_0)$ regularno za sve x_0 za koji vrijedi da je $\|x_0 - x^\star\| \le \min\{r, 1/(2\beta L)\}$, odnosno da je za svako takvo x_0 dobro definirano $x_1 = x_0 - F'(x_0)^{-1}F(x_0)$. Za tako definiran x_1 vrijedi

$$\begin{aligned} x_1 - x^\star &= x_0 - x^\star - F'(x_0)^{-1}F(x_0) \\ &= F'(x_0)^{-1}\left(F(x^\star) - F(x_0) - F'(x_0)(x^\star - x_0)\right) \end{aligned} \tag{4.49}$$

Kako bismo omeđili normu gornjeg izraza iskoristit ćemo jednakost

$$\begin{aligned} &F(x^\star) - F(x_0) - F'(x_0)(x^\star - x_0) = \\ &\qquad = \int_0^1 \left(F'(x_0 + t(x^\star - x_0)) - F'(x_0)\right) d\left(x_0 + t(x^\star - x_0)\right), \end{aligned}$$

i činjenicu da F' ispunjava Lipschitzov uvjet, što nam zajedno daje

$$\|F(x^\star) - F(x_0) - F'(x_0)(x^\star - x_0)\| \le L\int_0^1 t\|x^\star - x_0\|^2 dt = \frac{L}{2}\|x^\star - x_0\|^2.$$

Koristeći gornju ocjenu iz (4.49) i (4.48) slijedi

$$\|x_1 - x^\star\| \le 2\beta\frac{L}{2}\|x^\star - x_0\|^2 = \beta L\|x_0 - x^\star\|^2\,. \tag{4.50}$$

Pokažimo sada da niz $x_{k+1} = x_k - \alpha_k F'(x_k)^{-1}F(x_k)$, $k = 0, 1, 2\ldots$ konvergira prema $x^\star$. Iz (4.50) lako se pokaže da vrijedi $\|x_{k+1} - x^\star\| \le \beta L\|x_k - x^\star\|^2$, odnosno da vrijedi (4.46). □

Primjedba 5. Treba napomenuti da se može dokazati i znatno općenitiji teorem koji opisuje karakteristike Newtonove metode od *teorema 7*. Međutim, dokaz takvog teorema znatno je složeniji i dulji stoga ćemo ga samo navesti (dokaz vidi u [20], str. 312.)

Teorem 8. *Neka funkcija $F : \mathbb{R}^n \to \mathbb{R}^n$ zadovoljava pretpostavke (4.44). Tada postoji otvoren skup $\mathcal{S}$ koji sadržava $x^\star$, takav da je za proizvoljni $x_0 \in \mathcal{S}$ niz (x_n) definiran Newtonovim iterativnim postupkom dobro definiran, ostaje u $\mathcal{S}$ i konvergira prema $x^\star$ superlinearnom brzinom, tj. postoji nul-niz (λ_n), takav da je*

$$\|x_{k+1} - x^\star\| \le \lambda_k\|x_k - x^\star\|, \qquad k = 0, 1, \ldots.$$

Ako funkcija F' dodatno zadovoljava Lipschitzov uvjet, onda je brzina konvergencije kvadratična, tj. postoji konstanta β takva da je

$$\|x_{k+1} - x^\star\| \le \beta\|x_k - x^\star\|^2, \qquad k = 0, 1, \ldots.$$

Primjedba 6. Može se pokazati (vidi [11]) da za niz (x_n), koji superlinearnom brzinom konvergira prema $x^\star$, vrijedi

$$\lim_{k\to\infty} \frac{\|x_{k+1}-x_k\|}{\|x_k-x^\star\|} = 1$$

uz uvjet da je $x_k \neq x^\star$ za sve $k = 0, 1, \ldots$.

Na osnovi *primjedbe* 6. uz kriterij $\|F(x_k)\| < \varepsilon_1$ za zaustavljanje procesa ima smisla koristiti dodatni kriterij $\|x_{k+1}-x_k\| < \varepsilon_2\|x_k\|$

Prethodnim teoremom iskazane su dvije osnovne prednosti Newtonove metode:

- postoji područje prihvaćanja $\mathcal{S}$ Newtonove iterativne metode,
- visoka brzina konvergencije u području prihvaćanja $\mathcal{S}$.

Nedostaci Newtonove metode su:

- potreba izbora dobre početne aproksimacije x_0,
- u svakom koraku treba računati matricu parcijalnih derivacija reda n.

U cilju prevladavanja naročito posljednjeg navedenog nedostatka, matricu $F'(x_k)$ (odnosno matricu $[F'(x_k)]^{-1}$) zamijenit ćemo nekom matricom B_k (odnosno matricom H_k), koja se lako računa, a *nalikuje* matrici $F'(x_k)$ (odnosno matricu $[F'(x_k)]^{-1}$). Tako promatramo iterativne procese

$$\begin{aligned} & x_{k+1} := x_k - \alpha_k B_k^{-1} F(x_k), \\ \text{odnosno} \quad & x_{k+1} := x_k - \alpha_k H_k F(x_k). \end{aligned} \tag{4.51}$$

Vrijedi sljedeći teorem (vidi [11]).

Teorem 9. *Neka funkcija $F : \mathbb{R}^n \to \mathbb{R}^n$ zadovoljava uvjete (4.44), a niz (B_k) neka je niz regularnih matrica. Pretpostavimo nadalje da za neki $x_0 \in D$ niz (x_n) definiran s (4.51) ostaje u D, $x_k \neq x^\star$ za sve $k = 0, 1, \ldots$ i konvergira prema $x^\star$. Tada niz (x_n) konvergira superlinearno prema $x^\star$ onda i samo onda ako*

$$\lim_{k\to\infty} \frac{\|(B_k - F'(x^\star))p_k\|}{\|p_k\|} = 0, \qquad p_k = x_{k+1} - x_k. \tag{4.52}$$

Uvjet (4.52) zahtijeva da niz matrica (B_n) konvergira prema $F'(x^\star)$ u smjeru vektora kretanja $p_k = x_{k+1} - x_k$.

Primjedba 7. Uvjet (4.52) ekvivalentan je uvjetu (vidi primjerice [15])

$$\lim_{k\to\infty} \frac{\|B_k p_k - y_k\|}{\|p_k\|} = 0, \qquad y_k := F(x_{k+1}) - F(x_k), \tag{4.53}$$

što znači da uvjet $B_k p_k \approx y_k$ također jamči superlinearnu konvergenciju.

Ne može se očekivati da $B_k p_k = y_k$, ali zato ćemo matricu B_{k+1} birati tako da bude ispunjen tzv. *kvazi-Newtonov uvjet*

$$B_{k+1} p_k = y_k \tag{4.54}$$

i da je za poznate p_k, y_k, B_k lako izračunati B_{k+1}. Iterativna metoda (4.51) pri čemu niz matrica (B_n) ispunjava kvazi-Newtonov uvjet naziva se *kvazi-Newtonova metoda*. Radi pojednostavljivanja uvedimo sljedeće oznake:

$$\begin{aligned} &B := B_k, \qquad \overline{B} := B_{k+1}, \qquad x := x_k, \qquad \overline{x} := x_{k+1} \\ &p := p_k = \overline{x} - x, \qquad y := y_k = F(\overline{x}) - F(x). \end{aligned}$$

Uz ove oznake kvazi-Newtonov uvjet glasi

$$\overline{B}\, p = y.$$

Prvi put je takav niz matrica uveo 1965. godine Broyden [7], u literaturi poznat pod nazivom *Broydenova korekcija ranga* 1

$$\overline{B} = B + \tfrac{(y-Bp)p^T}{p^T p}. \tag{4.55}$$

U dokazima teorema koji slijede, zbog jednostavnosti ćemo pretpostaviti da je $\alpha_k = 1$, za sve k. Vrijedi sljedeći teorem

Teorem 10. *Neka funkcija $F : \mathbb{R}^n \to \mathbb{R}^n$ zadovoljava uvjete (4.44) i (4.45).*

1. *Tada postoji $\delta > 0$ takvo da ako $\|B_k - F'(x_k)\| \le \delta$ i $\|x_0 - x^\star\| \le \delta$, niz definiran sa $x_{k+1} = x_k - B_k^{-1} F(x_k)$ dobro je definiran i konvergira ka $x^\star$.*
2. *Ako pri tome vrijedi da $\|B_k - F'(x_k)\| \to 0$, tada je konvergencija superlinearna.*
3. *Ako postoji konstanta C takva da $\|B_k - F'(x_k)\| \le C\|F(x_k)\|$, tada je konvergencija kvadratična.*

Dokaz. Uvedimo oznaku $\beta = \|F'(x^\star)^{-1}\| < \infty$. Izaberemo li δ dovoljno malo, možemo postići:

$$\|x - x^\star\| \le \delta, \qquad \|B - F'(x)\| \le \delta \Rightarrow \|B^{-1}\| \le 2\beta\,.$$

Gornja nejednakost slijedi korištenjem sličnog računa kao u dokazu *teorema* 7. Zaista, ako izaberemo δ takvo da je $\delta < 1/(2\beta)$, onda iz

$$\|F'(x^\star)^{-1}\left(B - F'(x^\star)\right)\| \le \beta\delta < \frac{1}{2}$$

vidimo da je $F'(x^\star)^{-1} B = I - F'(x^\star)^{-1}\left(B - F'(x^\star)\right)$ regularna, te vrijedi ocjena

$$\|B^{-1} F'(x^\star)\| < \frac{1}{1 - \frac{1}{2}}\,.$$

Sada sve zajedno daje

$$\|B^{-1}\| = \|B^{-1}F'(x^\star)F'(x^\star)^{-1}\| \le 2\beta\,.$$

Neka je L Lipschitzova konstanta za F'. Dodatnim smanjivanjem δ možemo postići da je ispunjena nejednakost

$$2\beta\left(\frac{L}{2}+1\right)\delta \le \frac{1}{2}\,.$$

Pretpostavimo nadalje da smo B_0 i x_0 izabrali u skladu s gore izabranim δ. Tada

$$\begin{aligned} x_1 - x^\star &= x_0 - x^\star - B_0^{-1}F(x_0) = B_0^{-1}\left(F(x^\star) - F(x_0) - B_0(x^\star - x_0)\right) \\ &= B_0^{-1}\int_0^1 \left(F'((1-t)x_0 + tx^\star) - B_0\right)(x^\star - x_0)dt\,. \end{aligned}$$

Koristeći Lipschitzovo svojstvo funkcije F', nejednakost trokuta i uvjet na B_0 vrijedi

$$\|F'((1-t)x_0 + tx^\star) - F'(x_0) + F'(x_0) - B_0\| \le Lt\|x^\star - x_0\| + \delta\,.$$

To znači da možemo pisati

$$\begin{aligned} \|x_1 - x^\star\| &\le 2\beta\left(\frac{L}{2}\|x^\star - x_0\| + \delta\right)\|x^\star - x_0\| \le 2\beta\left(\frac{L}{2}+1\right)\delta\|x^\star - x_0\| \\ &\le \frac{1}{2}\|x^\star - x_0\|\,. \end{aligned}$$

Vidimo da gornja nejednakost povlači $\|x_1 - x^\star\| \le \delta$, pa ponavljajući ovaj postupak lako se dokazuje *1.* Iz gornjeg razmatranja vidimo da vrijedi

$$\|x_{k+1} - x^\star\| \le 2\beta\left(\frac{L}{2}\|x^\star - x_k\| + \|B_k - F'(x_k)\|\right)\|x^\star - x_k\|\,,$$

što daje superlinearnu konvergenciju, odnosno pretpostavku *2.* Uz dodatnu pretpostavku iz *3.* da je $\|B_k - F'(x_k)\| \le C\|F(x_k)\|$, slijedi

$$\|B_k - F'(x_k)\| \le C\|F(x_k) - F(x^\star)\| \le \tilde{C}\|x_k - x^\star\|\,,$$

što povlači kvadratnu konvergenciju. □

Neki primjeri Kvazi-Newtonovih metoda:

- Zamijenimo $\partial F^i(x)/\partial x_j$ s podijeljenom razlikom $[F^i(x + he_j) - F^i(x)]/h$ za neko malo h ili u zamjenu koristimo neki drugi diferencijalni kvocijent. Iz prethodnog teorema slijedi da će konvergencija biti osigurana za dovoljno male h.

- Koristimo jednom izračunatu derivaciju (Jacobijevu matricu) F' u više iteracijskih koraka. Ili kao u nekim metodama, poput Broydenove koju ćemo prikazati u sljedečem poglavlju, koristimo određeni postupak pomoću kojega "adekvatno mijenjamo" prethodni Jacobian.
- Izaberemo $B_i = \theta_i F'(x_i) + (1-\theta_i)I$. Tada konvergencija metode ovisi o θ_i i to metoda konvergira ako je θ_i dovoljno blizu 1, te je brzina konvergencije superlinearna ako $\theta_i \to 1$.

Broydenova metoda

Broydenova je metoda (objavio ju je C.G. Broyden 1965.) važan primjer kvazi-Newtonove metode. To je jedna od mogućih generalizacija metode sekante na n-dimenzionalne prostore. U slučaju jedne nelinearne jednadžbe, metoda sekante zamjenjuje $f'(x_i)$ u Newtonovoj metodi s aproksimacijom $[f(x_i) - f(x_{i-1})]/(x_i - x_{i-1})$. Na taj način dolazi se do sljedećeg iterativnog postupka: za dobivanje iteracija

$$x_{i+1} = x_i - \frac{x_i - x_{i-1}}{f(x_i) - f(x_{i-1})} f(x_i)\,.$$

Naravno, ova se ideja ne može izravno generalizirati na $\mathbb{R}^n$, budući da u $\mathbb{R}^n$ ne možemo dijeliti s vektorom $f(x_i) - f(x_{i-1})$. Umjesto toga, možemo uzeti u obzir jednadžbu

$$B_i(x_i - x_{i-1}) = F(x_i) - F(x_{i-1})\,.$$

Međutim, ta jednadžba ne određuje matricu B_i nego samo njezino djelovanje na $x_i - x_{i-1}$. Stoga, kako bismo u potpunosti odredili B_i, Broydenova metoda određuje djelovanje i na vektore ortogonalne na $x_i - x_{i-1}$ i to tako da je ona jednaka B_{i-1}. Broydenova metoda je tzv. metoda ažuriranja, tj. nova matrica B_i se određuje (ažurira) korištenjem prethodne B_{i-1}.

U svrhu definiranja Broydenove metode trebat će nam rezultat sljedećeg teorema:

Teorem 11. *Neka su zadani vektori $s \neq 0$ i $v \in \mathbb{R}^n$ i matrica $C \in R^{n\times n}$, tada postoji jedinstvena matrica $B \in R^{n\times n}$ takva da je*

$$Bs = v\,, \qquad Bz = Cz\,, \quad \textit{za sve } z \textit{ takve da } \; s^T z = 0.$$

Broyden je 1965. godine prvi put uveo takve matrice, koje se po njemu u literaturi nazivaju *Broydenovim korekcijama ranga* 1 i definiraju se

$$B = C + \frac{(v - Cs)s^T}{s^T s}. \tag{4.56}$$

Važno je primjetiti da matricu B dobivamo pomoću C dodavanjem matrice ranga 1. Nadalje, kao što ćemo pokazati u sljedećem teoremu, matrica B je na određeni od svih matrica koje s prebacuju u v način najbliža matrici C.

Teorem 12. *Za danu matricu $C \in \mathbb{R}^{n\times n}$ i vektore $s \neq 0$ i $v \in \mathbb{R}^n$ matrica B definirana sa (4.56) jedinstveno je rješenje problema minimizacije*

$$\min_{\tilde{B}} \left\{ \|\tilde{B} - C\|_F : \tilde{B}s = v \right\},$$

gdje je $\|\cdot\|_F$ Frobeniusova norma.

Dokaz. Najprije primijetimo da za proizvoljnu matricu $\tilde{B}$, za koju vrijedi kvazi-Newtonov uvjet $\tilde{B}s = v$, vrijedi

$$\begin{aligned} \|B - C\|_F &= \left\| \frac{(v - Cs)s^T}{s^T s} \right\|_F = \left\| \frac{(\tilde{B} - C)ss^T}{s^T s} \right\|_F \\ &\leq \|\tilde{B} - C\|_F \left\| \frac{ss^T}{s^T s} \right\|_F \\ &= \|\tilde{B} - C\|_F \end{aligned}$$

U posljednjoj jednakosti iskoristili smo

$$\left\| ss^T \right\|_F = \sum_{ij} (s_i s_j)^2 = \left(\sum_i s_i^2 \right) \left(\sum_j s_j^2 \right) = (s^T s)^2.$$

Kako je skup svih $\tilde{B} \in \mathbb{R}^{n\times n}$ takvih da je $\tilde{B}s = v$ konveksan, a funkcija $\Phi : \mathbb{R}^{n\times n} \mapsto \mathbb{R}$, $\Phi(A) = \|B - A\|_F$ strogo konveksna, matrica B na kojoj se postiže minimum je jedinstvena. □

Sada smo spremni navesti Broydenovu metodu.

Algoritam 3 . Broydenova metoda

Izabrati $x_0 \in \mathbb{R}^n$, $B_0 \in \mathbb{R}^{n\times n}$
for i:= 0, 1, ... **do**
  $x_{i+1} = x_i - B_i^{-1} F(x_i)$;
  $s_i = x_{i+1} - x_i$;
  $v_i = F(x_{i+1}) - F(x_i)$;
  $B_{i+1} = B_i + \frac{1}{s_i^T s_i}(v_i - B_i s_i)s_i^T$;
end for

Primjedba 8. Uobičajeno je za B_0 uzeti $F'(x_0)$. U jednodimenzionalnom slučaju Broydenova metoda svodi se na metodu sekante.

Ključ učinkovitosti Broydenove metode je u tome što se matrica B_{i+1} od matrice B_i razlikuje samo za matricu ranga 1. Ali, kao što će biti prikazano u sljedećem teoremu, jednom kada se efikasno izračuna djelovanje inverzne matrice na vektor a (npr., supstitucije unaprijed i unazad, za matricu koja je rastavljena na produkt donje i gornje trokutaste), tada se efikasno može izračunati i inverz te matrice promijenjene matricom ranga 1.

Dokažimo prvo sljedeću lemu.

Lema 7. *Za zadane $v, w \in \mathbb{R}^n$ vrijedi*

$$\det(I + vw^T) = 1 + (v, w).$$

Dokaz. Označimo $P = I + vw^T$. Ako je $v = 0$, dokaz je trivijalan. Zato pretpostavimo da je $v \neq 0$.

Neka je λ svojstvena vrijednost, a x odgovarajući svojstveni vektor od P. Tada mora biti

$$Px = \lambda x \Rightarrow (I + vw^T)x = \lambda x \Rightarrow x + v(w^T x) = \lambda x.$$

Ova jednakost može biti ispunjena tako da je ili $w \perp x$ ili $v \parallel x$.

Dakle, svaki svojstveni vektor od P je ili okomit na w ili paralelan s v.

1. Ako je $w \perp x$, onda je $\lambda = 1$;

2. Ako je $v \parallel x$, tj. $x = \alpha v$, tada je

$$\alpha v + vw^T \alpha v = \lambda \alpha v \Rightarrow v(1 + w^T v) = \lambda v \Rightarrow \lambda = 1 + (v, w).$$

Budući da može postojati samo jedan svojstveni vektor kolinearan s v, to znači da matrica P ima jednu svojstvenu vrijednost $1 + (v, w)$, dok su preostalih $(n - 1)$ jednake 1. Time je lema dokazana. □

Teorem 13 (Sherman-Morrison-Woodbury). *Neka je $u, v \in \mathbb{R}^n$, a $B \in \mathbb{R}^{n \times n}$ regularna matrica. Tada je $\tilde{B} = B + uv^T$ regularna onda i samo onda ako je $\sigma = 1 + v^T B^{-1} u \neq 0$. Ako je $\sigma \neq 0$, vrijedi*

$$\tilde{B}^{-1} \equiv \left(B + uv^T\right)^{-1} = B^{-1} - \frac{1}{\sigma} B^{-1} uv^T B^{-1} \tag{4.57}$$

Dokaz. Budući da je B regularna matrica, imamo

$$\det(B + uv^T) = \det B \det(I + B^{-1} uv^T)$$

pa je $\det(B + uv^T) \neq 0$ ako i samo ako je $\det(I + B^{-1}uv^T) \neq 0$. Koristeći prethodnu lemu vrijedi $\det(I + B^{-1}uv^T) \neq 0$ ako i samo ako je $1 + v^T B^{-1} u \neq 0$, što nam daje traženu karakterizaciju regularnosti od $\tilde{B}$.

Nadalje, neka je za bilo koji vektor $y \in \mathbb{R}^n$ dan $x = \tilde{B}^{-1} y$, odnosno $\tilde{B}x = y$, to znači da

$$Bx + (v^T x)u = y\,. \tag{4.58}$$

Množenjem gornje jednakosti s lijeva s $v^T B^{-1}$ dobivamo $(v^T x)(1 + v^T B^{-1} u) = v^T B^{-1} y$. Kako je $1 + v^T B^{-1} u \neq 0$, dobivamo

$$v^T x = \frac{v^T B^{-1} y}{1 + v^T B^{-1} u}\,.$$

Kombinacijom gornjeg izraza sa (4.58) slijedi

$$Bx = y - \frac{v^T B^{-1} y}{1 + v^T B^{-1} u} u\,.$$

Sada množenjem gornje jednakosti s lijeva s B^{-1} i korištenjem $x = \tilde{B}^{-1}y$, dobivamo

$$\tilde{B}^{-1}y \equiv x = B^{-1}y - \frac{v^T B^{-1} y}{1 + v^T B^{-1} u} B^{-1} u = \left(B^{-1} - \frac{B^{-1} u v^T B^{-1}}{1 + v^T B^{-1} u} \right) y\,,$$

iz čega slijedi Sherman-Morrison-Woodburyeva fromula (4.57). □

Stoga kako bismo odredili umnožak $\tilde{B}^{-1}$ s nekim vektorom treba nam samo umnožak B^{-1} s tim vektorom te skalarni produkt vektora v s tim vektorom.

Dobar način implementacije Sherman-Morrison-Woodburyeve formule u Broydenovoj metodi može se dobiti spremanjem $H_i \doteq B_i^{-1}$, što je efikasnije nego spremanje B_i. Algoritam tada postaje:

Algoritam 4 . Modificirana Broydenova metoda

Izabrati $x_0 \in \mathbb{R}^n$, $H_0 \in \mathbb{R}^{n \times n}$
for i:= 0, 1, ... **do**
  $x_{i+1} = x_i - H_i F(x_i)$;
  $s_i = x_{i+1} - x_i$;
  $v_i = F(x_{i+1}) - F(x_i)$;
  $H_{i+1} = H_i + \frac{1}{s_i^T H_i v_i}(s_i - H_i v_i) s_i^T H_i$;
end for

Primijetite da u slučaju ako su H_0 i B_0^{-1} jednaki, tada gornji algoritam predstavlja osnovni Broydenov algoritam.

Konvergencija Broydenove metode

Označimo sa $x^\star$ rješenje nelinearnog sustava opisanog sa $F(x) = 0$. Neka x_i i B_i označavaju vektor i matricu koje su određene i-tim korakom Broydenove metode. Označimo sa

$$e_i = x_i - x^\star\,, \qquad M_i = B_i - F'(x^\star)\,. \tag{4.59}$$

Grubo govoreći, za konvergenciju Broydenove metode ključne su sljedeće dvije činjenice:

(1) e_{i+1} će biti malo u odnosu na e_i, u slučaju da M_i nije preveliko,

(2) M_{i+1} neće biti puno veći od M_i, ako su e_i mali.

Preciznije objašnjenje temelji se na sljedećim jednakostima, koje slijede izravno iz oznaka uvedenih sa (4.59) i definicije B_i,

$$e_{i+1} = -B_i^{-1}\left[F(x_i) - F(x^\star) - F'(x^\star)(x_i - x^\star)\right] + B_i^{-1} M_i e_i\,, \tag{4.60}$$

i

$$M_{i+1} = M_i \left(I - \frac{1}{s_i^T s_i} s_i s_i^T \right) + \frac{1}{s_i^T s_i}\left(v_i - F'(x^\star) s_i s_i^T \right)\,. \tag{4.61}$$

Sljedeći teorem govori o lokalnoj converegenciji Broydenove metode, te o njezinoj brzini konvergencije koja je u najgorem slučaju linearna.

Teorem 14. *Neka je funkcija F diferencijabilna na kugli $\mathcal{D}$ oko rješenja $x^\star \in \mathbb{R}^n$, ($F(x^\star) = 0$), te neka je $\|F'(x) - F'(x^\star)\| \leq L\|x - x^\star\|$, za sve $x \in \mathcal{D}$. Pretpostavimo da je $F'(x^\star)$ regularna, pri čemu vrijedi $\|F'(x^\star)\| \leq \beta$. Neka su $x_o \in \mathcal{D}$, i $B_0 \in \mathbb{R}^{n\times n}$ izabrani tako da vrijedi*

$$\|M_0\| + 2L\|e_0\| \leq \frac{1}{8\beta}\,.$$

Tada su nizovi x_i i B_i određeni Broydenovim algoritmom 3. dobro definirani, a za niz grešaka vrijedi

$$\|e_{i+1}\| \leq \frac{1}{2}\|e_i\|\,, \qquad i = 0, 1, 2, \ldots.$$

Dokaz. Kako bismo mogli dokazati teorem trebat će nam sljedeće tvrdnje:

Tvrdnja **1.** Pokažimo da ako su nizovi x_i i B_i određeni Broydenovim algoritmom 3. dobro definirani i ako je $\|M_i\| \leq 1/(2\beta)$, tada su B_i regularne matrice (odnosno x_{i+1} je dobro definiran), te

$$\|B_i^{-1}\| \leq 2\beta\,, \qquad \|e_{i+1}\| \leq (L\beta\|e_i\| + 2\beta\|M_i\|)\,\|e_i\|\,.$$

Zaista, $F'(x^\star)^{-1}B_i = I + F'(x^\star)^{-1}M_i$, a kako je $\|M_i\| \leq 1/(2\beta)$, $\|F'(x^\star)^{-1}M_i\| \leq 1/(2\beta)$, vidimo da je B_i invertibilna, pri čemu vrijedi $\|B_i^{-1}\| \leq 2\beta$. Stoga slijedi da je x_{i+1} dobro definiran. Ocjena za $\|e_{i+1}\|$ dobiva se pomoću ocjene $\|B_i^{-1}\| \leq 2\beta$ i (4.60).

Primijetite da iz pretpostavki teorema i gore navedenog slijedi da je x_1 dobro definiran i da vrijedi $\|e_1\| \leq 1/2\|e_0\|$.

Tvrdnja **2.** Ako su $B_0, \ldots, B_i$ dobro definirane i invertibilne, onda vrijedi

$$\|M_{i+1}\| \leq \|M_i\| + \max\{\|e_i\|, \|e_{i+1}\|\}\,.$$

Gornju ćemo ocjenu dokazati koristeći (4.61). Prvi dio na desnoj strani u (4.61) je umnožak M_i s ortogonalnim projektorom (na ortogonalni komplement od s_i), tako da je u bilo kojoj unitarno invarijantoj normi taj produkt omeđen s $\|M_i\|$. Uočimo da se dio drugog člana može izraziti kao

$$v_i - F'(x^\star)s_i = \int_0^1 \left[F'((1-t)x_i + tx_{i+1}) - F'(x^\star)\right] s_i dt\,.$$

Kako je $\|F'((1-t)x_i + tx_{i+1}) - F'(x^\star)\| \leq L\max\{\|e_i\|, \|e_{(i+1)}\|\}$, slijedi da je

$$\|v_i - F'(x^\star)s_i\| \leq L\max\{\|e_i\|, \|e_{i+1}\|\}\|s_i\|^2 \frac{1}{\|s_i\|^2}\,,$$

što znači da je drugi dio u izrazu (4.61) omeđen s

$$L\max\{\|e_i\|, \|e_{i+1}\|\}\,,$$

čime je dokazana i tvrdnja **2.**

Ostaje nam dokazati teorem. Koristeći indukciju po i pokazat ćemo da je niz $x_0, \ldots, x_{i+1}$ dobro definiran i da vrijedi:

$$\|e_i\| \le \frac{1}{8L\beta}, \quad \|M_i\| \le \frac{1}{\delta\beta}, \quad \|e_{i+1}\| \le \frac{1}{2}\|e_i\|.$$

Gornje ocjene očito vrijede za $i = 0$. Pretpostavimo da su istinite i za sve indekse manje ili jednake od i. Zamjenom i sa $i+1$ direktno iz pretpostavke indukcije slijedi $\|e_{i+1}\| \le \frac{1}{8L\beta}$.

Uzastopnom primjenom tvrdnje **2.**, uzimajući u obzir $\|e_{i+1}\| \le \|e_i\| \le \ldots$ imamo

$$\begin{aligned}\|M_{i+1}\| &\le \|M_0\| + L\left(\|e_0\| + \|e_1\| + \ldots + \|e_{i+1}\|\right)\\ &\le \|M_0\| + L\|e_0\|(1 + \frac{1}{2} + \ldots) = \|M0\| + 2L\|e_0\| \le \frac{1}{8\beta},\end{aligned}$$

što odgovara drugoj nejednakosti. Konačno primjenom tvrdnje **1.** slijedi treća nejednakost, čime je teorem dokazan. □

Minimizacija glatke funkcije

Promatrajmo sada problem minimizacije bez ograničenja za dovoljno glatku funkciju $f : \mathbb{R}^n \to \mathbb{R}$. Pretpostavimo da je $x^\star$ njen lokalni minimum. Prema *teoremu 3.* tada je $f'(x^\star) = 0$, a $f''(x^\star)$ pozitivno je definitna Hesseova matrica. Uz oznaku $F(x) := f'(x)$, analogno uvjetima (4.44) i (4.45) definirat ćemo uvjete na funkciju $f : \mathbb{R}^n \to \mathbb{R}$:

(i) f je dvostruko neprekidno diferencijabilna na otvorenom konveksnom skupu D, tj. $f \in C^2(D)$;

(ii) postoji $x^\star \in D$ tako da $f'(x^\star) = 0$, a $f''(x^\star)$ je pozitivno definitna matrica;

(iii) f'' zadovoljava Lipschitzov uvjet $\|f''(x) - f''(x^\star)\| \le L\|x - x^\star\|, \quad L > 0, \quad \forall x \in D$.

Uz oznake $g(x) := f'(x)$ i $g_k = g(x_k)$ iterativni postupak (4.51) postaje

$$\begin{aligned} & x_{k+1} := x_k - \alpha_k B_k^{-1} g_k,\\ \text{odnosno}\quad & x_{k+1} := x_k - \alpha_k H_k g_k,\end{aligned}$$

gdje je sada B_k aproksimacija Hessijana $f''(x_k)$, a $H_k = B_k^{-1}$, pri čemu duljinu koraka α_k možemo odrediti nekom metodom izbora duljine koraka iz 4.1. Budući da matrice B_k trebaju aproksimirati Hessijan $f''(x_k)$, za njih se mora pretpostaviti simetričnost. Slično kao ranije, uvedimo oznake

$$B := B_k, \qquad \overline{B} := B_{k+1}$$

$$p := p_k = x_{k+1} - x_k \qquad y := y_k = g_{k+1} - g_k \quad (\text{ zbog } g(x) = F(x))$$

Jedna mogućnost za definiranje korekcije ranga 1 je *simetrična formula ranga* 1 koju je uveo Davidon (1959.) (vidi [8]):

$$\overline{B} = B + \frac{(y - Bp)(y - Bp)^T}{(y - Bp, p)}, \qquad (y - Bp, p) \neq 0 \tag{4.62}$$

Zadatak 27. Pokažite da ako postoje matrice: $H := B^{-1}$ i $\overline{H} := \overline{B^{-1}}$ i ako je B simetrična, tada vrijedi

$$\overline{H} = H + \frac{(p-Hy)(p-Hy)^T}{(p-Hy,y)} \tag{4.63}$$

Primjer 17. Potražimo lokalni minimum funkcije

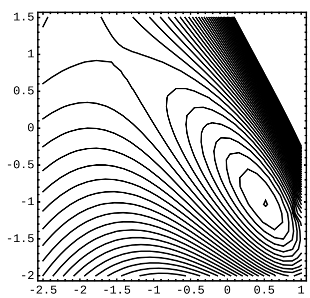
$$f(x_1, x_2) = e^{x_1}\left(4x_1^2 + 2x_2^2 + 4x_1x_2 + 2x_2 + 1\right)$$

Gradijent, Hessijan i `ContourPlot` dobit ćemo na sljedeći način:

```
In[1]:=
f[x_] := Exp[x[1]](4x[1]^2 + 2x[2]^2 + 4x[1]x[2] + 2x[2] + 1); n = 2;
g[x_] := Table[D[f[x], x[i]], {i, n}]
H[x_] := Table[D[g[x][[i]], x[j]], {i, n}, {j, n}]
  Print["f'=", Simplify[g[x] /. {x[1] -> x1, x[2] -> x2}] // MatrixForm,
  "   H=", Simplify[H[x] /. {x[1] -> x1, x[2] -> x2}] // MatrixForm]
  ContourPlot[f[x] /. {x[1] -> x1, x[2] -> x2}, {x1, -2.5, 1}, {x2, -2, 1.5},
  Contours -> 40, PlotPoints -> 30, ContourShading -> None];
```

$$f' = \begin{bmatrix} e^{x_1}(1 + 4x_1^2 + 6x_2 + 2x_2^2 + 4x_1(2 + x_2)) \\ e^{x_1}(2 + 4x_1 + 4x_2) \end{bmatrix},$$

$$H = \begin{bmatrix} e^{x_1}(9 + 4x_1^2 + 10x_2 + 2x_2^2 + 4x_1(4 + x_2)) & e^{x_1}(3 + 2x_1 + 2x_2) \\ e^{x_1}(3 + 2x_1 + 2x_2) & e^{x_1} \end{bmatrix}$$



Slika 4.5. `ContourPlot` funkcije $f(x_1, x_2) = e^{x_1}\left(4x_1^2 + 2x_2^2 + 4x_1x_2 + 2x_2 + 1\right)$

Zadavanjem ulaznih podataka

```
In[2]:= n = 2; x0 = {1., 1}; k = 0; eps = 0.005;
          w = Table[x[i] -> x0[[i]], {i, n}]; B0 = IdentityMatrix[n];
            B0 = H[x] /. w
              g0 = g[x] /. w
```

i korištenjem modula `Korak` (str. 40.) za određivanje duljine koraka

```
In[3]:= While[
    p0 = -LinearSolve[B0, g0];
    alpha = Korak[n, f, p0, x0];
    x1 = x0 + alpha p0;
    w1 = Table[x[i] -> x1[[i]], {i, n}]; g1 = g[x] /. w1;
    y0 = g1 - g0;
    B1 = B0+Outer[Times, (y0 - B0.p0),(y0 - B0.p0)]/((y0 - B0.p0).p0);
    f1 = f[x] /. w1;
    Abs[f1] > eps,
      B0 = B1; g0 = g1;
      x0 = x1; k = k + 1;
      Print["x", k, "=", x1, "  fx=", f[x] /. w1,
      "  grad=", g1, "  alpha=", alpha] ];
      Print["x", k + 1, "=", x1, "  fx=",f[x] /. w1,
      "  grad=", g1, "  alpha=", alpha]
```

nakon 18 iteracija dobivamo

$$x_{18} = [-11.43, 19.80]^T, \quad f(x_{18}) = 0.00479, \quad f'(x_{18}) = [0.0047, 0.00038]^T, \quad \alpha_{18} = 1.$$

Slično *teoremu 12.*, vrijedi

Teorem 15. *Neka je $M \in \mathbb{R}^{n\times n}$ simetrična regularna matrica, $y, p \in \mathbb{R}^n \quad (p \neq 0)$ i $c := M^{-2}p$. Ako je B simetrična matrica, tada je rješenje problema minimizacije*

$$\min_{\hat{B}} \left\{ \|M(\hat{B} - B)M\|_F : \hat{B} = \hat{B}^T, \quad \hat{B}p = y \right\} \tag{4.64}$$

zadano s

$$\overline{B} = B + \frac{(y - Bp)c^T + c(y - Bp)^T}{c^T p} - \frac{(y - Bp)^T p}{(c^T p)^2} cc^T. \tag{4.65}$$

Matricom M definirana je težinska Frobeniusova norma, $\overline{B}$ je također simetrična matrica, koja ispunjava kvazi-Newtonov uvjet $\overline{B}p = y$ i po težinskoj Frobeniusovoj normi najmanje se razlikuje od matrice B. Formulu (4.65) zovemo korekcija ranga 2 jer se $\overline{B}$ dobiva tako da matrici B dodamo jednu matricu ranga 2.

Zadatak 28. Dva različita dokaza ovog teorema mogu se vidjeti u [15] i [11]. Izvedite i usporedite oba dokaza.

Pokazat ćemo (vidi [11]) kako se izvodi kvazi-Newtonova korekcija ranga 2 zadana s (4.65) jer je to način na koji se mogu izvesti i druge poznate korekcije ranga 2. Polazimo od simetrične matrice $B \in \mathbb{R}^{n\times n}$, te slično Broydenovoj korekciji ranga 1 (4.56) kao mogućeg kandidata za $\overline{B}$ promatramo

$$C_1 = B + \tfrac{(y-Bp)c^T}{(c,p)}.$$

koji ispunjava kvazi-Newtonov uvjet $C_1 p = y$. Budući da C_1 općenito nije simetrična matrica, definiramo simetričnu matricu

$$C_2 = \tfrac{1}{2}\left(C_1 + C_1^T\right).$$

Budući da C_2 ne ispunjava kvazi-Newtonov uvjet, ponovimo postupak i tako rekurzivno definiramo niz (C_n)

$$\begin{aligned} C_{2k+1} &= C_{2k} + \tfrac{(y-C_{2k}p)c^T}{(c,p)}, \\ C_{2k+2} &= \tfrac{1}{2}\left(C_{2k+1} + C_{2k+1}^T\right), \end{aligned} \tag{4.66}$$

gdje je $C_0 = B$. Vrijedi

Lema 8. *Neka je $B \in \mathbb{R}^{n\times n}$ simetrična matrica i $c, p, y \in \mathbb{R}^n$ sa svojstvom $(c, p) \neq 0$. Tada niz (C_n) zadan rekurzivnom formulom (4.66) uz $C_0 = B$ konvergira prema $\overline{B}$ zadanom s (4.65).*

Dokaz. Dokažimo da niz (C_{2k}) konvergira. Uz oznaku $G_k = C_{2k}$ iz (4.66) dobivamo

$$\begin{aligned} G_{k+1} = C_{2(k+1)} &= \tfrac{1}{2}\left(C_{2k+1} + C_{2k+1}^T\right) = \tfrac{1}{2}\left(G_k + \tfrac{(y-G_kp)c^T}{(c,p)} + G_k^T + \tfrac{c(y-G_kp)^T}{(c,p)}\right) \\ &= G_k + \tfrac{w_kc^T + cw_k^T}{2(c,p)}, \end{aligned}$$

gdje je $w_k = y - G_k p$. Koristeći dobivenu jednakost dobivamo

$$\begin{aligned} w_{k+1} = y - G_{k+1}p &= y - G_kp - \tfrac{1}{2}\tfrac{w_kc^T + cw_k^T}{(c,p)}p = w_k - \tfrac{1}{2}\tfrac{w_kc^Tp}{(c,p)} - \tfrac{1}{2}\tfrac{cw_k^Tp}{(c,p)} \\ &= \tfrac{1}{2}w_k - \tfrac{1}{2}\tfrac{c(p^Tw_k)}{(c,p)}, \end{aligned}$$

što možemo pisati kao

$$w_{k+1} = Pw_k, \qquad P = \tfrac{1}{2}\left[I - \tfrac{cp^T}{(c,p)}w_k\right]$$

Prema *Lemil* 7. matrica P ima jednu svojstvenu vrijednost 0 $\left[\tfrac{1}{2}\left(1 - (\tfrac{c}{(c,p)}, p)\right) = 0\right]$ i svojstvenu vrijednost $\tfrac{1}{2}$ kratnosti $(n-1)$. Dakle, spektralni radijus matrice P je $\rho(P) = \tfrac{1}{2} < 1$. Prema *Neumannovoj lemi* (vidi primjerice [20], str. 45.) tada postoji matrica $(I-P)^{-1}$ i vrijedi

$$(I-P)^{-1} = \lim_{k\to\infty}\sum_{i=0}^{k} P^i.$$

Kako je s druge strane

$$w_k = Pw_{k-1} = \cdots = P^k w_0 = P^k(y - Bp),$$

onda imamo

$$\sum_{k=0}^{\infty} w_k = \sum_{k=0}^{\infty} P^k(y - Bp) = (I - P)^{-1}(y - Bp).$$

Također vrijedi

$$\begin{aligned} \sum_{k=0}^{n}(G_{k+1} - G_k) = G_{n+1} - G_0 \quad &\Longrightarrow \quad \sum_{k=0}^{\infty}(G_{k+1} - G_k) = \lim_{n\to\infty}(G_{n+1} - G_0) \\ &\Longrightarrow \quad \sum_{k=0}^{\infty}(G_{k+1} - G_k) = \lim_{k\to\infty} G_k - B \end{aligned}$$

Zato vrijedi

$$\begin{aligned} & \lim_{k\to\infty} G_k = \\ = \; & B + \sum_{k=0}^{\infty}(G_{k+1} - G_k) = B + \sum_{k=0}^{\infty} \tfrac{1}{2} \tfrac{w_k c^T + c w_k^T}{(c,p)} = B + \tfrac{1}{2(c,p)} \sum_{k=0}^{\infty} w_k c^T + \tfrac{1}{2(c,p)} \sum_{k=0}^{\infty} c w_k^T \\ = \; & B + \tfrac{1}{2(c,p)} \left[(I - P)^{-1}(y - Bp)c^T + c(y - Bp)^T(I - P^T)^{-1}\right] = \\ & \Big\{\text{prema } \textit{lemi 13.}: \; I - P = I - \tfrac{1}{2}\left[I - \tfrac{cp^T}{(c,p)}\right] = \tfrac{1}{2}\left[I + \tfrac{cp^T}{(c,p)}\right], \\ & \sigma = 1 + p^T \tfrac{c}{(c,p)} = 2, \; (I - P)^{-1} = 2\left[I - \tfrac{1}{2}\tfrac{cp^T}{(c,p)}\right], \; (I - P^T)^{-1} = 2\left[I - \tfrac{1}{2}\tfrac{pc^T}{(c,p)}\right]\Big\}, \\ = \; & B + \tfrac{1}{2(c,p)} \left\{2\left[I - \tfrac{1}{2}\tfrac{cp^T}{(c,p)}\right](y - Bp)c^T + c(y - Bp)^T 2\left[I - \tfrac{1}{2}\tfrac{pc^T}{(c,p)}\right]\right\} \\ = \; & B + \tfrac{1}{(c,p)} \left\{(y - Bp)c^T + c(y - Bp)^T - \tfrac{1}{2}\tfrac{cp^T}{(c,p)}(y - Bp)c^T - \tfrac{1}{2}c(y - Bp)^T \tfrac{pc^T}{(c,p)}\right\} \\ = \; & B + \tfrac{1}{(c,p)} \left\{(y - Bp)c^T + c(y - Bp)^T - \tfrac{c(y-Bp)^T pc^T}{(c,p)}\right\} \end{aligned}$$

iz čega neposredno slijedi (4.65). □

Specijalno za $c = p$ dobiva se *Powell–Broydenova simetrična korekcija* (Powell Symmetric Broyden - PSB)

$$\overline{B}_{PSB} = B + \frac{(y - Bp)p^T + p(y - Bp)^T}{(p,p)} - \frac{((y - Bp), p)pp^T}{(p,p)^2}. \qquad (4.67)$$

Korekcije koje čuvaju pozitivnu definitnost

Do sada smo promatrali korekcije ranga 1 ili 2 koje su čuvale simeričnost matrice B i zadovoljavale kvazi-Newtonov uvjet. Dodatni zahtjev koji se postavlja bit će čuvanje pozitivne definitnosti. Kvazi-Newtonov uvjet zajedno s uvjetom čuvanja pozitivne definitnosti matrice B pobliže određuje kut između vektora p i y:

$$(y, p) > 0. \qquad \left[(y,p) \overset{\overline{B}p=y}{=} (\overline{B}p, p)\right]$$

U cilju ispitivanja nasljednosti pozitivne definitnosti koristit ćemo sljedeću lemu (vidi primjerice [11]):

Lema 9. *Neka je $A \in \mathbb{R}^{n\times n}$ simetrična matrica sa svojstvenim vrijednostima*

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n,$$

i neka je $A_T = A + \sigma uu^T$ za neki $u \in \mathbb{R}^n$. Ako je $\sigma \ge 0$, onda A_T ima svojstvene vrijednosti λ_i^T takve da je

$$\lambda_1 \le \lambda_1^T \le \lambda_2 \le \cdots \le \lambda_n \le \lambda_n^T,$$

a ako je $\sigma \le 0$, onda A_T ima svojstvene vrijednosti λ_i^T takve da je

$$\lambda_1^T \le \lambda_1 \le \lambda_2^T \le \cdots \le \lambda_n^T \le \lambda_n.$$

Teorem 16. *Neka je $B \in \mathbb{R}^{n\times n}$ simetrična pozitivno definitna matrica i neka su $c, p, y \in \mathbb{R}^n$ takvi da je $(c, p) \ne 0$. Tada je $\overline{B}$ definiran s (4.65) pozitivno definitna matrica onda i samo onda ako je* $\det \overline{B} > 0$.

Dokaz. Ako je $\overline{B} > 0$, onda je iz opće teorije poznato da je $\det \overline{B} > 0$. Za dokaz obrata, matricu $\overline{B}$ napišimo najprije u obliku

$$\overline{B} = B + \frac{(y-Bp)c^T + c(y-Bp)^T}{c^Tp} - \frac{(y-Bp)^Tp}{(c^Tp)^2}cc^T$$
$$\overline{B} = B + vw^T + wv^T,$$

gdje je

$$v = \frac{y - Bp}{(c,p)} - \frac{1}{2}\frac{(y-Bp,p)}{(c,p)^2}c, \qquad w = c.$$

Zbog toga je

$$\Big[(v+w)(v+w)^T - (v-w)(v-w)^T = vv^T + ww^T$$
$$+2vw^T - \left(vv^T + ww^T - 2wv^T\right)\Big]$$

$$\overline{B} = B + \tfrac{1}{2}\left((v+w)(v+w)^T - (v-w)(v-w)^T\right)$$

Na taj smo način $\overline{B}$ napisali kao zbroj jedne pozitivno definitne matrice (B) i dvije simetrične matrice ranga 1. Prema *lemi* 9. matrica

$$C = B + \tfrac{1}{2}(v+w)(v+w)^T$$

pozitivno je definitna jer je $\sigma = \frac{1}{2} > 0$.

Nadalje, također prema *lemi* 9., s obzirom da je $C > 0$ i $\sigma = -\frac{1}{2} < 0$ matrica

$$\overline{B} = C - \tfrac{1}{2}(v-w)(v-w)^T$$

može imati najviše jednu svojstvenu vrijednost nepozitivnu. Kako je s druge strane

$$0 < \det \overline{B} = \lambda_1 \cdots \lambda_n,$$

i ta jedna svojstvena vrijednost mora biti pozitivna. Dakle, $\overline{B} > 0$. □

U cilju ispitivanja za koji će $c \in \mathbb{R}^n$ rekurzivna formula (4.65) čuvati pozitivnu definitnost sljedeća lema ispituje determinantu matrice $\overline{B}$.

Lema 10. *Neka je $u_i \in \mathbb{R}^n$, $i = 1, 2, 3, 4$. Tada vrijedi*

$$\det\left(I + u_1u_2^T + u_3u_4^T\right) = (1 + (u_1, u_2))\,(1 + (u_3, u_4)) - (u_1, u_4)(u_2, u_3).$$

Dokaz. Ako trenutno pretpostavimo da je $(u_1, u_2) \neq -1$, tj. da je sukladno Sherman-Morrisonovoj lemi $I + u_1u_2^T$ regularna matrica, onda možemo pisati

$$I + u_1u_2^T + u_3u_4^T = \left(I + u_1u_2^T\right)\left(I + (I + u_1u_2^T)^{-1}u_3u_4^T\right).$$

Zato je $\det\left(I + u_1u_2^T + u_3u_4^T\right) = D_1 \cdot D_2$, gdje je

$$D_1 = \det\left(I + u_1u_2^T\right) = 1 + (u_1, u_2),$$
$$D_2 = \det\left(I + (I + u_1u_2^T)^{-1}u_3u_4^T\right).$$

Kako je $\left((I + u_1u_2^T\right)^{-1} = I - \frac{1}{1+(u_1,u_2)}u_1u_2^T$, onda

$$\begin{aligned} D_2 &= \det\left[I + \left(I - \tfrac{u_1u_2^T}{1+(u_1,u_2)}\right)u_3u_4^T\right] = \det\left[I + u_3u_4^T - \tfrac{u_1u_2^Tu_3u_4^T}{1+(u_1,u_2)}\right] \\ &= \det\left[I + \left(u_3 - \tfrac{(u_2,u_3)}{1+(u_1,u_2)}u_1\right)u_4^T\right] = 1 + \left(u_3 - \tfrac{(u_2,u_3)}{1+(u_1,u_2)}u_1, u_4\right) \\ &= 1 + (u_3, u_4) - \tfrac{(u_2,u_3)}{1+(u_1,u_2)}(u_1, u_4), \end{aligned}$$

iz čega slijedi tražena jednakost. □

Primjenom *leme* 10. na (4.65) uz pretpostavku da je postoji B^{-1} i oznaku $H := B^{-1}$, dobivamo

$$\det \overline{B} = \det B\left[\left((c, Hy)^2 - (c, Hc)(y, Hy) + (c, Hc)(y, p)\right)\tfrac{1}{(c,p)^2}\right]. \tag{4.68}$$

U svrhu dokaza ove formule najprije ćemo (4.65)

$$\overline{B} = B + \tfrac{(y-Bp)c^T + c(y-Bp)^T}{c^Tp} - \tfrac{(y-Bp)^Tp}{(c^Tp)^2}cc^T$$

napisati u obliku [$(y-Bp)^T p$ je broj pa ga možemo "staviti unutar" cc^T]

$$\overline{B} = B + \frac{1}{(c,p)}\left[(y-Bp)c^T + c(y-Bp)^T\left(I - \frac{pc^T}{(c,p)}\right)\right].$$

Uvođenjem oznake $H := B^{-1}$ dobivamo

$$\overline{B} = B\left[I + \frac{H(y-Bp)}{(c,p)}c^T + Hc\frac{(y-Bp)^T}{(c,p)}\left(I - \frac{pc^T}{(c,p)}\right)\right].$$

Sada vrijedi

$$\det \overline{B} = \det B \det\left(I + u_1u_2^T + u_3u_4^T\right),$$

gdje je

$$u_1 = \frac{Hy-p}{(c,p)}, \quad u_2 = c,$$
$$u_3 = Hc, \qquad u_4 = \left(I - \frac{cp^T}{(c,p)}\right)\frac{(y-Bp)}{(c,p)}.$$

Prema *lemi* 10. je

$$\det\left(I + u_1u_2^T + u_3u_4^T\right) = (1 + (u_1,u_2))\,(1 + (u_3,u_4)) - (u_1,u_4)(u_2,u_3),$$

gdje je

$$\begin{aligned}
1 + (u_1,u_2) &= 1 + c^T\frac{Hy-p}{(c,p)} = 1 + \frac{c^THy}{(c,p)} - \frac{c^Ty}{(c,p)} = \frac{c^THy}{(c,p)}\\
1 + (u_3,u_4) &= 1 + c^T\left(H - \frac{Hcp^T}{(c,p)}\right)\frac{(y-Bp)}{(c,p)}\\
&= 1 + \frac{c^T}{(c,p)}\left(Hy - \frac{Hcp^Ty}{(c,p)} - s + \frac{Hcp^TBp}{(c,p)}\right) =\\
&= 1 + \frac{c^THy}{(c,p)} - \frac{cTHcp^Ty}{(c,p)^2} - 1 + \frac{cTHcp^TBp}{(c,p)^2},\\
(1 + (u_1,u_2))\,(1 + (u_3,u_4)) &= \frac{(c,Hy)^2}{(c,p)^2} - \frac{(c,Hy)(c,Hc)(p,y)}{(c,p)^3} + \frac{(c,Hy)(c,Hc)(p,Bp)}{(c,p)^3}\\
(u_1,u_4) &= \frac{y^TH-p^T}{(c,p)^2}\left(I - \frac{cp^T}{(c,p)}\right)(y-Bp)\\
&= \frac{y^THy}{(c,p)^2} - \frac{p^Ty}{(c,p)^2} - \frac{y^THcp^Ty}{(c,p)^3} + \frac{y^THcp^TBp}{(c,p)^3}
\end{aligned}$$

iz čega slijedi

$$\det\left(I + u_1u_2^T + u_3u_4^T\right) = \frac{(c,Hy)^2}{(c,p)^2} - \frac{(c,Hc)(y,Hy)}{(c,p)^2} + \frac{(c,Hc)(y,p)}{(c,p)^2}$$

i tražena formula (4.68).

Ako pretpostavimo da je $B > 0$, onda je i $H := B^{-1} > 0$, pa možemo definirati $v = H^{1/2}y$ i $w = H^{1/2}c$. Time (4.68) postaje

$$\det \overline{B} = \det B\left[\left((v,w)^2 - \|v\|^2\|w\|^2 + \|w\|^2(y,p)\right)\frac{1}{(c,p)^2}\right]. \qquad (4.69)$$

Na osnovi *teorema 16.* zaključujemo da je $\overline{B} > 0$ onda i samo onda ako vrijedi

$$\|w\|^2(y,p) > \|v\|^2\|w\|^2 - (v,w)^2. \tag{4.70}$$

Jednostavan način kako se može zadovoljiti ova nejednakost koju uveo Davidon (1959.) (kasnije objašnjeno i implementirano u radu Fletchera i Powella, 1963.). Naime, ako su w, v (odnosno c, y) linearno zavisni, onda je (4.70) zadovoljeno ako je $(y,p) > 0$. Zato se poznata *Davidon-Fletcher-Powell* korekcija (DFP) ranga 2 dobiva iz (4.65) uz $c = y$

$$\begin{aligned} \overline{B}_{DGP} &= B + \tfrac{(y-Bp)y^T + y(y-Bp)^T}{(y,p)} - \tfrac{(y-Bp,p)}{(y,p)^2}yy^T. \\ &= \left(I - \tfrac{yp^T}{(y,p)}\right) B \left(I - \tfrac{py^T}{(y,p)}\right) + \tfrac{yy^T}{(y,p)}. \end{aligned} \tag{4.71}$$

Zadatak 29. Provjerite valjanost formula (4.71)!

Vrijedi sljedeći teorem:

Teorem 17. *Neka je $B \in \mathbb{R}^{n\times n}$ regularna simetrična matrica i neka je $\overline{B}_{DFP} \in \mathbb{R}^{n\times n}$ definiran s (4.71) za neke vektore $y, p \in \mathbb{R}^n$, koji zadovoljavaju $(y,p) \neq 0$.*

Tada je $\overline{B}_{DFP}$ regularna onda i samo onda ako je $(y, Hy) \neq 0$, gdje je $H := B^{-1}$.

Ako je $\overline{B}_{DFP}$ regularna, tada se $\overline{H}_{DFP} := \overline{B}_{DFP}^{-1}$ može zapisati kao

$$\overline{H}_{DFP} = H + \tfrac{pp^T}{(y,p)} - \tfrac{Hyy^TH}{(y,Hy)}. \tag{4.72}$$

Nadalje, ako je $B > 0$, onda je i $\overline{B}_{DFP} > 0$ onda i samo onda ako vrijedi $(y,p) > 0$.

Zadatak 30. Provjerite valjanost formula (4.71) i dokažite *teorem 17.*

Inverzne korekcije

U prethodnim razmatranjima nastojali smo dobiti rekurzivne formule za matrice–korekcije, koje sve više nalikuju na Hessijan minimizirajuće funkcije. Sada ćemo pokušati definirati rekurzivne formule za matrice–korekcije, koje će sve više nalikovati inverznim Hessijan minimizirajućim funkcijama. Te matrice nazivat ćemo *inverzne korekcije*.

U tu svrhu pretpostavimo da posjedujemo neku aproksimaciju inverznog Hessijana H $[\nabla^2 f(x)]^{-1}$ i pomoću nje pokušajmo pronaći što bolju aproksimaciju $\overline{H}$ inverznog Hessijana $[\nabla^2 f(\overline{x})]^{-1}$, gdje je $\overline{x} = x + p$. Analogno kvazi-Newtonovom uvjetu (4.54) ($\overline{B}p = y$) postavimo *inverzni kvazi-Newtonov uvjet*

$$\overline{H}y = p. \tag{4.73}$$

Odmah je jasno da opća inverzna korekcija ranga 1, koja zadovoljava inverzni kvazi-Newtonov uvjet (4.73) mora biti oblika

$$\overline{H} = H + \frac{(p-Hy)d^T}{(d,y)}, \qquad \text{za neki} \quad d \in \mathbb{R}^n, \quad (d,y) \neq 0. \tag{4.74}$$

Zadatak 31. Pokažite da korekcija ranga 1 zadana s (4.56) i inverzna korekcija ranga 1 zadana s (4.74) predstavljanju istu korekciju ako je $c = B^T d$.

Kao u slučaju prethodno razmatranih korekcija koje aproksimiraju Hessijan $(\nabla^2 f(x))$ i ovdje se najprije postavlja pitanje čuvanja nasljeđa simetričnosti, tj. vrijedi li

$$H \quad \text{je simetrična matrica} \quad \Longrightarrow \quad \overline{H} \quad \text{je simetrična matrica ?}$$

Zadatak 32. Pokažite da je inverzna korekcija ranga 1 koja zadovoljava inverzni kvazi-Newtonov uvjet (4.73) i čuva simetričnost također zadana formulom (4.63).

Zadatak 33. Slično kao u dokazu *leme 8.* izvedite formulu za inverznu kvazi-Newtonovu korekcija ranga 2

$$\overline{H} = H + \frac{(p - Hy)d^T + d(p - Hy)^T}{(d, y)} - \frac{(p - Hy)^T y}{(d, y)^2} dd^T. \tag{4.75}$$

Najvažniju varijantu inverzne kvazi-Newtonove korekcije ranga 2 (4.75), koja pri tome čuva i pozitivnu definitnost, nezavisno su napravili Broyden, Fletcher, Goldfarb i Shano 1970. godine, pa se ona u literaturi naziva *Broyden–Fletcher–Goldfarb–Shano korekcija* ili skraćeno *BFGS-korekcija*. Dobiva se tako da u (4.75) stavimo $d = p$. Može se pokazati da u tom slučaju vrijedi

$$\overline{H}_{BFGS} = \left(I - \frac{py^T}{(y, p)} \right) H \left(I - \frac{yp^T}{(y, p)} \right) + \frac{pp^T}{(y, p)}. \tag{4.76}$$

kao i sljedeći teorem:

Teorem 18. *Neka je $H \in \mathbb{R}^{n\times n}$ regularna simetrična matrica i neka je $\overline{H}_{BFGS} \in \mathbb{R}^{n\times n}$ definirana s (4.76) za neke vektore $y, p \in \mathbb{R}^n$, $(y, p) \neq 0$.*

Tada je $\overline{H}_{BFGS}$ regularna onda i samo onda ako je $(p, Bp) \neq 0$, gdje je $B = H^{-1}$.

Ako je $\overline{H}_{BFGS}$ regularna, onda se $\overline{B}_{BFGS} := \overline{H}_{BFGS}^{-1}$ može zapisati kao

$$\overline{B}_{BFGS} = B + \frac{yy^T}{(y, p)} - \frac{Bpp^T B}{(p, Bp)}.$$

Nadalje, ako je $H > 0$, onda je $\overline{H}_{BFGS} > 0$ onda i samo onda ako je $(y, p) > 0$.

Zadatak 34. Provjerite valjanost formule (4.76) i dokažite *teorem 18.*

POGLAVLJE 5

Problem uvjetne minimizacije

U ovom poglavlju pokazat ćemo kako se tehnike iz prethodnih poglavlja mogu primijeniti na optimizacijske probleme uvjetne minimizacije.

Do sada nam je x_k uglavnom predstavljalo x_k-tu aproksimaciju, a tu ćemo oznaku koristiti i u ovom poglavlju. Budući da ćemo u ovom poglavlju često računati i s komponentama k-te aproksimacije, oznaka $(x_k)_i$ predstavljat će nam i-tu koordinatu k-te aproksimacije.

Kod uvjetne minimizacije pretpostavit ćemo da su ograničenja jednostavna. U tu svrhu neka su $L_1, L_2, \ldots, L_n$ i $U_1, U_2, \ldots, U_n$ realni brojevi takvi da vrijedi

$$-\infty < L_i < U_i < +\infty, \quad i = 1, \ldots, n.$$

Problem uvjetne minimizacije problem je određivanja lokalnog minimuma $x^\star$ funkcije f s n varijabli pri čemu je zadano ograničenje:

$$x^\star \in \Omega = \{x \in \mathbb{R}^n : L_i \leq (x)_i \leq U_i\}.$$

Odnosno, zahtijevamo da je $x^\star$ takav da vrijedi

$$f(x^\star) \leq f(x) \quad \text{za sve } x \in \Omega \text{ blizu } x^\star.$$

Problem se obično zapisuje kao

$$\min_{x \in \Omega} f(x). \tag{5.1}$$

Budući da je skup Ω kompaktan, uvijek postoji rješenje gore navedenog minimizacijskog problema [24].

Nejednakosti $L_i \leq (x)_i \leq U_i, \quad i = 1, \ldots, n$ zovemo ograničenjima, a reći ćemo da je i-to ograničenje aktivno za $x \in \Omega$ ako je ili $(x)_i = L_i$ ili $(x)_i = U_i$. U suprotnom, reći ćemo da je i-to ograničenje neaktivno. Skup svih indeksa i tako da je i-to ograničenje aktivno (neaktivno) zvat ćemo skupom aktivnih (neaktivnih) indeksa za x i označivat ćemo ga s $\mathcal{A}(x)$ ($\mathcal{N}(x)$).

5.1 Nužni i dovoljni uvjeti za optimalnost

U 2. poglavlju već smo proučavali nužne i dovoljne uvjete da bi točka $x^\star$ bila minimum funkcije. Posebnu pažnju posvetili smo konveksnim funkcijama. Ovdje ćemo neke od tih rezultata generalizirati za slučaj uvjetne minimizacije.

Ako promatramo problem s ograničenjem za realnu funkciju jedne varijable, tada se domena funkcije restringira na segment $[a, b]$. Tada moramo promijeniti nužne uvjete tako da oni dopuštaju mogućnost da se minimum $x^\star$ nalazi u jednom od krajeva. Ako je $x^\star = a$ lokalni minimum, tada ne mora biti $f'(a) = 0$, međutim, budući da je a lokalni minimum, $f(x) \geq f(a)$ za sve $a \leq x$ dovoljno blizu a, stoga je i $f'(a) \geq 0$. Međutim ništa ne možemo reći o f''. Ako je f diferencijabilna na (a, b), tada su nužni uvjeti za sve tri mogućnosti $x^\star = a$, $x^\star = b$ i $a < x^\star < b$ iskazani sljedećim teoremom.

Teorem 19. *Neka je f derivabilna funkcija jedne varijable na segmentu $[a, b]$. Neka je $x^\star$ lokalni minimum funkcije f na $[a, b]$. Tada je*

$$f'(x^\star)(x - x^\star) \geq 0 \quad \text{za sve} \quad x \in [a, b]. \tag{5.2}$$

Nadalje, ako je f dva puta derivabilna na $[a, b]$, tada vrijedi

$$f''(x)(x^\star - a)(b - x^\star) \geq 0.$$

Izraz (5.2) poslužit će nam kao motivacija za sljedeću definiciju stacionarnosti.

Definicija 11. Točka $x^\star \in \Omega$ stacionarna je za problem (5.1) ako je

$$\nabla f(x^\star)^T (x - x^\star) \geq 0 \quad \text{za sve} \quad x \in \Omega.$$

Uočimo da stacionarna točka funkcije definirana u poglavlju 2. (x je stacionarna točka funkcije definirane na otvorenom skupu ako je $\nabla f(x) = 0$) zadovoljava uvjet (5.2). Štoviše, vrijedi da $x^\star \in \mathrm{Int}\Omega$ zadovoljava gornji uvjet ako i samo ako $\nabla f(x^\star) = 0$.

Slično kao u slučaju kada nema ograničenja, može se pokazati činjenica da optimalnost povlači stacionarnost.

Teorem 20. *Neka je funkcija f diferencijabilna na konveksnom skupu Ω i neka je $x^\star$ rješenje problema (5.1). Tada je $x^\star$ stacionarna točka za problem (5.1).*

Dokaz. Neka je $x^\star$ rješenje problema (5.1) i $y \in \Omega$. Budući da je Ω konveksan, dužina određena točkama $x^\star$ i y je u Ω. Stoga, funkcija

$$\Phi(t) = f(x^\star + t(y - x^\star))$$

definirana je za svaki $t \in [0, 1]$ i ima lokalni minimum $t = 0$, stoga po teoremu 19

$$0 \leq \Phi'(0)(t - 0) \quad \text{za sve} \quad t \in [0, 1],$$

odnosno

$$0 \leq \Phi'(0) = \nabla f(x^\star)^T (y - x^\star),$$

što smo i trebali pokazati. $\square$

Kako bismo ilustrirali ulogu druge derivacije promotrit ćemo funkciju dvije varijable. Neka je $n = 2$ i neka funkcija f'' zadovoljava Lipschitzov uvjet $\|f''(x) - f''(y)\| \le \lambda\|x - y\| \quad \forall x, y \in \Omega = [0, 1] \times [0, 1]$. Ako je $x^\star$ rješenje od (5.1) i nijedno ograničenje nije aktivno, tada je $f''(x^\star)$ pozitivno semidefinitna matrica uz isto argumentiranje kao kada nemamo ograničenja. Međutim, ako je jedno ili više ograničenja aktivno, tada ne možemo ništa zaključiti o pozitivnoj definitnosti matrice $f''(x^\star)$. Pretpostavimo da je minimum funkcije u $x^\star = (\xi, 0)$, gdje je $0 < \xi < 1$. Ništa ne možemo reći o $\partial^2 f(x^\star)/\partial x_2^2$, ali funkcija $\phi(t) = f(t, 0)$ definirana na $[0, 1]$ mora zadovoljavati

$$\phi''(\xi) = \partial^2 f(x^\star)/\partial x_1^2 \ge 0.$$

Doprinosi drugih parcijalnih derivacija koji odgovaraju neaktivnim ograničenjima moraju biti nenegativni, dok ništa ne možemo reći o doprinosima koji odgovaraju aktivnim ograničenjima.

Kako bismo upravo izloženu ideju preciznije iskazali, definirat ćemo reducirani Hessijan.

Definicija 12. Neka je f dva puta diferencijabilna u $x \in \Omega$. Reducirani Hessijan $f_R''(x)$ matrica je čiji su elementi definirani s

$$(f_R''(x))_{ij} = \begin{cases} \delta_{ij}, & \text{ako je} \quad i \in \mathcal{A}(x) \text{ ili } j \in \mathcal{A}(x) \\ (f''(x))_{ij}, & \text{inače} \end{cases}.$$

Sljedeći teorem daje nužni uvjet koristeći pojam reduciranog Hessijana.

Teorem 21. *Neka f zadovoljava Lipschitzov uvjet $\|f''(x) - f''(y)\| \le \lambda\|x - y\|$ $\forall x, y \in \mathbb{R}^n$ i neka je $x^\star$ rješenje problema (5.1). Tada je reducirani Hessijan $f_R''(x^*)$ pozitivno semidefinitna matrica.*

Dokaz. Pretpostavimo da ima m neaktivnih indeksa i $n - m$ aktivnih. Particioniramo $x \in \Omega$ (po potrebi ispermutiramo varijable) tako da je $x = (u, v)$, gdje u odgovara neaktivnim indeksima, dok v odgovara aktivnim.

Tada preslikavanje $\phi(u) = f(u, v^\star)$ ima lokalni minimum $u^\star \in \mathbb{R}^m$ (ako je minimum funkcije bez ograničenja) i stoga je $\phi''(u^*)$ pozitivno semidefinitna matrica. Budući da je tada reducirani Hessijan jednak

$$f_R''(x^*) = \begin{pmatrix} \phi''(u^*) & 0 \\ 0 & I \end{pmatrix}$$

slijedi da je i $f_R''(x^*)$ pozitivno semidefinitna matrica. □

Neka $\mathcal{P}$ označava projekciju na Ω. To je preslikavanje koje varijabli x pridružuje najbližu točku (u l_2 normi) iz skupa Ω. Tada vrijedi

$$(\mathcal{P}(x))_i = \begin{cases} L_i, & \text{ako je} \quad (x)_i \le L_i \\ (x)_i, & \text{ako je} \quad L_i < (x)_i < U_i \\ U_i, & \text{ako je} \quad (x)_i \ge U_i \end{cases}.$$

Jedna karakterizacija stacionarne točke dana je sljedećim teoremom.

Teorem 22. *Neka je f neprekidno diferencijabilna. Točka $x^\star \in \Omega$ stacionarna je za problem (5.1) ako i samo ako je*

$$x^\star = \mathcal{P}(x^\star - \alpha \nabla f(x^\star)),$$

za sve $\alpha \geq 0$.

Ovaj ćemo teorem dokazati u poglavlju analize konvergencije jer ćemo tada imati sve tvrdnje koje su potrebne za dokaz.

U nastavku ćemo navesti nekoliko rezultata vezanih uz dovoljne uvjete koristeći pojam reduciranog Hessijana. Ako je $x^\star$ stacionarna točka te $i \in \mathcal{N}(x^\star)$, tada je $x^\star \pm te_i \in \Omega$ za sve t dovoljno male (e_i je i-ti kanonski vektor).

Budući da je

$$\frac{df(x^\star \pm te_i)}{dt}|_{t=0} = \pm \nabla f(x^\star) e_i \geq 0,$$

stoga

$$(\nabla f(x^\star))_i = 0 \quad \text{za sve} \quad i \in \mathcal{N}(x^\star).$$

Pojam nedegenerirane stacionarne točke problema koristit ćemo u formulaciji dovoljnih uvjeta. Navedeni pojam važan je i za definiranje zaustavnog kriterija.

Definicija 13. Točka $x^\star \in \Omega$ je nedegenerirana stacionarna točka problema (5.1) ako je $x^\star$ stacionarna točka i vrijedi

$$(\nabla f(x^\star))_i \neq 0 \quad \text{za sve} \quad i \in \mathcal{A}(x^\star).$$

Ako je $x^\star$ i rješenje problema (5.1), tada kažemo da je $x^\star$ nedegenerirani lokalni minimum.

Ako je $\mathcal{S}$ proizvoljan skup indeksa, definiramo

$$(\mathcal{P}_{\mathcal{S}} x)_i = \begin{cases} (x)_i, & \text{ako je} \quad (x)_i \in \mathcal{S} \\ 0, & \text{inače} \end{cases}.$$

Prvo ćemo dokazati lemu u kojoj se koristi pojam nedegeneriranosti, a koristit ćemo je u nastavku poglavlja.

Lema 11. *Neka je $x^\star$ nedegenerirana stacionarna točka. Pretpostavimo da $\mathcal{A}(x^\star)$ nije prazan. Tada postoji σ takav da vrijedi*

$$\nabla f(x^\star)^T (x - x^\star) = \nabla f(x^\star)^T \mathcal{P}_{\mathcal{A}(x^\star)}(x - x^\star) \geq \sigma \|\mathcal{P}_{\mathcal{A}(x^\star)}(x - x^\star)\|$$

za sve $x \in \Omega$.

Dokaz. Postoji $\sigma > 0$ tako da ako je $i \in \mathcal{A}(x^\star)$, tada nedegeneriranost i stacionarnost povlače da je ili

$$(x^\star)_i = L_i \quad \text{i} \quad (\nabla f(x^\star))_i \geq \sigma \quad \text{ili} \quad (x^\star)_i = U_i \quad \text{i} \quad (\nabla f(x^\star))_i \leq -\sigma.$$

Ako je $x \in \Omega$, tada za sve $i \in \mathcal{A}(x^\star)$

$$(\nabla f(x^\star))_i (x - x^\star)_i \geq \sigma |(x - x^\star)_i|.$$

Budući da je $\|x\|_1 \geq \|x\|$, vrijedi nejednakost

$$\nabla f(x^\star)^T \mathcal{P}_{\mathcal{A}(x^\star)}(x - x^\star) \geq \sigma \|\mathcal{P}_{\mathcal{A}(x^\star)}(x - x^\star)\|,$$

što je i trebalo pokazati. □

Iskazat ćemo teorem koji daje dovoljan uvjet, a sličan je pripadnom teoremu iz slučaja kada nemamo ograničenja.

Teorem 23. *Neka je $x^\star \in \Omega$ nedegenerirana stacionarna točka problema (5.1). Neka je f dva puta diferencijabilna u okolini točke $x^\star$. Pretpostavimo da je reducirani Hessijan u $x^\star$ pozitivno definitna matrica. Tada je $x^\star$ rješenje problema (5.1) (stoga je i nedegenerirani lokalni minimum).*

Dokaz. Neka je $x \in \Omega$. Definirajmo funkciju ϕ s $\phi(t) = f(x^\star + t(x - x^\star))$. Pokazat ćemo tvrdnju tako da pokažemo da je

$$\text{ili} \quad (i) \quad \phi'(0) > 0 \quad \text{ili} \quad \text{(ii)} \quad \phi'(0) = 0, \phi''(0) > 0.$$

Neka je $e = x - x^\star$ i uočimo da je

$$\phi'(0) = \nabla f(x^\star)^T e = \nabla f(x^\star)^T (\mathcal{P}_{\mathcal{A}(x^\star)} e + \mathcal{P}_{\mathcal{N}(x^\star)} e).$$

Stacionarnost točke povlači da $\nabla f(x^\star)^T \mathcal{P}_{\mathcal{N}(x^\star)} e = 0$.
Ako je $\mathcal{P}_{\mathcal{A}(x^\star)} e \neq 0$, tada nedegeneriranost povlači da je

$$\nabla f(x^\star)^T \mathcal{P}_{\mathcal{A}(x^\star)} e > 0$$

i stoga vrijedi (i).

U suprotnom, ako je $\mathcal{P}_{\mathcal{A}(\S^\star)} e = 0$, tada je

$$\phi''(0) = (x - x^\star)^T \mathcal{P}_{\mathcal{N}(x^\star)} f''(x^\star) \mathcal{P}_{\mathcal{N}(x^\star)} (x - x^\star) > 0,$$

stoga vrijedi (ii). □

5.2 Gradijentna metoda s projekcijom

U ovom poglavlju definirat ćemo algoritam gradijentne metode s projekcijom koji je prirodna primjena gradijentne metode na problem uvjetne minimizacije. Ako imamo aproksimaciju x_c, tada je nova aproksimacija dana s

$$x_+ = \mathcal{P}(x_c - \alpha \nabla f(x_c)),$$

gdje je α duljina koraka određena algoritmom, slično kao u slučaju kada nema ograničenja. U ovom poglavlju, radi jednostavnosti zapisivanja, za $\alpha > 0$ uvodimo $x(\alpha)$ koji označava sljedeću aproksimaciju ako je duljina koraka α. U slučaju gradijentne metode imamo

$$x(\alpha) = \mathcal{P}(x - \alpha \nabla f(x)).$$

Duljinu koraka α određivat ćemo iz uvjeta

$$f(x(\alpha)) - f(x) \leq \tfrac{-c}{\alpha} \|x - x(\alpha)\|^2. \tag{5.3}$$

Pri određivanju duljine koraka potrebno je odrediti najmanji prirodni broj m tako da za $\alpha = \beta^m$ vrijedi gornja nejednakost. Parametar c obično se stavlja na 10^{-4} [10].

Algoritam za gradijentnu metodu dan je sljedećim algoritmom.

Algoritam 5 . Algoritam gradijentne metode sa projekcijom

---

for i:= $1, \ldots, nmax$ **do**
  izračunaj f i ∇f; provjeri zaustavni kriterij
  odredi najmanji prirodni broj m tako da vrijedi nejednakost (5.3) za $\alpha = \beta^m$
  $x = x(\alpha)$
end for
if $n = nmax$ i zaustavni kriterij nije zadovoljen **then**
  Ispiši: "Algoritam nije odredio zadovoljavajuću aproksimaciju";
end if

U nastavku ćemo uvesti zaustavni kriterij.

Zaustavni kriterij

Budući da promatramo minimizaciju funkcije uz ograničenja, moramo prilagoditi zaustavni kriterij koji će na odgovarajući način uključiti ograničenja. Primijetimo da prilikom rješavanja problema (5.1) ∇f ne mora biti nula. Zaustavni kriterij koji se nameće u slučaju minimizacije funkcije s ograničenjima uključuje razliku x i $x(1)$. Ako je razlika mala, onda možemo zaustaviti iterativni proces. Kako bismo potvrdili da je razlika x i $x(1)$ dobra mjera za zaustavljanje, izvest ćemo dovoljne uvjete koji potvrđuju opravdanost ovog zaustavnog kriterija.

Prisjetimo se oznake $e = x - x^\star$. Prvo ćemo iskazati pomoćni rezultat koji nam je potreban za dokaz teorema.

Lema 12. *Neka je f dva puta neprekidno diferencijabilna na Ω i neka je $x^\star$ nedegenerirana stacionarna točka problema (5.1). Neka je $\alpha \in (0,1]$. Tada za x dovoljno blizu $x^\star$ vrijedi*

1. $\mathcal{A}(x) \subset \mathcal{A}(x^\star)$ *i* $(x)_i = (x^\star)_i$ *za sve* $i \in \mathcal{A}(x)$,
2. $\mathcal{A}(x(\alpha)) = \mathcal{A}(x^\star)$ *i* $(x(\alpha))_i = (x^\star)_i$ *za sve* $i \in \mathcal{A}(x^\star)$.

Dokaz. Neka je $\delta_1 = \min_{i \in \mathcal{N}(x^\star)} \{(U_i - (x^\star)_i), (x^\star)_i - L_i, (U_i - L_i)/2\}$. Ako je $i \in \mathcal{N}(x^\star)$ i $\|e\| < \delta_1$, tada $L_i < (x)_i < U_i$. Stoga

$$\mathcal{N}(x^\star) \subset \mathcal{N}(x),$$

čime smo dokazali tvrdnju da je $\mathcal{A}(x) \subset \mathcal{A}(x^\star)$. Štoviše, kako je $i \in \mathcal{A}(x) \subset \mathcal{A}(x^*)$, $(x)_i = L_i$ ili $x_i = U_i$ te zbog činjenice da je

$$\|e\| < \delta_1 \leq \min\{(U_i - L_i)/2\},$$

imamo $(x)_i = (x^\star)_i$ za sve $i \in \mathcal{A}(x)$ čime smo dokazali tvrdnju 1.

Neka je $i \in \mathcal{A}(x^\star)$. Prema *lemi* 11. i zbog neprekidnosti od ∇f postoji δ_2 tako da ako je $\|e\| < \delta_2$, tada

$$(\nabla f(x^\star + e))_i (x - x^\star)_i > \sigma |x - x^\star|_i / 2.$$

Stoga, ako je

$$\|e\| < \delta_3 < \min\{\sigma/2, \delta_2\},$$

tada je $i \in \mathcal{A}(x(\alpha))$ i $(x(\alpha))_i = (x^\star)_i$, čime smo dokazali da je $\mathcal{A}(x^\star) \subset \mathcal{A}(x(\alpha))$. Stoga, treba još pokazati da $\mathcal{A}(x^\star) \supset \mathcal{A}(x(\alpha))$.

Iz definicije projekcije $\mathcal{P}$ imamo

$$\|\mathcal{P}(x) - \mathcal{P}(y)\| \leq \|x - y\|$$

za sve $x, y \in \mathbb{R}^n$. Budući da je $\nabla^2 f$ neprekidna, postoji Lipschitzova konstanta L za ∇f u Ω. Ako iskoristimo činjenicu da je $x^\star$ stacionarna točka uz *teorem* 22. imamo

$$x^\star = x^\star(\alpha) = \mathcal{P}(x^\star - \alpha \nabla f(x^\star)).$$

Nadalje,

$$\begin{aligned} \|x^\star - x(\alpha)\| &= \|\mathcal{P}(x^\star - \alpha \nabla f(x^\star)) - \mathcal{P}(x - \alpha \nabla f(x))\| \\ &\leq \|e\| + \alpha \|\nabla f(x^\star) - \nabla f(x)\| \\ &\leq (1 + L\alpha)\|e\|. \end{aligned} \tag{5.4}$$

Ako bi postojao $i \in \mathcal{A}(x(\alpha)) \cap \mathcal{N}(x^\star)$, tada bi vrijedilo

$$\|x^\star - x(\alpha)\| \geq \delta_1 = \min_{i \in \mathcal{N}(x^\star)} \{(U_i - (x^\star)_i), ((x^\star)_i - L_i)\}. \tag{5.5}$$

Međutim, ako je

$$\|e\| < \delta_4 = \min\{\delta_3, \delta_1/(1 + \alpha L)\},$$

tada nejednakost (5.4) povlači da (5.5) ne može vrijediti, stoga je pretpostavka da je $i \in \mathcal{A}(x(\alpha)) \cap \mathcal{N}(x^\star)$ kriva. Znači, ako je $i \in \mathcal{A}(x(\alpha))$, onda je $i \in \mathcal{A}(x^\star)$, čime smo dokazali obrnutu inkluziju. $\square$

Teorem 24. *Neka je f dva puta neprekidno diferencijabilna na Ω, neka je $x^\star$ nedegenerirana stacionarna točka problema (5.1) i neka $\mathcal{A}(x^\star)$ nije prazan skup. Tada postoje δ i M tako da ako je $\|e\| < \delta$ i $\mathcal{A}(x) = \mathcal{A}(x^\star)$ te vrijedi*

$$\frac{\|e\|}{M} \leq \|x - x(1)\| \leq M\|e\|.$$

Dokaz. Neka je L Lipschitzova konstanta od ∇f u Ω. Koristeći stacionarnost imamo

$$\begin{aligned}
\|x - x(1)\| &= \|e - (x(1) - x^\star(1))\| \\
&\leq \|e\| + \|\mathcal{P}(x - \nabla f(x)) - \mathcal{P}(x^\star - \nabla f(x^\star))\| \\
&\leq 2\|e\| + \|\nabla f(x) - \nabla f(x^\star)\| \leq (2 + L)\|e\|.
\end{aligned}$$

Stoga, uz konstantu $M \geq 2 + L$ vrijedi desna nejednakost.

Kako bismo dokazali lijevu nejednakost primijenit ćemo *lemu* 12. Neka je δ_1 takav da je $\|e\| < \delta_1$, tada *lema* 12. vrijedi za $\alpha = 1$. Pomoću leme dobivamo

$$(x - x(1))_i = \begin{cases} (\nabla f(x))_i, & \text{ako je} \quad i \in \mathcal{N}(x^\star) \\ (e_i) = 0, & \text{ako je} \quad i \in \mathcal{A}(x^\star) \end{cases}.$$

Ostaje nam provjeriti slučaj kada je $i \in \mathcal{N}(x) = \mathcal{N}(x^\star)$. Dovoljni uvjeti povlače da postoji $\mu > 0$ tako da vrijedi

$$u^T \mathcal{P}_{\mathcal{N}(x^\star)} \nabla^2 f(x^\star) \mathcal{P}_{\mathcal{N}(x^\star)} u \geq \mu \|\mathcal{P}_{\mathcal{N}(x^\star)} u\|^2$$

za sve $u \in \mathbb{R}^n$. Stoga postoji δ_2 tako da je $\|e\| < \delta_2$. Tada vrijedi

$$u^T \mathcal{P}_{\mathcal{N}(x^\star)} \nabla^2 f(x) \mathcal{P}_{\mathcal{N}(x^\star)} u \geq \mu \|\mathcal{P}_{\mathcal{N}(x^\star)} u\|^2/2$$

za sve $u \in \mathbb{R}^n$. Nadalje, budući da je $e = \mathcal{P}_{\mathcal{N}(x^\star)} e$, imamo

$$\begin{aligned}
\|\mathcal{P}_{\mathcal{N}(x^\star)}(x - x(1))\|^2 &= \int_0^1 e^T \mathcal{P}_{\mathcal{N}(x^\star)} \nabla^2 f(x + te) e \, dt \\
&= \int_0^1 e^T \mathcal{P}_{\mathcal{N}(x^\star)} \nabla^2 f(x + te) \mathcal{P}_{\mathcal{N}(x^\star)} e \, dt \\
&\geq \mu \|\mathcal{P}_{\mathcal{N}(x^\star)} e\|^2/2.
\end{aligned}$$

Stoga vrijedi da je $\|x - x(1)\| \geq \min\{1, \sqrt{\mu/2}\}\|e\|$. Ako odaberemo $M = \max\{2 + L, 1, \sqrt{2/\mu}\}$, tada su zadovoljene obje nejednakosti. □

Zaustavni kriterij definirat ćemo tako da u sebi sadrži relativnu i apsolutnu mjeru stacionarnosti $\|x - x(1)\|$. Neka je $r_0 = \|x_0 - x_0(1)\|$, a relativna i apsolutna tolerancija neka su dane sa τ_r i τ_a, respektivno. Tada se u algoritmu gradijentne metode s projekcijom za zaustavni kriterij uzima nejednakost

$$\|x - x(1)\| \leq \tau_a + \tau_r r_0. \tag{5.6}$$

Analiza konvergencije

Za razliku od analize konvergencije za standardnu gradijentnu metodu sada posebnu pažnju moramo posvetiti ograničenjima. Analizu ćemo započeti lemom.

Lema 13. *Za sve* $x, y \in \Omega$

$$(y - x(\alpha))^T (x(\alpha) - x + \alpha \nabla f(x)) \geq 0.$$

Dokaz. Koristeći definiciju $\mathcal{P}$ vidimo da je

$$\|x(\alpha) - x + \alpha \nabla f(x)\| \leq \|y - x + \alpha \nabla f(x)\|$$

za sve $y \in \Omega$. To znači da je $t = 0$ lokalni minimum funkcije

$$\phi(t) = \|(1-t)x(\alpha) + ty - x + \alpha \nabla f(x)\|^2 / 2$$

na $[\,0,\, 1\,]$ pa vrijedi

$$0 \leq \phi'(0) = (y - x(\alpha))^T (x(\alpha) - x + \alpha \nabla f(x)),$$

što je i trebalo pokazati. □

Nejednakost iz prethodne leme obično ćemo koristiti u ekvivalentnoj formi

$$(x - x(\alpha))^T (y - x(\alpha)) \leq \alpha \nabla f(x)^T (y - x(\alpha)). \tag{5.7}$$

Ako u prethodni izraz uvrstimo $y = x$ dobit ćemo sljedeći korolar.

Korolar 3. *Za sve* $x \in \Omega$ *i* $\alpha \geq 0$ *vrijedi*

$$\|x - x(\alpha)\|^2 \leq \alpha \nabla f(x)^T (x - x(\alpha)).$$

Može se pokazati da ako za duljinu koraka koristimo nejednakost (5.3) uz pretpostavku o Lipschitz neprekidnosti, tada će duljina koraka biti pozitivan broj što je preciznije pokazano sljedećim teoremom.

Teorem 25. *Pretpostavimo da je* ∇f *Lipschitz neprekidna s Lipschitzovom konstantom* L *i neka je* $x \in \Omega$. *Tada je nejednakost za određivanje koraka (5.3) zadovoljena za svaki* α *koji zadovoljava*

$$0 < \alpha \leq \tfrac{2(1-c)}{L}.$$

Dokaz. Uzmimo da je $y = x - x(\alpha)$. Tada imamo

$$f(x - y) - f(x) = f(x(\alpha)) - f(x) = -\int_0^1 \nabla f(x - ty)^T y \, dt.$$

Nadalje vrijedi

$$f(x(\alpha)) = f(x) + \nabla f(x)^T (x(\alpha) - x) - \int_0^1 (\nabla f(x - ty) - \nabla f(x))^T y \, dt.$$

Nakon sređivanja prethodni izraz može se zapisati kao

$$\alpha(f(x) - f(x(\alpha))) = \alpha\nabla f(x)^T(x - x(\alpha)) + \alpha E,$$

gdje je $E = \int_0^1 (\nabla f(x - ty) - \nabla f(x))^T y\, dt$. Ako iskoristimo Lipschitz neprekidnost dobit ćemo

$$\|E\| \leq L\|x - x(\alpha)\|^2/2.$$

Nadalje imamo

$$\alpha(f(x) - f(x(\alpha))) \geq \alpha\nabla f(x)^T(x - x(\alpha)) - \alpha L\|x - x(\alpha)\|^2/2.$$

Ako sada iskoristimo tvrdnju iz *Korolara 3.* dobit ćemo

$$\alpha(f(x) - f(x(\alpha))) \geq (1 - \alpha L/2)\|x - x(\alpha)\|^2,$$

iz čega slijedi tvrdnja teorema. □

Prema oznakama iz algoritma za gradijentnu metodu s projekcijom možemo zaključiti da će određivanje koraka biti zadovoljeno, ako ne prije, onda kada je

$$\beta^m \leq \tfrac{2(1-c)}{L} \leq \beta^{m-1} \quad \Rightarrow \quad \beta^m \geq \tfrac{2\beta(1-c)}{L}.$$

Stoga je donja granica za duljinu koraka

$$\bar{\alpha} = \tfrac{2\beta(1-c)}{L}. \tag{5.8}$$

Teorem 26. *Pretpostavimo da je ∇f Lipchitz neprekidna s konstantom L. Neka je $\{x_n\}$ niz generiran gradijentnom metodom s projekcijom. Tada je svako gomilište niza stacionarna točka.*

Dokaz. Budući da je $\{f(x_n)\}$ padajući niz, a funkcija f ograničena je odozdo na Ω, $f(x_n)$ ima limes $\lim_{k\to\infty} f(x_k) = f^\star$. Zbog ograde za duljinu koraka imamo

$$\|x_n - x_{n+1}\|^2 \leq \alpha(f(x_n) - f(x_{n+1}))/c \leq (f(x_n) - f(x_{n+1}))/c \to 0$$

kada $n \to \infty$.

Neka je $y \in \Omega$ i $n \geq 0$. Koristeći nejednakost (5.7) slijedi

$$\begin{aligned} \nabla f(x_n)^T(x_n - y) &= \nabla f(x_n)^T(x_{n+1} - y) + \nabla f(x_n)^T(x_n - x_{n+1}) \\ &\leq \alpha_n^{-1}(x_n - x_{n+1})^T(x_{n+1-y}) + \nabla f(x_n)^T(x_n - x_{n+1}). \end{aligned}$$

Ponovno zbog donje ograde za duljinu koraka (5.8) imamo

$$\nabla f(x_n)^T(x_n - y) \leq \|x_n - x_{n+1}\|(\alpha_n^{-1}\|x_{n+1} - y\| + \|\nabla f(x_n)\|), \tag{5.9}$$

$$\nabla f(x_n)^T(x_n - y) \leq \|x_n - x_{n+1}\|(\bar{\alpha}^{-1}\|x_{n+1} - y\| + \|\nabla f(x_n)\|). \tag{5.10}$$

Ako $x_{n_l} \to x^\star$ kao podniz $\{x_n\}$, tada možemo pogledati limese u gornjim nejednakostima (5.9-5.10) kada $l \to \infty$ i time dokazati teorem. □

Algoritam gradijentne metode s projekcijom ima važno svojstvo da ako konvergira prema nedegeneriranom lokalnom minimumu, tada je aktivni skup $\mathcal{A}^n$ od x_n isti kao i $\mathcal{A}(x^\star)$ nakon konačno mnogo iteracija.

Teorem 27. *Pretpostavimo da je f dva puta neprekidno diferencijabilna i da aproksimacije $\{x_n\}$ dobivene gradijentnom metodom s projekcijom konvergiraju prema nedegeneriranom lokalnom minimumu $x^\star$. Tada postoji n_0 tako da je $\mathcal{A}(x_n) = \mathcal{A}(x^\star)$ za sve $n \geq n_0$.*

Dokaz. Neka je $\bar{\alpha}$ donja ograda za duljinu koraka. Neka je δ takav da vrijede tvrdnje iz *leme* 12. za sve $\alpha = \bar{\alpha}$ (stoga i za sve $\alpha \geq \bar{\alpha}$). Neka je n_0 takav da je $\|e_n\| < \delta$ za sve $n \geq n_0 - 1$, tada vrijede tvrdnje teorema. □

Na kraju ovog poglavlja, budući da smo pripremili sve potrebne rezultate, dokazat ćemo *teorem* 22.:

Dokaz. Korolar 3. daje nejednakost

$$0 \leq \|x^\star - x^\star(\alpha)\|^2 \leq \alpha \nabla f(x^\star)^T (x^\star - x^\star(\alpha)).$$

Ako u definiciju stacionarnosti uvrstimo $x = x^\star(\alpha)$, imamo

$$\nabla f(x^\star)^T (x^\star - x^\star(\alpha)) \leq 0$$

i stoga $x^\star = x^\star(\alpha)$.

Kako bismo dokazali obrat, pretpostavimo da je $x^\star = x^\star(\alpha)$ za sve $\alpha \geq 0$. To povlači da $x^\star$ ostaje isti kada primijenimo gradijentnu metodu s projekcijom. Stoga je $x^\star$ i stacionarna točka.

□

Definirajmo funkciju

$$F(x) = x - \mathcal{P}(x - \nabla f(x)).$$

Iz definicije aproksimacije $x(\alpha)$ imamo

$$F(x) = x - x(1).$$

Ako u *teorem* 22. uvrstimo $\alpha = 1$, dobit ćemo sljedeći korolar.

Korolar 4. *Neka je f neprekidno diferencijabilna funkcija na Ω. Tada, ako je $x^\star$ stacionarna, onda je $F(x^\star) = 0$.*

5.3 Newtonova metoda s projekcijom

Prvo ćemo definirati $\mathcal{A}^\epsilon(x)$, ϵ-aktivan skup u x sa

$$\mathcal{A}^\epsilon(x) = \{i : U_i - (x)_i \leq \epsilon \quad \text{ili} \quad (x)_i - L_i \leq \epsilon\}.$$

Slično kao u prethodnom poglavlju, neka se komplement skupa označava sa $\mathcal{N}^\epsilon(x)$ i zvat ćemo ga ϵ-neaktivnim skupom.

Neka je dan ϵ takav da je $0 \leq \epsilon < \min\{(U_i - L_i)/2 : i = 1, 2, \dots, n\}$, x_c i Hessijan u točki x_c (označen s H_c). Tada možemo uvesti reducirani Hessijan $\mathcal{R}_c$ kao matricu definiranu s

$$\begin{aligned}\mathcal{R}_c &= \mathcal{P}_{\mathcal{A}^\epsilon(x_c)}(x_c) + \mathcal{P}_{\mathcal{N}^\epsilon(x_c)}(x_c) H_c \mathcal{P}_{\mathcal{N}^\epsilon(x_c)}(x_c) \\ &= \left\{ \begin{array}{ll} \delta_{ij}, & \text{ako je} \quad i \in \mathcal{A}^\epsilon(x_c) \quad \text{ili} \quad j \in \mathcal{A}^\epsilon(x_c) \\ (H_c)_{ij}, & \text{inače.} \end{array} \right. .\end{aligned}$$

Kada je važna eksplicitna ovisnost o x_c, $\epsilon = \epsilon_c$, H_c reducirani Hessijan označavat ćemo sa $\mathcal{R}(x_c, \epsilon_c, H_c)$.

Primijetimo da je $\nabla^2_R f(x_c) = \mathcal{R}(x_c, 0, \nabla^2 f(x_c))$.

Ako je početna iteracija x_0 dovoljno blizu nedegeneriranom lokalnom minimumu $x^\star$, tada će Newtonova metoda s projekcijom raditi s korakom duljine 1, odnosno $\alpha = 1$. Nadalje, ako se ϵ_n pažljivo odabere, tada će niz aproksimacija konvergirati q-kvadratično ka $x^\star$.

Jedna preporuka za ϵ u n-tom koraku (ϵ_n) je [18]

$$\epsilon_n = \min\{\|x_n - x_n(1)\|, (U_i - L_i)/2\}.$$

Ako imamo aproksimaciju x_c, ako je $\mathcal{R}(x, \epsilon, \nabla^2_R f(x))$ simetrična pozitivno definitna matrica, tada je nova aproksimacija dobivena Newtonovim algoritmom s projekcijom dana s

$$x(\alpha) = \mathcal{P}(x + \alpha p),$$

gdje se p dobija kao rješenje sustava

$$\mathcal{R}(x, \epsilon, \nabla^2_R f(x)) p = -\nabla f(x_c),$$

a α određuje duljinu koraka. Duljina koraka određuje se kao u slučaju kod gradijentne metode s projekcijom.

Može se pokazati da ako je $0 < \epsilon < \min\{(U_i - L_i)/2 : i = 1, 2, \dots, n\}$ i ako je $\mathcal{R}(x, \epsilon, \nabla^2_R f(x))$ simetrična pozitivno definitna matrica, tada je

$$f(x(\alpha)) < f(x),$$

za dovoljno mali α. Štoviše, u [18] je pokazano da je zadovoljen dovoljni uvjet

$$f(x(\alpha)) - f(x) \leq -C \nabla f(x)^T (x - x(\alpha)),$$

za dovoljno mali α.

U sljedećem algoritmu dan je formalan opis Newtonovog algoritma s projekcijom. Napomenimo da ako je početna aproksimacija x_0 daleko od $x^\star$ i reducirani Hessijan nije simetrična pozitivno definitna matrica, tada odabir duljine koraka (stoga i cjelokupni iterativni proces) može biti problem. Algoritam koristi činjenicu da je

$$\mathcal{R}(x, \epsilon, \nabla_R^2 f(x)) = \mathcal{R}(x, \epsilon, \nabla^2 f(x))$$

koja slijedi iz činjenice da je $\mathcal{A}(x) \subset \mathcal{A}^\epsilon(x)$.

Algoritam 6 . Newtonova metoda s projekcijom

Ulazni parametri: $x, f, nmax, \tau$.
for i:= 0, 1, ..., $nmax$ **do**
  (a) Izračunaj f i ∇f; provjeri zaustavni kriterij (5.6);
  (b) $\epsilon = \|x - x(1)\|$;
  (c) Izračunaj $\mathcal{R} = \mathcal{R}(x, \epsilon, \nabla_R^2 f(x))$. Ako $\mathcal{R}$ nije simetrična pozitivno definitna matrica, tada prekini optimizacijski proces;
  (d) Riješi sustav $\mathcal{R}p = -\nabla f(x_c)$;
  (e) Odredi najmanji prirodni broj m tako da vrijedi nejednakost (5.3) za $\alpha = \beta^m$;
  (f) $x = x(\alpha)$.
end for
if $n = nmax$ i zaustavni kriterij nije zadovoljen **then**
  Ispiši: "Algoritam nije odredio zadovoljavajuću aproksimaciju";
end if

Sljedećim teoremom pokazat ćemo da ako smo blizu minimuma, onda je brzina konvergencije kao kod standardne Newtonove metode.

Teorem 28. *Neka je $x^\star$ nedegenerirani lokalni minimum. Ako je x_0 dovoljno blizu $x^\star$ i $\mathcal{A}(x_0) = \mathcal{A}(x^\star)$, tada će Newtonova metoda s projekcijom (uz $\epsilon_n = \|x_n - x_n(1)\|$) konvergirati kvadratičnom brzinom.*

Dokaz. Naša je pretpostavka da je aktivni skup identificiran, odnosno

$$\mathcal{A}(x_c) = \mathcal{A}(x_+) = \mathcal{A}(x^\star),$$

što povlači da

$$\mathcal{P}_{\mathcal{A}(x_c)} e_c = \mathcal{P}_{\mathcal{A}(x_c)} e_+ = 0.$$

Stoga kako bismo dokazali rezultat moramo procijeniti $\mathcal{P}_{\mathcal{N}(x_c)} e_+$.
Neka je

$$\delta^\star = \min_{i \in \mathcal{N}(x^\star)} \{|(x)_i - U_i|, |(x)_i - L_i|\} > 0.$$

Pogrešku e uzimamo tako da je

$$\|e\| \leq \delta^\star / M,$$

gdje je M konstanta iz *teorema* 24. Tada *teorem* 24. možemo primijeniti kako bismo zaključili da je $\epsilon_c < \delta^\star$ i $\|e_c\| < \delta^\star$. Tada svaki indeks $i \in \mathcal{A}^{\epsilon_c}(x_c)$ također mora biti i u $\mathcal{A}(x_c) = \mathcal{A}(x^\star)$.
Stoga
$$\mathcal{A}^{\epsilon_c}(x_c) = \mathcal{A}(x_c) = \mathcal{A}(x^\star),$$
iz čega slijedi da je
$$\mathcal{R}(x_c, \epsilon_c, \nabla_R^2 f(x_c)) = \nabla_R^2 f(x_c).$$
Stoga za $\|e_c\|$ dovoljno mali, Newtonova metoda sa projekcijom daje
$$x_+ = \mathcal{P}(x_c - (\nabla_R^2 f(x_c))^{-1} \nabla f(x_c)).$$
Korištenjem fundamentalnog teorema matematičke analize imamo
$$\nabla f(x_c) = \nabla f(x^\star) + \nabla^2 f(x_c) e_c + E_1, \tag{5.11}$$
gdje je
$$E_1 = \int_0^1 (\nabla^2 f(x^\star + te_c) - \nabla^2 f(x_c)) e_c \, dt$$
i tada je $\|E_1\| \leq K_1 \|e_c\|^2$ za neku konstantu $K_1 > 0$.
Iz nužnih uvjeta imamo
$$\mathcal{P}_{\mathcal{N}(x)} \nabla f(x^\star) = \mathcal{P}_{\mathcal{N}(x^\star)} \nabla f(x^\star) = 0. \tag{5.12}$$
Iz činjenice da je $\mathcal{N}(x_c) = \mathcal{N}(x^\star)$ dobivamo ekvivalentne tvrdnje
$$e_c = \mathcal{P}_{\mathcal{N}(x_c)} e_c \quad \text{i} \quad \mathcal{P}_{\mathcal{A}(x_c)} e_c = 0. \tag{5.13}$$
Nadalje, uzimajući u obzir (5.11), (5.12) i (5.13) imamo
$$\begin{aligned}
\mathcal{P}_{\mathcal{N}(x_c)} \nabla f(x_c) =& \mathcal{P}_{\mathcal{N}(x_c)} \nabla^2 f(x_c) \mathcal{P}_{\mathcal{N}(x_c)} e_c + \mathcal{P}_{\mathcal{N}(x_c)} E_1 \\
=& \mathcal{P}_{\mathcal{A}(x_c)} e_c + \mathcal{P}_{\mathcal{N}(x_c)} \nabla^2 f(x_c) \mathcal{P}_{\mathcal{N}(x_c)} e_c + \mathcal{P}_{\mathcal{N}(x_c)} E_1 \\
=& \nabla_R^2 f(x_c) e_c + \mathcal{P}_{\mathcal{N}(x_c)} E_1.
\end{aligned}$$
Sada koristeći definiciju od ∇_R^2 dobivamo
$$\mathcal{P}_{\mathcal{N}(x_c)} (\nabla_R^2 f(x_c))^{-1} \nabla f(x_c) = (\nabla_R^2 f(x_c))^{-1} \mathcal{P}_{\mathcal{N}(x_c)} \nabla f(x_c) = e_c + E_2,$$
gdje je $\|E_2\| \leq K_2 \|e_c\|^2$ za neku konstantu $K_2 > 0$.

Budući da je $\mathcal{P}_{\mathcal{N}(x_c)} \mathcal{P} w = \mathcal{P} \mathcal{P}_{\mathcal{N}(x_c)} w$ za sve $w \in \mathbb{R}^n$
$$\begin{aligned}
\mathcal{P}_{\mathcal{N}(x_c)} x_+ =& \mathcal{P}_{\mathcal{N}(x_c)} \mathcal{P}(x_c - (\nabla_R^2 f(x_c))^{-1} \nabla f(x_c)) \\
=& \mathcal{P} \mathcal{P}_{\mathcal{N}(x_c)} (x_c - (\nabla_R^2 f(x_c))^{-1} \nabla f(x_c)) \\
=& \mathcal{P}(x^\star - E_2).
\end{aligned}$$
Stoga $e_+ \leq K_2 \|e_c\|^2$, što je i trebalo pokazati. $\square$

POGLAVLJE 6

Metode direktnog traženja

U ovom poglavlju proučavat ćemo metode direktnog traženja. To su metode koje kreću od vrijednosti funkcije f u točkama koje su na početku zadane nakon čega se iz tih podataka određuje sljedeća točka u cilju minimizacije funkcije. Obradit ćemo metode koje eksplicitno ne koriste aproksimaciju gradijenta.

6.1 Nelder-Meadov algoritam

Nelder-Meadov simpleks algoritam u pojedinom koraku algoritma formira simpleks S koji kao svoje vrhove sadrži aproksimacije oko optimalne točke. Simpleks, u N dimenzija, konveksna je ljuska $N+1$ točaka. Npr. u dvije dimenzije simpleks je trokut, a u tri dimenzije tetraedar.

Kod Nelder-Meadovog simpleks algoritma vrhovi $x_1, x_2, \ldots, x_{N+1}$ sortirani su u odnosu na vrijednosti funkcije f, odnosno

$$f(x_1) \leq f(x_2) \leq \cdots \leq f(x_{N+1}). \tag{6.1}$$

x_1 zovemo najboljim vrhom, a vrh x_{N+1} najlošijim vrhom. Ukoliko je vrijednost funkcije u nekim vrhovima jednaka vrijednosti funkcije u x_1, tada najbolji vrh nije jedinstveno određen.

U danom koraku algoritma cilj je zamijeniti najlošiji vrh x_{N+1} novim vrhom oblika

$$x(\mu) = (1+\mu)\overline{x} - \mu\, x_{N+1}, \tag{6.2}$$

pri čemu je $\overline{x}$ centroid konveksne ljuske podataka $(x)_1, \ldots, (x)_N$, odnosno

$$\overline{x} = \tfrac{1}{N}\sum_{i=1}^{N}(x)_i. \tag{6.3}$$

Vrijednost μ nalazi se u sljedećim intervalima

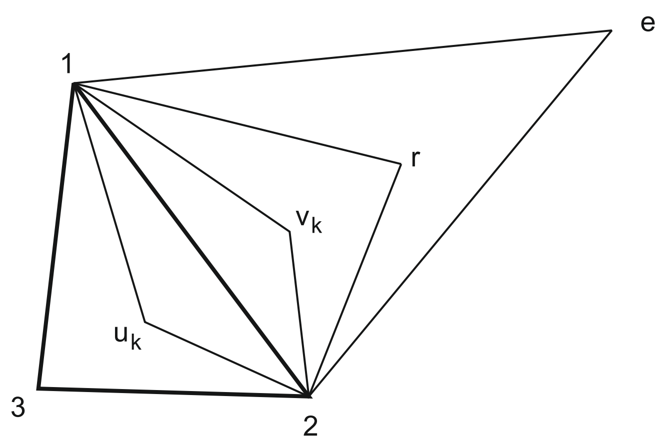
$$-1 < \mu_{uk} < 0 < \mu_{vk} < \mu_r < \mu_e.$$

Vrijednost μ određuje kakav će se korak algoritma izvršavati danom koraku algoritma. U sljedećem je algoritmu opisano o kojim se koracima radi. Ova formulacija algoritma dopušta kraj algoritma ako je $f(x_{N+1}) - f(x_1)$ dovoljno malo (manje od unaprijed zadane tolerancije $\epsilon > 0$) ili je broj pozivanja funkcije f prešao određeni, unaprijed zadani broj. Označimo ga sa k_{max}.

Jedan mogući odabir vrijednosti μ je

$$\{\mu_r, \mu_e, \mu_{vk}, \mu_{uk}\} = \{1, 2, 1/2, -1/2\}.$$

slika 6.1. pokazuje moguće transformacije opisane u navedenom algoritmu u slučaju dvije dimenzije ($N = 2$). Vrhovi početnog simpleksa numerirani su s 1, 2 i 3.



Slika 6.1. Novi vrhovi kod Nelder-Meadovog simpleks algoritma

Za Nelder-Meadov algoritam ne može se jamčiti konvergencija, čak ni za glatke funkcije. Međutim, performanse Nelder-Meadovog algoritma u praksi su vrlo dobre.

Slučaj kada Nelder-Meadov algoritam ne konvergira je slučaj kada dolazi do stagnacije u neoptimalnoj točki. Tada je potrebno pokrenuti algoritam na novom skupu točaka, pri čemu se od prijašnjih točaka zadržava jedino najbolja točka. Više o kriteriju prepoznavanja stagnacije i algoritmima koji opisuju novi skup točaka može se naći u [18].

Algoritam 7 . Algoritam Nelder-Mead

Ulaz: S-početni simpleks, f, $\epsilon > 0$, k_{max}
Izračunaj vrijednost funkcije f u vrhovima simpleksa S i sortiraj vrhove tako da vrijedi (6.1), $f_{br} = N + 1$
while $f(x_{N+1}) - f(x_1) > \epsilon$ **do**
Izračunaj $\overline{x}$ prema (6.3), $x(\mu_r)$ prema (6.2), $f_r = f(x(\mu_r))$, $f_{br} = f_{br} + 1$
(a) Refleksija:
if $f_{br} = k_{max}$ **then**; **exit**; **end if**
if $f(x_1) \leq f_r < f(x_N)$ **then**
zamijeni x_{N+1} sa $x(\mu_r)$ i idi na korak (f)
end if
(b) Ekspanzija:
if $f_{br} = k_{max}$ **then**; **exit**; **end if**
if $f_r < f(x_1)$ **then**
izračunaj $f_e = f(x(\mu_e))$, $f_{br} = f_{br} + 1$
end if
if $f_e < f_r$ **then**
zamijeni x_{N+1} sa $x(\mu_e)$; **else** zamijeni x_{N+1} sa $x(\mu_r)$ i idi na korak (f)
end if
(c) Vanjska kontrakcija:
if $f_{br} = k_{max}$ **then**; **exit**; **end if**
if $f(x_N) \leq f_r < f(x_{N+1})$ **then**
izračunaj $f_k = f(x(\mu_{vk}))$, $f_{br} = f_{br} + 1$
end if
if $f_k \leq f_r$ **then**
zamijeni x_{N+1} sa $x(\mu_{vk})$ i idi na korak (f); **else** idi na korak (e)
end if
(d) Unutarnja kontrakcija:
if $f_{br} = k_{max}$ **then**; **exit**; **end if**
if $f_r \geq f(x_{N+1})$ **then**
izračunaj $f_k = f(x(\mu_{uk}))$, $f_{br} = f_{br} + 1$
end if
if $f_k < f(x_{N+1})$ **then**
zamijeni x_{N+1} sa $x(\mu_{uk})$ i idi na korak (f); **else** idi na korak (e)
end if
(e) Stezanje:
if $f_{br} \geq k_{max} - N$ **then**, **exit**, **end if**
for $i = 2, \ldots, N+1$ **do**
$x_i = x_1 - (x_i - x_1)/2$, izračunaj $f(x_i)$,$f_{br} = f_{br} + 1$
end for
(f) Sortiraj vrhove od S tako da vrijedi (6.1)
end while

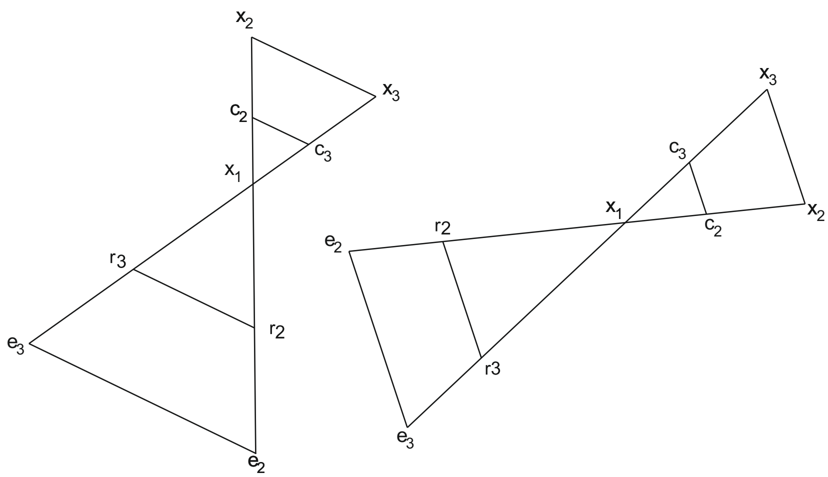
6.2 MDS algoritam

U cilju poboljšanja Nelder-Meadovog algoritma u MDS algoritmu (eng. multi-directional search) zahtijeva se da je novi simpleks kongruentan polaznom simpleksu.

Slika 6.2. pokazuje u slučaju dimenzije 2 dva tipa simpleksa koji se pojavljuju u ovom algoritmu. Počevši s uređenim simpleksom S^c s vrhovima x_1, x_2, x_3 ako primijenimo korak refleksije dobit ćemo simpleks S^r s vrhovima x_1, r_2 i r_3. Ako je najbolja vrijednost funkcije u vrhovima simpleksa S^r bolja od najbolje vrijednosti $f(x_1)$ u S^0, tada se S^r privremeno prihvaća i nastavlja se s korakom ekspanzije. *slika* 6.2. pokazuje i korake ekspanzije te kontrakcije koji se koriste u algoritmu.

Korak ekspanzije u ovom slučaju razlikuje se od ekspanzije u Nelder-Meadovom algoritmu jer se u MDS algoritmu dobija N novih točaka, dobija se veći simpleks koji je sličan polaznom. Simpleks dobiven ekspanzijom S^e, na *slici* 6.2., ima vrhove x_1, e_2, e_3 i prihvaća se umjesto simpleksa S^r ako je najbolja vrijednost funkcije u vrhovima S^e bolja od najbolje vrijednosti funkcije u vrhovima S^r.

U slučaju da najbolja vrijednost funkcije u vrhovima S^r nije bolja od najbolje vrijednosti funkcije u vrhovima S^c, tada se na simpleks primjenjuje korak kontrakcije, a vrhove novog simpleksa označit ćemo sa x_1, c_2, c_3. Nakon što je novi simpleks generiran, vrhovi se numeriraju tako da je simpleks uređen. Kod novog simpleksa vrhovi se ponovo uređuju, što nam daje novi simpleks S^+.



Slika 6.2. Novi vrhovi kod MDS algoritma

Ako su vrhovi simpleksa S^c poznati, onda broj dodatnih evaluacija koji je potreban da se izračuna S^+ iznosi $2N$.

Slično kao kod Nelder-Meadovog algoritma imamo parametre ekspanzije i kontrakcije (μ_e i μ_k). Tipične vrijednosti tih parametara su 2 i 1/2.

Algoritam 8 Algoritam MDS

Ulaz: S-početni simpleks, f, $\epsilon > 0$, k_{max}
  Izračunaj vrijednost funkcije f u vrhovima simpleksa S i sortiraj vrhove tako da vrijedi (6.1), $f_{br} = N + 1$
  while $f(x_{N+1}) - f(x_1) > \epsilon$ **do**
    (a) Refleksija:
    if $f_{br} = k_{max}$ **then**
      exit
    end if
    for $j = 2, \ldots, N + 1$ **do**
      $r_j = x_1 - (x_j - x_1)$, izračunaj $f(r_j)$, $f_{br} = f_{br} + 1$
      if $f(x_1) > \min_j\{f(r_j)\}$ **then**
        idi na korak (b)
      else
        idi na korak (c).
      end if
    end for
    (b) Ekspanzija:
    for $j = 2, \ldots, N + 1$ **do**
      $e_j = x_1 - \mu_e(x_j - x_1)$, izračunaj $f(e_j)$, $f_{br} = f_{br} + 1$
    end for
    if $\min_j\{f(r_j)\} > \min_j\{f(e_j)\}$ **then**
      for $j = 2, \ldots, N + 1$ **do**
        $x_j = e_j$
      end for
    else
      for $j = 2, \ldots, N + 1$ **do**
        $x_j = r_j$
      end for
    end if
    idi na korak (d)
    (c) Kontrakcija:
    for $j = 2, \ldots, N + 1$ **do**
      $x_j = x_1 + \mu_k(x_j - x_1)$, izračunaj $f(x_j)$, $f_{br} = f_{br} + 1$
    end for
    Sortiraj vrhove tako da vrijedi (6.1)
  end while

6.3 Powellova metoda u više dimenzija

Razne metode razvijene su za slučaj jednodimenzionalne minimizacije, stoga se u slučaju višedimenzionalne minimizacije mogu generirati metode koje rješavaju problem višedimenzionalne minimizacije koristeći niz jednodimenzionalnih minimizacija. Različite metode razlikuju se u činjenici kako u svakom sljedećem koraku biraju smjer n u kojem traže minimum. Sve metode toga tipa podrazumijevaju postojanje "rutine" koja rješava problem jednodimenzionalne minimizacije.

Algoritam 9 Algoritam jednodimenzionalne minimizacije

Ulaz: vektori P i n, funkcija f
  Odredi skalar λ koji minimizira $f(P + \lambda n)$
  Zamijeni P sa $P + \lambda n$
  Zamijeni n sa λn

U ovom poglavlju proučavat ćemo metode koje biraju uzastopne smjerove tako da eksplicitno ne računaju gradijent funkcije. Gore navedeni algoritam jednodimenzionalne minimizacije može, ali ne mora, koristiti gradijent. Međutim, ako se gradijent koristi već u algoritmu jednodimenzionalne minimizacije bilo bi poželjno da se koristi i u odabiru smjerova.

Jednostavan primjer takvog odabira smjerova su jedinični kanonski vektori: $e_1, e_2, \ldots, e_N$. U tom će slučaju metoda, koristeći algoritam jednodimenzionalne minimizacije, prvo odrediti minimum u smjeru prvog smjera, tada se iz dobivene točke traži minimum u drugom smjeru, itd., kružeći po cijelom skupu smjerova sve dok se funkcija ne prestane smanjivati. Za velik broj funkcija ova je metoda dosta dobra, ali za određene funkcije ova metoda nije efikasna. U cilju dobivanja što bolje metode cilj je odrediti smjerove koji će biti bolji od smjerova e_i.

Metode koje se temelje na skupu smjerova sadrže pravilo kako kreiraju sljedeći smjer i njihov skup smjerova ili (i) sadrže pojedine smjerove koji vode minimumu funkcije ili (ii) sadrže određeni broj smjerova koji se ne "podudaraju" s dodatnim svojstvom da se može izbjeći beskonačno kruženje kroz smjerove.

U Powellovoj kvadratično konvergentnoj metodi kreira se N međusobno konjugiranih smjerova. Metoda je opisana sljedećim "osnovnim postupkom":

Algoritam 10 Powellov algoritam minimizacije

Ulaz: početni smjerovi $u_i = e_i$, $i = 1, \ldots, N$
Početnu poziciju označi sa P_0
for $i = 1, \ldots, N$ **do**
pomakni P_{i-1} u minimum po smjeru u_i i dobivenu točku označi sa P_i;
end for
for $i = 1, \ldots, N-1$ **do**
$u_i = u_{i+1}$;
$u_N = P_N - P_0$
pomakni P_N u minimum po smjeru u_N i dobivenu točku označi sa P_0
end for

U navedenom algoritmu dolazi do izbacivanja vektora u_1 u cilju dodavanja vektora $P_N - P_0$ a time se dolazi do smjerova koji postaju linearno zavisni što je problem u navedenom algoritmu. Kada se to dogodi, algoritam će odrediti minimum funkcije f na nekom potprostoru N dimenzionalnog prostora što može rezultirati netočnim minimumom. Stoga se gore navedeni algoritam mora popraviti.

U cilju izbjegavanja linearne zavisnosti kod Powellovog algoritma mogu se koristiti sljedeće metode:

i) Može se napraviti ponovna inicijalizacija smjerova u_i tako da im se nakon svakih N ili $N+1$ iteracija osnovne procedure pridruže vektori e_i.

ii) Skup smjerova može se izjednačiti sa stupcima određene ortogonalne matrice. Umjesto da se odbace postojeće informacije vezane uz konjugirane smjerove, u ovom slučaju kreira se matrica čiji stupci nakon ortogonalizacije čine nove smjerove.

iii) Ovaj pristup nema svojstvo kvadratične konvergencije, ali cilj je pokušati pronaći nekoliko dobrih smjerova koji vode u minimum umjesto N nužno konjugiranih smjerova. Osnovna ideja je modificirati Powellovu metodu s tim da i dalje uzimamo novi smjer $P_N - P_0$. Promjena je u tome da odbacujemo stari smjer u smjeru kojeg je funkcija imala najveći pad. Ta činjenica čini se paradoksalnom, budući da je taj smjer bio najbolji u prethodnoj iteraciji. Međutim, to je vjerojatno glavna komponenta u novom smjeru. Stoga, ako zaboravimo navedeni smjer, izbjeći ćemo linearnu zavisnost. Postoje neke iznimke vezane uz tu osnovnu metodu, a više o tome može se pronaći u [23].

Literatura

[1] M. Alić, G. Nogo, *Optimizacija: Uvod u teoriju nužnih i dovoljnih uvjeta ekstrema*, Odjel za matematiku, Sveučilište u Osijeku, Osijek, 2004.

[2] M. Avriel, *Nonlinear Programming: Analysis and Methods 2ed.*, Dover Publications, Inc. Mineola, New York, 2003.

[3] M. S. Bazaraa, H. D. Sherali, C. M. Shetty, *Nonlinear Programming. Theory and Algorithms. 3^{rd} Edition*, Wiley, New Jersey, 2006.

[4] J. F. Bonnans, J.C. Gilbert, C. Lemaréchal, C.A. Sagastizábal, *Numerical Optimization. Theoretical and Practical Aspects*, Springer-Verlag, Berlin, 2006.

[5] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, 2004.

[6] R. P. Brent, *Algorithms for Minimization without Derivatives*, Dover Publications, New York, 1973.

[7] C. G. Broyden, *A class of methods for solving nonlinear simultaneous equation*, Math. Comput., 19(1965), 577-593

[8] W. C. Davidon, *Variable metric method for minimization*, Rep. ANL-5990 Rev, Argonne National Laboratories, Argonne, Ill., 1959.

[9] V. F. Demjanov, F. P. Vasilev, *Nedifferenciruemaja optimizacija*, Nauka, Moskva, 1981.

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

[11] J. E. Dennis Jr., J. J. More, *Quasi-Newton methods, motivation and theory*, SIAM Review, 19(1977), 46-89

[12] Douglas N. Arnold, *A Concise Introduction to Numerical Analysis*, School of Mathematics, University of Minnesota, Minneapolis, MN 55455, http://umn.edu/ arnold/

[13] P. E. Gill, W. Murray, M. H. Wright, *Practical Optimization*, Academic Press, 1981.

[14] J. B. Hiriart-Urruty, C. Lemaréchal, *Fundamentals of Convex Analysis*, Springer-Verlag, Berlin - Heidelberg, 2001.

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

[16] D. Jukić, R. Scitovski, *Metode optimizacije – predavanja*, Tuzla, 2005.

[17] D. Kincaid, W. Cheney, *Numerical Analysis*, Brooks/Cole Publishing company, New York, 1996.

[18] C. T. Kelley, *Iterative Methods for Optimization*, SIAM, Philadelphia, 1999.

[19] P. Lancaster, K. Šalkauskas, *Curve and Surface Fitting: An Introduction*, Academic Press, London, 1986.

[20] J. M. Ortega, W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970. (postoji i ruski prijevod)

[21] R. Plato, *Concise Numerical Mathematics*, Providence: American Mathematical Society, 2003.

[22] B. N. Psheniynii, Yu. M. Danilin, *Chislennye metody v ekstremaljnyh zadachah*, Nauka, Moskva, 1975.

[23] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, Cambridge, 1989.

[24] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, New York, 1953.

[25] A. Ruszczynski, *N*onlinear Optimization, Princeton University Press, Princeton and Oxford, 2006.

[26] R. Scitovski, *Numerička matematika*, Odjel za matematiku, Sveučilište u Osijeku, Osijek, 2004.

[27] R. Scitovski, Š. Ungar, D. Jukić, *Approximating surfaces by moving total least squares method*, Applied Mathematics and Computation 93 (1998.) 219-232.

[28] J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, New York, 1993.

[29] F. P. Vasiljev, *Lekcii po metodam resheniya ekstremalnyh zadach*, Moskovski univerzitet, Moskva, 1974.

[30] S. Wolfram, *The Mathematica Book*, Wolfram Media, Champaign, 2011.

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

Metode optimizacije

Namjena ovog teksta je upoznati studente s glavnim metodama jednodimenzionalne i višedimenzionalne minimizacije sa i bez ograničnja. Posebna je pozornost posvećena metodama minimizacije nediferencijabilnih funkcija. Pri tome je izbjegavano dokazivanje zahtjevnih teorema, osim u slučajevima konstruktivnih dokaza koji sami po sebi upućuju na izgradnju ideja ili metoda.

Navedeni optimizacijski problemi imaju veliku primjenu u raznim dijelovima života. Na primjer, često se javljaju problemi poput optimalnog oblikovanja određenih mehaničkih sustava (oblikovanje dijelova automobilskih motora, nosivih konstrukcija u građevinarstvu, ...), problem modeliranja ponašanja tržišta, problemi iz teorije upravljanja (smirivanje sustava, optimalno upravljanje, ...) i mnogi drugi.

Upravo činjenica da se razni problemi optimizicije pojavljuju u raznim dijelovima ljudske djelatnosti osigurava ovom tekstu široku primjenu.