Subdiferencijal svojstvene vrijednosti simetrične matrice

Markek, Goran

Master's thesis / Diplomski rad

2015

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:508738

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Goran Markek

SUBDIFERENCIJAL SVOJSTVENE VRIJEDNOSTI SIMETRIčNE MATRICE

Diplomski rad

Voditelj rada:
Prof. dr. sc. Marko Vrdoljak

Zagreb, veljača 2015.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ____________________________, predsjednik
2. ____________________________, član
3. ____________________________, član

Povjerenstvo je rad ocijenilo ocjenom ________________.

Potpisi članova povjerenstva:

1. ____________________________
2. ____________________________
3. ____________________________

Sadržaj

Sadržaj **iii**

Uvod **1**

1 Konveksna analiza **2**
1.1 Derivacija u smjeru i subdiferencijal konveksne funkcije 2
1.2 Neka svojstva subdiferencijala 9
1.3 Geometrijska interpretacija subdiferencijala 11

2 Neglatka analiza **14**
2.1 Clarkeova generalizacija derivacije u smjeru 14
2.2 Subdiferencijalni račun 19

3 Svojstvene vrijednosti simetrične matrice **26**
3.1 Derivacija najveće svojstvene vrijednosti 26
3.2 Michel-Penotova derivacija u smjeru 29
3.3 Clarkeov i Michel-Penotov subdiferencijal svojstvene vrijednosti simetrične matrice 36

Bibliografija **42**

Uvod

U teoriji optimizacije vrlo često se javljaju problemi koji zahtijevaju rad s neglatkim funkcijama, bilo zbog prirode problema, ili zbog strukture prikupljenih podataka. U nekim slučajevima, metode koje ne koriste gradijent funkcije bit će dovoljne za pronalazak rješenja, ali ponekad će greška biti prevelika i javlja se potreba za modifikacijom zadanih funkcija.
U ovom radu cilj je poopćiti dobro poznate pojmove i rezultate iz klasične analize, kao što su na primjer pojam diferencijabilnosti, derivacije u smjeru i gradijenta realne funkcije. Želimo aproksimirati funkcije u točkama u kojima nisu diferencijabilne, ali tako da zadržimo svojstvo konveksnosti aproksimacije. Zanimljiv primjer je funkcija d_C koja mjeri udaljenost točke $x \in \mathbb{R}^n$ od nepraznog i zatvorenog skupa $C \subset \mathbb{R}^n$ definirana sa:

$$d_C(x) = \min\{|x - c| : c \in C\},$$

gdje je $|\cdot|$ euklidska norma. Clarke u svojoj knjizi (vidi [2], točka 2.5) pokazuje da ako je C konveksan skup, funkcija d_C nije diferencijabilna u nekim točkama svoje domene. Srodan, ali mnogo jednostavniji primjer je funkcija apsolutne vrijednosti koju spominjemo u ovom radu. Kao prvi korak promatramo konveksne funkcije u okviru konveksne analize. Također definiramo subdiferencijal konveksne funkcije i proučavamo njegova svojstva.
Nakon toga, u drugom poglavlju prelazimo na lokalno Lipschitzove, ne nužno konveksne funkcije i obična derivacija u smjeru tada gubi svojstvo konveksnosti. Zbog toga uvodimo Clarkeovu derivaciju u smjeru i poopćujemo diferencijalni račun i vrlo važan teorem srednje vrijednosti.
Zadnje poglavlje započinjemo primjenom nekih teorema iz prva dva poglavlja na funkciju najveće svojstvene vrijednosti simetrične matrice. Zatim definiramo još jednu generaliziranu derivaciju u smjeru, Michel-Penotovu derivaciju u smjeru. Do kraja rada, cilj je dokazati jednakost Clarkeovog i Michel-Penotovog subdiferencijala m-te najveće svojstvene vrijednosti simetrične matrice.

Poglavlje 1

Konveksna analiza

1.1 Derivacija u smjeru i subdiferencijal konveksne funkcije

Svatko tko se bavi teorijom optimizacije zna da je svojstvo konveksnosti poželjno i prilično pojednostavljuje problem. Geometrijski gledano, svojstvo diferencijabilnosti je mogućnost linearne aproksimacije grafa funkcije (pomoću hiperravnine). Za konveksne funkcije hiperravnine će uvijek biti ispod grafa funkcije. Upravo na tome će se i bazirati definicija subdiferencijala za konveksne funkcije. Konveksne funkcije se najčešće definiraju na konveksnim skupovima, ali budući da se uvijek mogu proširiti sa $+\infty$ na cijeli $\mathbb{R}^n$ za domenu ćemo uzeti $\mathbb{R}^n$.

Definicija 1.1.1 (Lokalno Lipschitzova funkcija)**.** *Funkcija* $f : \mathbb{R}^n \to \mathbb{R}$ *je* lokalno Lipschitzova s konstantom K u $x \in \mathbb{R}^n$ *ako postoji* $\epsilon > 0$ *takav da za sve* $y, z \in B(x, \epsilon)$ *vrijedi*

$$|f(y) - f(z)| \leq K\|y - z\|. \tag{1.1}$$

Definicija 1.1.2 (Konveksna funkcija)**.** *Za funkciju* $f : \mathbb{R}^n \to \mathbb{R}$ *kažemo da je* konveksna *ako za sve* $x, y \in \mathbb{R}^n$ *i* $\lambda \in [0, 1]$ *vrijedi*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Teorem 1.1.3. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ konveksna funkcija. Tada je funkcija f lokalno Lipschitzova u x, za svaki $x \in \mathbb{R}^n$.*

Dokaz. Vidi [9] str. 8. □

Sljedeći bitan pojam koji želimo proučiti je derivacija funkcije u smjeru.

Derivacija u smjeru

Definicija 1.1.4 (Derivacija u smjeru)**.** Derivacija funkcije f u smjeru $v \in \mathbb{R}^n$, u točki $x \in \mathbb{R}^n$ *dana je sa*

$$f'(x, v) = \lim_{t \to 0+} \frac{f(x + tv) - f(x)}{t}.$$

Ako se prisjetimo glatke analize, znamo da ako je funkcija f diferencijabilna u točki $x \in \mathbb{R}^n$, tada derivacija u smjeru, u točki x, postoji za svaki $v \in \mathbb{R}^n$ i derivacija u smjeru je tada linearno preslikavanje

$$f'(x, v) = \nabla f(x) \cdot v,$$

gdje je vektor $\nabla f(x)$ gradijent funkcije f u točki x. U slučaju da f nije diferencijabilna, tada naravno, nije niti $f'(x, \cdot)$ linearna, ali može se pokazati da je sublinearna (pod uvjetom da je f konveksna).

Definicija 1.1.5 (Sublinearna funkcija)**.** *Funkcija $\sigma : \mathbb{R}^n \to \mathbb{R}$ je* sublinearna *ako su zadovoljena sljedeća dva uvjeta*

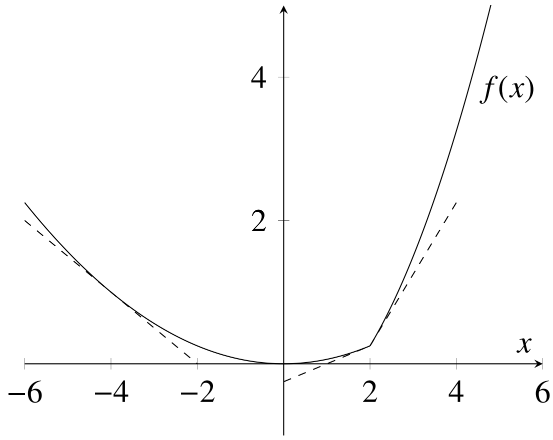
$$\sigma(x + y) \leq \sigma(x) + \sigma(y), \;\; x, y \in \mathbb{R}^n \qquad \textit{(subaditivnost),}$$

$$\sigma(tx) = t\sigma(x), \;\; x \in \mathbb{R}^n, t > 0 \qquad \textit{(pozitivna homogenost).}$$

U klasičnoj analizi, najjednostavnije, netrivijalne funkcije su linearne funkcije, dok u konveksnoj analizi to mjesto zauzimaju sublinearne funkcije. Iz gornje definicije lako se vidi da vrijedi

$$\sigma(ax + by) \leq a\sigma(x) + b\sigma(y), \;\; x, y \in \mathbb{R}^n, a, b \geq 0, \tag{1.2}$$

pa prema tome sublinearne funkcije smatramo generalizacijom linearnih funkcija. Na slici 1.1 vidimo primjer linearne i sublinarne aproksimacije. Nadalje, sublinearne funkcije su i konveksne budući da u 1.2 možemo odabrati $a, b \in [0, 1]$ tako da vrijedi $a + b = 1$. Geometrijski gledano, ako je $f : \mathbb{R}^n \to \mathbb{R}$ diferencijalna u x onda u toj točki postoji jedinstvena potporna hiperravnina određena gradijentom funkcije f u x, pa kažemo da je funkcija tangencijalno linearna. U suprotnom, ako je funkcija samo konveksna, umjesto



Slika 1.1: Primjer konveksne funkcije koja nije diferenicjabilna u $x = 2$.

hiperravnine imamo konus i kažemo da je funkcija f tangencijalno sublinearna.
Još jedno korisno svojstvo sublinearnih funkcija je korespondencija sa nepraznim zatvorenim konveksnim skupovima. Naime, ako za konveksni skup S definiramo funkciju $\sigma_S : \mathbb{R}^n \to \mathbb{R}$ na sljedeći način

$$\sigma_S(x) = \sup\{s \cdot x : s \in S\}, \tag{1.3}$$

tada će to biti sublinearna funkcija i zove se potporna funkcija skupa S u točki x. Takve funkcije se promatraju u teoremima separacije, kao na primjer u geometrijskoj verziji Hahn-Banachovog teorema. Ovdje navodimo analitičku verziju Hahn-Banachovog teorema.

Teorem 1.1.6 (Hahn-Banach). *Neka je X linearni vektorski prostor i $Y \subset X$ linearni potprostor. Ako je $p : X \to \mathbb{R}$ sublinearan funkcional i $f : Y \to \mathbb{R}$ linearni funkcional takav da je $f(y) \leq p(y)$, za svaki $y \in Y$ onda postoji linearni funkcional $F : X \to \mathbb{R}$ takav da vrijedi*

$$F(y) = f(y),\ y \in Y \qquad i \qquad F(x) \leq p(x),\ x \in X.$$

U sljedećem teoremu sažeta su najvažnija svojstva derivacije u smjeru za konveksne funkcije.

Teorem 1.1.7. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ konveksna funkcija s Lipschitzovom konstantom K u točki $x \in \mathbb{R}^n$. Tada*
(i) postoji derivacija od f u svakom smjeru $v \in \mathbb{R}^n$ i vrijedi

$$f'(x, v) = \inf_{t>0} \frac{f(x + tv) - f(x)}{t},$$

(ii) funkcija $v \mapsto f'(x, v)$ je pozitivno homogena i subaditivna na $\mathbb{R}^n$ i vrijedi

$$f'(x, v) \leq K\|v\|,$$

(iii) Funkcija $f'(x, v)$ je odozgo poluneprekidna kao funkcija od (x, v) i Lipschitzova s konstantom K kao funkcija od v u $\mathbb{R}^n$,
(iv) $-f'(x, -v) \leq f'(x, v)$.

Dokaz. (i) Uzmimo proizvoljan $v \in \mathbb{R}^n$ i definiramo funkciju $\phi : \mathbb{R} \to \mathbb{R}$

$$\phi(t) := \frac{f(x + tv) - f(x)}{t}.$$

Pokažimo prvo da $\phi(t)$ opada kada $t \downarrow 0$. Neka su $t_1, t_2, \epsilon > 0$ tako da vrijedi $0 < t_1 < t_2$. Zbog konveksnosti od f imamo

$$\begin{aligned}
\phi(t_2) - \phi(t_1) &= \frac{t_1 f(x + t_2 v) - t_2 f(x + t_1 v) + (t_2 - t_1) f(x)}{t_1 t_2} \\
&= \frac{(\frac{t_1}{t_2} f(x + t_2 v) + (1 - \frac{t_1}{t_2}) f(x)) - f(\frac{t_1}{t_2}(x + t_2 v) + (1 - \frac{t_1}{t_2}) x)}{t_1} \\
&\geq 0.
\end{aligned}$$

Neka je $0 < t < \epsilon$, sada vrijedi

$$\begin{aligned}
\phi(t) - \phi(-\frac{\epsilon}{2}) &= \frac{\frac{1}{2} f(x + tv) + \frac{1}{2} f(x) + \frac{t}{\epsilon} f(x - \frac{\epsilon}{2} v) + (1 - \frac{t}{\epsilon}) f(x) - 2f(x)}{\frac{t}{2}} \\
&\geq \frac{\frac{1}{2} f(x + \frac{t}{2} v) + \frac{1}{2} f(x - \frac{t}{2} v) - f(x)}{\frac{t}{4}} \\
&\geq \frac{f(x) - f(x)}{\frac{t}{4}} = 0,
\end{aligned}$$

što znači da je funkcija ϕ ograničena odozdo za $0 < t < \epsilon$, a budući da opada za $t \downarrow 0$ limes $\lim_{t\downarrow 0} \phi(t)$ postoji i vrijedi

$$f'(x, v) = \inf_{t>0} \frac{f(x + tv) - f(x)}{t}.$$

(ii) Zbog Lipschitzovog svojstva funkcije f vrijedi

$$\begin{aligned} |f'(x, v)| &\leq \lim_{t\to 0+} \frac{|f(x + tv) - f(x)|}{t} \\ &\leq \lim_{t\to 0+} \frac{K\|x + tv - x\|}{t} \\ &\leq K\|v\|. \end{aligned}$$

Za $\lambda > 0$ vrijedi

$$\begin{aligned} f'(x, \lambda v) &= \lim_{t\to 0+} \frac{f(x + t\lambda v) - f(x)}{t} \\ &= \lim_{t\to 0+} \lambda \frac{f(x + t\lambda v) - f(x)}{t\lambda} \\ &= \lambda \lim_{t\to 0+} \frac{f(x + t\lambda v) - f(x)}{t\lambda} = \lambda f'(x, v), \end{aligned}$$

odnosno, funkcija $v \mapsto f'(x, v)$ je pozitivno homogena. Za dokaz subaditivnosti uzmimo proizvoljne $v, w \in \mathbb{R}^n$, zbog konveksnosti funkcije f vrijedi

$$\begin{aligned} f'(x, v + w) &= \lim_{t\to 0+} \frac{f(x + t(v + w)) - f(x)}{t} \\ &= \lim_{t\to 0+} \frac{f(\frac{1}{2}(x + 2tv) + \frac{1}{2}(x + 2tw)) - f(x)}{t} \\ &\leq \lim_{t\to 0+} \frac{f(x + 2tv) - f(x)}{2t} + \lim_{t\to 0+} \frac{f(x + 2tw) - f(x)}{2t} \\ &= f'(x, v) + f'(x, w), \end{aligned}$$

Dakle, $f'(x, \cdot)$ je ograničena i sublinearna funkcija.
(iii) Neka su $(x_i), (v_i) \subset \mathbb{R}^n$ nizovi takvi da $x_i \to x$ i $v_i \to v$ i definiramo niz $(t_i) \subset \mathbb{R}^n$ na sljedeći način

$$t_i := K\|x_i - x\|^{\frac{1}{2}} + \frac{1}{i}.$$

Uočimo da $t_i \to 0$ i $t_i > 0$. Sada računamo

$$\begin{aligned} f'(x_i, v_i) &= \inf_{t>0} \frac{f(x_i + tv_i) - f(x_i)}{t} \\ &\leq \frac{f(x_i + t_i v_i) - f(x_i)}{t_i} \\ &= \frac{f(x + t_i v) - f(x)}{t_i} + \frac{f(x_i + t_i v_i) - f(x + t_i v)}{t_i} + \frac{f(x) - f(x_i)}{t_i} \end{aligned}$$

Za $i \to \infty$, za srednji član, zbog Lipschitzovosti, vrijedi

$$\begin{aligned} \frac{|f(x_i + t_i v_i) - f(x + t_i v)|}{t_i} &\leq \frac{K\|x_i - x\| + K\|t_i v_i - t_i v\|}{t_i} \\ &\leq \|x_i - x\|^{\frac{1}{2}} + K\|v_i - v\| \to 0. \end{aligned}$$

Za zadnji član vrijedi

$$\frac{|f(x) - f(x_i)|}{t_i} \leq \frac{K\|x - x_i\|}{t_i} \leq \|x - x_i\|^{\frac{1}{2}} \to 0.$$

Uzmimo sada limes superior za $i \to \infty$

$$\limsup_{i\to\infty} f'(x_i, v_i) \leq \limsup_{i\to\infty} \frac{f(x + t_i v) - f(x)}{t_i} = f'(x, v),$$

i time je poluneprekidnost odozgo dokazana.
Za dokaz Lipschitzovosti funkcije $f'(x, \cdot)$ uzmimo $v, w \in \mathbb{R}^n$ za koje vrijedi $x + tv, x + tw \in B(x, \epsilon)$ i sada imamo

$$f(x + tv) - f(x + tw) \leq Kt\|v - w\|,$$

i uzimanjem limesa vrijedi

$$\lim_{t\to 0+} \frac{f(x + tv) - f(x)}{t} \leq \lim_{t\to 0+} \frac{f(x + tw) - f(x)}{t} + K\|v - w\|,$$

odnosno

$$f'(x, v) - f'(x, w) \leq K\|v - w\|.$$

Zamjenom uloga v i w imamo

$$f'(x, w) - f'(x, v) \leq K\|v - w\|,$$

iz čega slijedi

$$|f'(x,v) - f'(x,w)| \leq K\|v - w\|.$$

(iv) Zbog subaditivnosti i pozitivne homogenosti imamo

$$0 = f'(x,0) = f'(x, \frac{1}{2}v - \frac{1}{2}v) \leq \frac{1}{2}f'(x,v) + \frac{1}{2}f'(x,-v),$$

pa stoga

$$-f'(x,-v) \leq f'(x,v).$$

□

Subdiferencijal konveksne funkcije

Kao što smo već spomenuli u uvodu ovog poglavlja, subdiferencijal konveksne funkcije zasniva se na činjenici da se sve linearne aproksimacije konveksne funkcije nalaze ispod grafa funkcije. Upravo ta činjenica je motivacija za sljedeću definiciju:

Definicija 1.1.8 (Subdiferencijal konveksne funkcije)**.** *Za konveksnu funkciju $f : \mathbb{R}^n \to \mathbb{R}$ i $x \in \mathbb{R}^n$, skup*

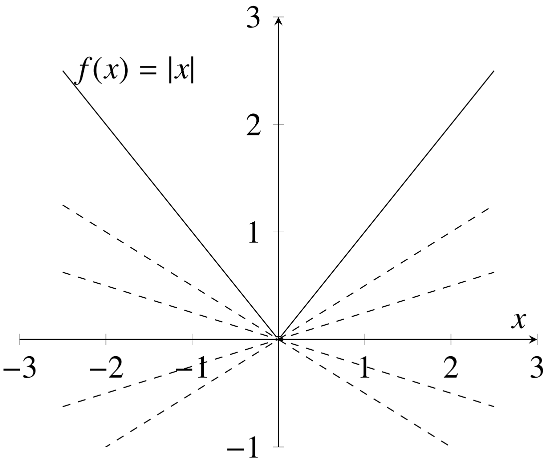
$$\partial_c f(x) = \{\xi \in \mathbb{R}^n : f(y) \geq f(x) + \xi \cdot (y - x),\ y \in \mathbb{R}^n\} \tag{1.4}$$

nazivamo subdiferencijalom funkcije f u točki x*, a elemente tog skupa* subgradijentima *funkcije f u točki x.*

Primjer 1.1.9. *($f(x) = |x|$) Apsolutna vrijednost $|\cdot| : \mathbb{R} \to \mathbb{R}$ je konveksna funkcija koja je derivabilna svuda osim u nuli. Izračunajmo njen subdiferencijal u nuli:*

$$\begin{aligned}\partial_c f(0) &= \{\xi \in \mathbb{R} : f(y) \geq f(0) + \xi(y - 0),\ ly \in \mathbb{R}\}\\ &= \{\xi \in \mathbb{R} : |y| \geq \xi y,\ y \in \mathbb{R}\}\\ &= [-1, 1],\end{aligned}$$

što je upravo skup vrijednosti nagiba svih pravaca koji prolaze kroz ishodište i nalaze se ispod grafa funkcije apsolutne vrijednosti, kao što vidimo na Slici 1.2.



Slika 1.2: Graf funkcije apsolutne vrijednosti.

1.2 Neka svojstva subdiferencijala

Sljedeći teorem nam daje poveznicu između obične derivacije u smjeru i subdiferencijala konveksne funkcije.

Teorem 1.2.1. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ konveksna funkcija. Tada za svaki $x \in \mathbb{R}^n$ vrijedi*
(i) $f'(x, v) = \max\{\xi \cdot v : \xi \in \partial_c f(x)\}$, za svaki $v \in \mathbb{R}^n$,
(ii)

$$\partial_c f(x) = \{\xi \in \mathbb{R}^n \ : \ f'(x, v) \geq \xi \cdot v, \ v \in \mathbb{R}^n\}, \tag{1.5}$$

(iii) $\partial_c f(x)$ je neprazan, konveksan i kompaktan skup sadržan u $B(0, K)$ gdje je K Lipschitzova konstanta funkcije f u točki x.
(iv) Ako je f dodatno i diferencijabilna onda vrijedi $\partial_c f(x) = \{\nabla f(x)\}$.

Dokaz. (i) Po definiciji konveksnog subdiferencijala za $y = x + tv, v \in \mathbb{R}^n$, vrijedi

$$\frac{f(x + tv) - f(x)}{t} \geq \frac{\xi \cdot vt}{t} = \xi \cdot v, \ \xi \in \partial_c f(x).$$

Dakle vrijedi

$$f'(x, v) \geq \max\{\xi \cdot v : \xi \in \partial_c f(x)\}.$$

Pokažimo da ne postoji $v \in \mathbb{R}^n$ tako da vrijedi stroga nejednakost. Ako pretpostavimo da takav v_1 postoji, onda po Hahn-Banachovom teoremu postoji $\xi_1 \in \mathbb{R}^n$ takav da za svaki $v \in \mathbb{R}^n$ vrijedi

$$f'(x, v) \geq \xi_1 \cdot v \quad \text{i} \quad f'(x, v_1) = \xi_1 \cdot v_1.$$

Svaki $y \in \mathbb{R}^n$ možemo rastaviti kao $y = x + tv$, za neki $v \in \mathbb{R}^n$ i $t > 0$, pa imamo

$$f(y) - f(x) = t\frac{f(x+tv) - f(x)}{t} \geq tf'(x,v) \geq t\xi_1 \cdot v = \xi_1 \cdot (y - x)$$

iz čega slijedi da je $\xi_1 \in \partial_c f(x)$ pa vrijedi

$$f'(x, v_1) > \max\{\xi \cdot v_1 : \xi \in \partial_c f(x)\} \geq \xi_1 \cdot v_1 = f'(x, v_1),$$

što je nemoguće, dakle, takav v_1 ne postoji.
(ii) Definiramo $H := \{\xi \in \mathbb{R}^n : f'(x,v) \geq \xi \cdot v,\ \ v \in \mathbb{R}^n\}$ i uzmimo proizvoljan $\xi \in H$. Zbog konveksnosti funkcije f i definicije od H, za svaki $v \in \mathbb{R}^n$ vrijedi

$$\begin{aligned}
\xi \cdot v &\leq f'(x,v) \\
&= \lim_{t \to 0+} \frac{f((1-t)x + t(x+v)) - f(x)}{t} \\
&\leq \lim_{t \to 0+} \frac{(1-t)f(x) + tf(x+v) - f(x)}{t} \\
&= f(x+v) - f(x),
\end{aligned}$$

za $t \leq 1$. Ako odaberemo $v = y - x$, za neki $y \in \mathbb{R}^n$, vidimo da $\xi \in \partial_c f(x)$.
Obratno, ako je $\xi \in \partial_c f(x)$ onda za svaki $y \in \mathbb{R}^n$ vrijedi

$$f'(x,y) = \lim_{t \to 0+} \frac{f(x+ty) - f(x)}{t} \geq \lim_{t \to 0+} \frac{\xi \cdot (x + ty - x)}{t} = \xi \cdot y,$$

stoga je $\xi \in H$.
(iii) Po Teoremu 1.1.7. $f'(x,\cdot)$ je pozitivno homogena i subaditivna, pa prema Hahn-Banachovom teoremu postoji vektor $\xi \in \mathbb{R}^n$ takav da je $\xi \cdot v \leq f'(x,v)$ za svaki $v \in \mathbb{R}^n$, dakle $\partial_c f(x)$ je neprazan skup.
Da bismo dokazali konveksnost uzmimo $\xi_1, \xi_2 \in \partial_c f(x)$ i $\lambda \in \mathbb{R}^n$. Tada po točki (i) vrijedi nejednakost

$$\begin{aligned}
(\lambda\xi_1 + (1 - \lambda\xi_2)) \cdot v &= \lambda\xi_1 \cdot v + (1-\lambda)\xi_2 \cdot v \\
&\leq \lambda f'(x,v) + (1-\lambda) f'(x,v) \\
&= f'(x,v)
\end{aligned}$$

Po točki (ii) vrijedi $\lambda\xi_1 + (1-\lambda)\xi_2 \in \partial_c f(x)$, što znači da je $\partial_c f(x)$ konveksan skup.
Uzmimo proizvoljni $\xi \in \partial_c f(x)$ i računamo

$$||\xi||^2 = |\xi \cdot \xi| \leq |f'(x,\xi)| \leq K|\xi|.$$

Dakle, $\partial_c f(x)$ je ograničen skup. Sada još samo treba pokazati da je i zatvoren. Za proizvoljan niz $(\xi_i) \subset \partial_c f(x)$ za koji vrijedi $\xi_i \to \xi$ želimo pokazati da je $\xi \in \partial_c f(x)$. Sada imamo

$$\xi \cdot v = (\lim_{i\to\infty} \xi_i) \cdot v = \lim_{i\to\infty} \xi_i \cdot v \leq \lim_{i\to\infty} f'(x, v) = f'(x, v).$$

Dakle, $\partial_c f(x)$ je kompaktan skup.
(iv) Po definiciji diferencijabilnosti, za svaki $v \in \mathbb{R}^n$ vrijedi

$$f'(x, v) = \nabla f(x) \cdot v,$$

pa po točki (ii) imamo

$$\partial_c f(x) = \{\nabla f(x)\}.$$

□

Napomena 1.2.2. *(a) Točka (i) prethodnog teorema daje nam alternativni način računanja derivacije u smjeru bez računanja limesa.*
(b) Prema točkama (i) i (ii) prethodnog teorema derivacija u smjeru i subdiferencijal jedinstveno određuju jedno drugo.
(c) Točka (iii) prethodnog teorema potvrđuje ono što je spomenuto u uvodu ovog poglavlja, subdiferencijal je generalizacija klasičnog pojma diferencijala.

Po prethodnom teoremu vidimo da je derivacija u smjeru zapravo potporna funkcija subdiferencijala konveksne funkcije u točki x (vidi formulu (1.3)).

1.3 Geometrijska interpretacija subdiferencijala

Znamo da ako je funkcija $f : \mathbb{R}^n \to \mathbb{R}$ diferencijabilna u točki $x \in \mathbb{R}^n$, postoji hiperravnina u $\mathbb{R}^{n+1}$ koja prolazi točkom $(x, f(x)$ i linearna je aproksimacija grafa funkcije f u okolini točki x. Ako je funkcija f konveksna, graf funkcije f mora se u potpunosti nalaziti iznad te hiperravnine. U slučaju da konveksna funkcija nije diferencijabilna u točki x, umjesto jedne, postoji više hiperravnina koje podupiru nadgraf funkcije f u točki $(x, f(x))$.

Definicija 1.3.1. *Za neprazan skup $C \subset \mathbb{R}^n$ je kažemo da je* konveksan *ako za sve $x, y \in C$ i $\lambda \in [0, 1]$ vrijedi*

$$(1 - \lambda)x + \lambda y \in C.$$

Definicija 1.3.2. Tangencijalni konus *zatvorenog konveksnog skupa* $C \subset \mathbb{R}^n$ *u točki* $x \in C$ *definiran je sa*

$$T_C(x) = \mathrm{cl}\,\{d \in \mathbb{R}^n : (\exists y \in C)(\exists \alpha \geq 0),\ d = \alpha(y - x)\},$$

gdje je cl$A$ *oznaka za zatvarač skupa* A.

Skup C iz gornje definicije u daljnjim razmatranjima bit će nadgraf (epigraf) konveksne funkcije definiran na sljedeći način:

Definicija 1.3.3. Epigraf funkcije $f : \mathbb{R}^n \to \mathbb{R}$ *je neprazan skup dan sa*

$$\mathrm{epi}\ f = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : r \geq f(x)\}.$$

Hiperravnine je moguće identificirati pomoću normale i točaka koje sadrže.

Definicija 1.3.4. Normalni konus *skupa* C *u točki* $x \in C$ *određen je sa*

$$N_C(x) = \{s \in \mathbb{R}^n : s \cdot (y - x) \leq 0,\ y \in C\}. \tag{1.6}$$

Radi lakšeg računanja možemo koristiti činjenicu da je tangencijalni konus ujedno i polarni konus normalnog konusa i obratno (vidi [4], str. 66.). To vrijedi u slučaju da je skup C konveksan skup.

Definicija 1.3.5. *Neka je* K *konveksan konus.* Polarni konus *konusa* K *zadan je s*

$$K^\circ = \{s \in \mathbb{R}^n : s \cdot x \leq 0,\ x \in K\}.$$

Sljedeća propozicija daje, uz (1.4) i (1.5), treći način računanja subdiferencijala konveksne funkcije.

Propozicija 1.3.6. *Neka je* $f : \mathbb{R}^n \to \mathbb{R}$ *konveksna funkcija.*
(i) Vektor $s \in \mathbb{R}^n$ *je subgradijent funkcije* f *u točki* $x \in \mathbb{R}^n$ *ako i samo ako je* $(s, -1) \in \mathbb{R}^n \times \mathbb{R}$ *normala na* epi f *u točki* $(x, f(x))$*, odnosno vrijedi*

$$N_{epi\ f}(x, f(x)) = \{(\lambda s, -\lambda) \in \mathbb{R}^n \times \mathbb{R} : s \in \partial_c f(x), \lambda \geq 0\}. \tag{1.7}$$

(ii) Tangencijalni konus na skup epi$f$ *u točki* $(x, f(x))$ *epigraf je funkcije* $v \mapsto f'(x, v)$*, dakle vrijedi*

$$T_{epi\ f}(x, f(x)) = \{(d, r) \in \mathbb{R}^n \times \mathbb{R} : r \geq f'(x, d)\}.$$

Dokaz. (i) Pretpostavimo da je $s \in \partial_c f(x)$. Po formuli (1.4) iz definicije subdiferencijala, za svaki $y \in \mathbb{R}^n$ vrijedi:

$$\begin{aligned} f(y) \geq f(x) + s \cdot (y - x) &\iff 0 \geq s \cdot (y - x) - (f(y) - f(x)) \\ &\iff 0 \geq (s, -1) \cdot (y - x, f(y) - f(x)) \\ &\iff 0 \geq (s, -1) \cdot [(y, f(y)) - (x, f(x))], \end{aligned}$$

pa prema tome, za svaki $r \in \mathbb{R}$ takav da $r \geq f(y)$, vrijedi

$$0 \geq (s, -1) \cdot [(y, r) - (x, f(x))], \quad (y, r) \in \operatorname{epi} f. \tag{1.8}$$

Po formuli (1.6) zaključujemo da je $(s, -1)$ normala na $\operatorname{epi} f$ u točki $(x, f(x))$. Da bismo pokazali obrat, potrebno je u (1.8) uzeti $r = f(y)$, a rezultat onda slijedi po gornjim ekvivalencijama.
Formula (1.7) slijedi iz činjenice da skup normala tvori konus koji sadrži ishodište.
(ii) Budući da je tangencijalni konus polarni konus konusa iz (1.7), sadrži sve $(d, r) \in \mathbb{R}^n \times \mathbb{R}$ tako da vrijedi

$$\lambda s \cdot d + (-\lambda) r \leq 0, \quad s \in \partial_c f(x) \quad \text{i} \quad \lambda \geq 0.$$

Za $\lambda = 0$ tvrdnja trivijalno vrijedi, a za $\lambda > 0$ možemo je ekvivalentno zapisati kao

$$r \geq \max\{s \cdot d : s \in \partial_c f(x)\} = f'(x, d).$$

□

Ranije smo izračunali da je subdiferencijal funkcije $f(x) = |x|$ u nuli jednak segmentu $[-1, 1]$. Pokušajmo ga sada izračunati pomoću gornje propozicije. Epigraf funkcije $f(x) = |x|$ je očito konveksan konus, a lako se vidi da je njegov polarni konus određen površinom ispod grafa funkcije $f(x) = -|x|$, pa je presjek tog polarnog konusa sa horizontalnim pravcem koji prolazi kroz točku $(0, -1)$ upravo segment $[-1, 1]$.
Derivacija u smjeru funkcije apsolutne vrijednosti u nuli lako se izračuna po definiciji

$$f'(0, v) = \lim_{t \to 0+} \frac{|0 + tv| - |0|}{t} = |v|,$$

iz čega slijedi da se tangencijalni konus funkcije $f(x) = |x|$ u nuli podudara sa epigrafom te iste funkcije.
U slučaju da je funkcija diferencijabilna u točki x tada je tangencijalni konus u toj točki cijeli poluprostor koji sadrži epigraf funkcije f.

Poglavlje 2

Neglatka analiza

U teoriji optimizacije mnoge metode rješavanja zahtijevaju diferencijabilnost promatranih funkcija. No, u praktičnim primjenama često nije moguće svesti problem na glatke funkcije bez velikih grešaka u aproksimaciji. To je upravo bila motivacija za razvitak neglatke analize.
Prethodne rezultate želimo poopćiti na proizvoljne lokalno Lipschitzove funkcije. Postoji više načina generalizacije, u ovom poglavlju slijedimo Clarkeov[1] pristup.
Nakon toga ćemo dati nekoliko pravila za lakše računanje sa subdiferencijalima, a glavni rezultat ovog poglavlja bit će teorem srednje vrijednosti. Jedan od razloga potrebe za generalizacijom je taj što obična derivacija u smjeru nije sublinearna za općenite Lipschitzove funkcije, što znači da će često subdiferencijal, definiran za konveksne funkcije, biti prazan skup.

2.1 Clarkeova generalizacija derivacije u smjeru

Clarkeova derivacija u smjeru

U ovom dijelu definiramo Clarkeovu derivaciju u smjeru i Clarkeov subdiferencijal, iznosimo glavna svojstva i uspoređujemo sa običnom derivacijom u smjeru i konveksnim subdiferencijalom.

Definicija 2.1.1 (Clarkeova derivacija u smjeru)**.** *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzova funkcija u $x \in \mathbb{R}^n$.* Clarkeova derivacija funkcije f u smjeru $v \in \mathbb{R}^n$ u točki $x \in \mathbb{R}^n$ *definirana je sa*

$$f^\circ(x, v) = \limsup_{y \to x, t \to 0+} \frac{f(y + tv) - f(y)}{t}.$$

[1]Frank Clarke, kanadski matematičar rođen 1948. godine

Možemo primijetiti da Clarkeova derivacija ne pretpostavlja postojanje limesa, točka y u brojniku nije fiksna, ali se nalazi u okolini točke x. Neka svojstva Clarkeove derivacije dana su u sljedećem teoremu.

Teorem 2.1.2 (Svojstva Clarkeove derivacije)**.** *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzova s konstantom K. Tada*
(i) funkcija $v \mapsto f^\circ(x, v)$ je pozitivno homogena i subaditivna na $\mathbb{R}^n$ te vrijedi

$$|f^\circ(x, v)| \leq K\|v\|,$$

(ii) $f^\circ(x, v)$ je odozgo poluneprekidna kao funkcija od $(x, v) \in \mathbb{R}^n \times \mathbb{R}^n$ i Lipschitzova s konstantom K kao funkcija od $v \in \mathbb{R}^n$,
(iii) $f^\circ(x, -v) = (-f)^\circ(x, v)$

Dokaz. (i) Zbog Lipschitzovosti funkcije f vrijedi

$$\begin{aligned} |f^\circ(x, v)| &\leq \left| \limsup_{y \to x, t \to 0+} \frac{f(y + tv) - f(y)}{t} \right| \\ &\leq \limsup_{y \to x, t \to 0+} \frac{|f(y + tv) - f(y)|}{t} \\ &\leq \limsup_{y \to x, t \to 0+} \frac{K\|y + tv - y\|}{t} = K\|v\|, \end{aligned}$$

za sve $y, y + tv \in B(x, \epsilon)$.
Za dokaz pozitivne homogenosti uzmimo $\lambda > 0$ i računamo

$$\begin{aligned} f^\circ(x, \lambda v) &= \limsup_{y \to x, t \to 0+} \frac{f(y + \lambda tv) - f(y)}{t} \\ &= \limsup_{y \to x, t \to 0+} \lambda \frac{f(y + \lambda tv) - f(y)}{\lambda t} \\ &= \lambda \limsup_{y \to x, t \to 0+} \frac{f(y + \lambda tv) - f(y)}{\lambda t} = \lambda f^\circ(x, v). \end{aligned}$$

Subaditivnost, za $v, w \in \mathbb{R}^n$, pokazujemo na sljedeći način

$$\begin{aligned} f^\circ(x, v + w) &= \limsup_{y \to x, t \to 0+} \frac{f(y + t(v + w) - f(y))}{t} \\ &= \limsup_{y \to x, t \to 0+} \frac{f(y + tv + tw) - f(y + tw) + f(y + tw) - f(y)}{t} \\ &\leq \limsup_{y \to x, t \to 0+} \frac{f((y + tw) + tv) - f(y + tw)}{t} + \limsup_{y \to x, t \to 0+} \frac{f(y + tw) - f(y)}{t} \\ &= f^\circ(x, v) + f^\circ(x, w). \end{aligned}$$

(ii) Neka su $(x_i), (v_i) \subset \mathbb{R}^n$ nizovi takvi da $x_i \to x$ i $v_i \to v$. Po svojstvu limes superiora postoje nizovi $(y_i) \subset \mathbb{R}^n$ i $(t_i) \subset \mathbb{R}^n$, $t_i > 0$ takvi da vrijedi

$$f^\circ(x_i, v_i) \leq \frac{f(y_i + t_i v_i) - f(y_i)}{t_i} + \frac{1}{i}$$

i

$$\|y_i - x_i\| + t_i \leq \frac{1}{i}, \quad i \in \mathbb{N}$$

Sada imamo

$$\begin{aligned} f^\circ(x_i, v_i) - \frac{1}{i} &= \limsup_{y \to x_i, t \to 0+} \frac{f(y + tv_i) - f(y)}{t} - \frac{1}{i} \\ &\leq \frac{f(y_i + t_i v_i) - f(y_i)}{t_i} \\ &= \frac{f(y_i + t_i v) - f(y_i)}{t_i} + \frac{f(y_i + t_i v_i) - f(y_i + t_i v)}{t_i}, \end{aligned}$$

gdje, za zadnji član, po Lispchitzovom svojstvu vrijedi

$$\frac{|f(y_i + t_i v_i) - f(y_i + t_i v)|}{t_i} \leq \frac{K\|t_i v_i - t_i v\|}{t_i} = K\|v_i - v\| \to 0,$$

za $i \to \infty$, pod uvjetom da su $y_i + t_i v_i, y_i + t_i v \in B(x, \epsilon)$. Uzimanjem limes superiora za $i \to \infty$ napokon dobivamo

$$\limsup_{i \to \infty} f^\circ(x_i, v_i) \leq \limsup_{i \to \infty} \frac{f(y_i + t_i v) - f(y_i)}{t_i} \leq f^\circ(x, v),$$

što je upravo definicija poluneprekidnosti odozgo. Preostalo je pokazati Lipschitzovost u drugoj varijabli. Neka su $v, w \in \mathbb{R}^n$ zadani. Za vektore $y + tv, y + tw \in B(x, \epsilon)$, zbog lokalne Lipschitzovosti funkcije f vrijedi

$$f(y + tv) - f(y + tw) \leq Kt\|v - w\|,$$

pa prema tome

$$\limsup_{y \to x, t \to 0+} \frac{f(y + tv) - f(y)}{t} \leq \limsup_{y \to x, t \to 0+} \frac{f(y + tw) - f(y)}{t} + K\|v - w\|,$$

iz čega slijedi

$$f^\circ(x, v) - f^\circ(x, v) \leq K\|v - w\|.$$

Zamjenom uloga od v i w imamo

$$f^\circ(x, w) - f^\circ(x, v) \leq K\|v - w\|,$$

pa konačno

$$|f^\circ(x, v) - f^\circ(x, w)| \leq K\|v - w\|.$$

(iii) Računamo

$$\begin{aligned} f^\circ(x, -v) &= \limsup_{y \to x, t \to 0+} \frac{f(y - tv) - f(y)}{t} = [u := y - tv] \\ &= \limsup_{y \to x, t \to 0+} \frac{(-f)(u + tv) - (-f)(u)}{t} \\ &= (-f)^\circ(x, v). \end{aligned}$$

□

Sada možemo definirati generalizirani subdiferencijal nekonveksne Lipschitzove funkcije.

Generalizirani subdiferencijal

Definicija 2.1.3 (Clarkeov subdiferencijal)**.** *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzova funkcija u $x \in \mathbb{R}^n$.* Clarkeov subdiferencijal funkcije f u točki x *je sljedeći skup*

$$\partial^\circ f(x) = \{\xi \in \mathbb{R}^n : f^\circ(x, v) \geq \xi \cdot v, \ v \in \mathbb{R}^n\}.$$

Svaki element $\xi \in \partial^\circ f(x)$ nazivamo subgradijentom *funkcije f u točki x.*

Primijetimo da je Clarkeov subdiferencijal definiran prema svojstvu subdiferencijala konveksne funkcije iz (1.5).

Teorem 2.1.4 (Svojstva Clarkeovog subdiferencijala)**.** *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzova funkcija u $x \in \mathbb{R}^n$ s konstantom K. Tada*
(i) $\partial^\circ f(x)$ je neprazan, konveksan i kompaktan skup takav da je $\partial^\circ f(x) \subset B(0, K)$,
(ii)

$$f^\circ(x, v) = \max\{\xi \cdot v \ : \ \xi \in \partial^\circ f(x)\}, \ v \in \mathbb{R}^n, \tag{2.1}$$

(iii) ako je f diferencijabilna u x onda vrijedi $\nabla f(x) \in \partial^\circ f(x)$.

Dokaz. (i) i (ii) se dokazuju na isti način kao i točke (i),(iii) u Teoremu 1.1.7.
(iii) Zbog diferencijabilnosti funkcije f u x za svaki $v \in \mathbb{R}^n$ vrijedi $f'(x, v) = \nabla f(x) \cdot v$, a budući da je $f' \leq f^\circ$, imamo

$$f^\circ(x, v) \geq \nabla f(x) \cdot v, \quad v \in \mathbb{R}^n$$

pa slijedi $\nabla f(x) \in \partial^\circ f(x)$. □

Za kraj ovog dijela dan je rezultat koji potvrđuje da se subdiferencijal konveksne funkcije i Clarkeov subdiferencijal podudaraju za konveksne funkcije.

Teorem 2.1.5. *Ako je funkcija $f : \mathbb{R}^n \to \mathbb{R}$ konveksna, tada*
(i) $f'(x, v) = f^\circ(x, v), \ v \in \mathbb{R}^n$,
(ii) $\partial_c f(x) = \partial^\circ f(x)$.

Dokaz. Dovoljno je dokazati (i) jer (ii) onda slijedi iz (1.5) i definicije konveksnog subdiferencijala.
Primijetimo da po definiciji od $f^\circ(x, \cdot)$ vrijedi $f^\circ(x, v) \geq f'(x, v), v \in \mathbb{R}^n$. S druge strane, ako fiksiramo $\delta > 0$, imamo

$$\begin{aligned} f^\circ(x, v) &= \limsup_{y \to x, t \to 0+} \frac{f(y + tv) - f(y)}{t} \\ &= \lim_{\epsilon \downarrow 0} \sup_{\|y - x\| < \epsilon\delta} \sup_{0 < t < \epsilon} \frac{f(y + tv) - f(y)}{t}. \end{aligned}$$

Iz dokaza Teorema 1.1.7 znamo da je funkcija $\phi(t) = \frac{f(y+tv)-f(y)}{t}$ neopadajuća, pa vrijedi

$$f^\circ(x, v) = \lim_{\epsilon \downarrow 0} \sup_{\|y - x\| < \epsilon\delta} \frac{f(y + \epsilon v) - f(y)}{\epsilon}.$$

Budući da je f Lipschitzova gornji supremum možemo, za $y \in B(x, \epsilon\delta)$, ograničiti na sljedeći način

$$\begin{aligned} \left| \frac{f(y + \epsilon v) - f(y)}{\epsilon} - \frac{f(x + \epsilon v) - f(x)}{\epsilon} \right| &\leq \left| \frac{f(y + \epsilon v) - f(x + \epsilon v)}{\epsilon} \right| + \left| \frac{f(x) - f(y)}{\epsilon} \right| \\ &\leq \frac{K}{\epsilon} \|y - x\| + \frac{K}{\epsilon} \|y - x\| \\ &\leq \frac{2K}{\epsilon} \epsilon\delta \\ &= 2K\delta, \end{aligned}$$

iz čega vidimo da vrijedi

$$f^\circ(x, v) \le \lim_{\epsilon \downarrow 0} \frac{f(x + \epsilon v) - f(x)}{\epsilon} + 2\delta K = f'(x, v) + 2\delta K.$$

Budući da je $\delta > 0$ bio proizvoljan, vrijedi

$$f^\circ(x, v) \le f'(x, v), \quad v \in \mathbb{R}^n.$$

□

2.2 Subdiferencijalni račun

Računanje subdiferencijala u praksi, pomoću definicija, često nije jednostavno. Slično, kao i za računanje diferencijala u glatkoj analizi, postoje formule za subdiferencijal linearne kombinacije, produkta i kvocijenta funkcija, no, budući da su to skupovi, u najopćenitijem slučaju vrijedi samo jedna inkluzija.

Definicija 2.2.1. *Za funkciju $f : \mathbb{R}^n \to \mathbb{R}$ kažemo da je* regularna *u $x \in \mathbb{R}^n$ ako za svaki $v \in \mathbb{R}^n$ derivacija u smjeru $f'(x, v)$ postoji i vrijedi*

$$f'(x, v) = f^\circ(x, v).$$

Neki od dovoljnih uvjeta regularnosti dani su u sljedećoj propoziciji.

Propozicija 2.2.2. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzova u $x \in \mathbb{R}^n$. Tada je f regularna u x ako f zadovoljava barem jedan od sljedećih uvjeta:*
(i) f je neprekidno diferencijabilna u x,
(ii) f je konveksna,
(iii) $f = \sum_{i=1}^m \lambda_i f_i$ za $\lambda_i > 0$ i f_i su regularne u x za $i = 1, \ldots, m$.

Dokaz. (i) Da je f neprekidno derivabilna u x znači da $x_i \to x$ povlači $\nabla f(x_i) \to \nabla f(x)$. Za svaki $v \in \mathbb{R}^n$ tada vrijedi

$$\begin{aligned} \lim_{x_i \to x} f'(x_i, v) &= \lim_{x_i \to x} \lim_{t \to 0+} \frac{f(x_i + tv) - f(x_i)}{t} \\ &= \lim_{x_i \to x} \nabla f(x_i) \cdot v \\ &= \nabla f(x) \cdot v = f'(x, v), \end{aligned}$$

i, nadalje

$$\begin{aligned} f'(x,v) &= \lim_{x_i \to x} f'(x_i, v) \\ &= \lim_{x_i \to x, t \to 0+} \frac{f(x_i + tv) - f(x_i)}{t} \\ &= \limsup_{x_i \to x, t \to 0+} \frac{f(x_i + tv) - f(x_i)}{t} = f^\circ(x,v). \end{aligned}$$

(ii) Tvrdnja slijedi iz Teorema 2.1.5.
(iii) Uočimo da je dovoljno dokazati za $m = 2$, jer za $m > 2$ koristi se indukcija po m. Ako je f regularna i $\lambda > 0$, tada za svaki $v \in \mathbb{R}^n$ vrijedi

$$(\lambda f)^\circ(x,v) = \lambda f^\circ(x,v) = \lambda f'(x,v) = (\lambda f)'(x,v).$$

Jasno je da $(f_1 + f_2)'$ postoji i vrijedi $(f_1 + f_2)' = f_1' + f_2'$. Po definiciji Clarkeove derivacije vrijedi $(f_1 + f_2)^\circ \geq (f_1 + f_2)'$, a za dokaz suprotne nejednakosti računamo

$$\begin{aligned} (f_1 + f_2)^\circ(x,v) &= \limsup_{y \to x, t \to 0+} \frac{(f_1 + f_2)(y + tv) - (f_1 + f_2)(y)}{t} \\ &= \limsup_{y \to x, t \to 0+} \frac{f_1(y + tv) + f_2(y + tv) - f_1(y) - f_2(y)}{t} \\ &\leq \limsup_{y \to x, t \to 0+} \frac{f_1(y + tv) - f_1(y)}{t} + \limsup_{y \to x, t \to 0+} \frac{f_2(y + tv) - f_2(y)}{t} \\ &= f_1^\circ(x,v) + f_2^\circ(x,v). \end{aligned}$$

Sada vrijedi

$$(f_1 + f_2)' = f_1' + f_2' = f_1^\circ + f_2^\circ \geq (f_1 + f_2)^\circ,$$

pa na kraju imamo

$$(f_1 + f_2)' = (f_1 + f_2)^\circ.$$

□

U sljedećim teoremima dane su osnovne formule za računanje sa subdiferencijalima.

Teorem 2.2.3. *Ako je f lokalno Lipschitzova u x, tada za svaki $\lambda \in \mathbb{R}$ vrijedi*

$$\partial^\circ(\lambda f)(x) = \lambda \partial^\circ f(x).$$

Dokaz. Očito je funkcija λf lokalno Lipschitzova u točki x. Za $\lambda \geq 0$ vrijedi $(\lambda f)^\circ = \lambda f^\circ$, iz čega slijedi $\partial^\circ(\lambda f)(x) = \lambda \partial^\circ f(x)$. Za $\lambda < 0$ dovoljno je pokazati za $\lambda = -1$, pa računamo

$$\begin{aligned}
\xi \in \partial^\circ(-f)(x) &\iff (-f)^\circ(x, v) \geq \xi \cdot v, \quad v \in \mathbb{R}^n \\
&\iff f^\circ(x, -v) \geq \xi \cdot v, \quad v \in \mathbb{R}^n \\
&\iff f^\circ(x, -v) \geq (-\xi) \cdot (-v), \quad (-v) \in \mathbb{R}^n \\
&\iff -\xi \in \partial^\circ f(x) \\
&\iff \xi \in -\partial^\circ f(x)
\end{aligned}$$

□

Teorem 2.2.4. *Neka su funkcije $f_i : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzove u x za $i = 1, \ldots, m$, tada za skalare $\lambda_i \in \mathbb{R}$ vrijedi*

$$\partial^\circ \left(\sum_{i=1}^{m} \lambda_i f_i \right)(x) \subset \sum_{i=1}^{m} \lambda_i \partial^\circ f_i(x),$$

te vrijedi jednakost ako je svaka funkcija f_i regularna u x i svaki skalar λ_i 0.

Dokaz. Dovoljno je dokazati za $m = 2$, općeniti slučaj se lako dokaže indukcijom. U dokazu Propozicije 2.2.2. (iii) izvedena je sljedeća nejednakost

$$(f_1 + f_2)^\circ(x, v) \leq f_1^\circ(x, v) + f_2^\circ(x, v),$$

pa po definiciji Clarkeovog subdiferencijala vrijedi

$$\partial^\circ(f_1 + f_2)(x) \subset \partial^\circ f_1(x) + \partial^\circ f_2(x).$$

Teorem 2.2.3. povlači

$$\begin{aligned}
\partial^\circ(\lambda_1 f_1 + \lambda_2 f_2)(x) &\subset \partial^\circ(\lambda_1 f_1)(x) + \partial^\circ(\lambda_2 f_2)(x) \\
&= \lambda_1 \partial^\circ f_1(x) + \lambda_2 \partial^\circ f_2(x),
\end{aligned}$$

i time je prvi dio teorema dokazan.
Pretpostavimo da su f_1 i f_2 regularne i $\lambda_1, \lambda_2 > 0$. Po Propoziciji 2.2.2. (iii) funkcija $\lambda_1 f_1 + \lambda_2 f_2$ je regularna i vrijedi

$$(\lambda_1 f_1 + \lambda_2 f_2)^\circ = (\lambda_1 f_1 + \lambda_2 f_2)' = \lambda_1 f_1' + \lambda_2 f_2' = \lambda_1 f_1^\circ + \lambda_2 f_2^\circ.$$

Na kraju, po definiciji Clarkeovog subdiferencijala vrijedi

$$\partial^\circ(\lambda_1 f_1 + \lambda_2 f_2)(x) = \lambda_1 \partial^\circ f_1(x) + \lambda_2 \partial^\circ f_2(x).$$

□

Ako se prisjetimo klasične analize, znamo da za ekstreme glatke funkcije vrijedi da je derivacija funkcije u tim točkama jednaka nuli. Sljedeći rezultat je upravo poopćenje tog svojstva glatkih funkcija na općenite Lipschitzove funkcije.

Teorem 2.2.5. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzova funkcija u $x \in \mathbb{R}^n$ i neka postiže lokalni ekstrem u x. Tada vrijedi*

$$0 \in \partial^\circ f(x).$$

Dokaz. Pretpostavimo da se postiže lokalni minimum u x. Tada postoji $\epsilon > 0$ tako da vrijedi

$$f(x + tv) - f(x) \geq 0, \quad 0 < t < \epsilon, \; v \in \mathbb{R}^n.$$

Sada po definiciji Clarkeove derivacije i zbog gornje nejednakosti vrijedi

$$f^\circ(x, v) = \limsup_{y \to x, t \to 0+} \frac{f(y + tv) - f(y)}{t} \geq \limsup_{t \to 0+} \frac{f(x + tv) - f(x)}{t} \geq 0,$$

pa prema tome

$$f^\circ(x, v) \geq 0 = 0 \cdot v, \quad v \in \mathbb{R}^n.$$

Dakle, $0 \in \partial^\circ f(x)$. U slučaju maksimuma, možemo promatrati funkciju $-f$, pa slično kao gore, vrijedi $0 \in \partial^\circ(-f)(x)$. Tada tvrdnja slijedi po Teoremu 2.2.3. □

Teorem srednje vrijednosti

Lema 2.2.6. *Neka su $x, y \in \mathbb{R}^n, x \neq y$ i neka je $f : \mathbb{R}^n \to \mathbb{R}$ Lipschitzova funkcija na segmentu $[x, y]$. Tada je funkcija $g : [0, 1] \to \mathbb{R}$, definirana sa $g(t) := f(x + t(y - x))$ Lipschitzova na $\langle 0, 1 \rangle$ i vrijedi*

$$\partial^\circ g(t) \subset \partial^\circ f(x + t(y - x)) \cdot (y - x).$$

Dokaz. Neka je $x_t := x + t(y - x)$. Za sve $t, t' \in \langle 0, 1 \rangle$ vrijedi

$$\begin{aligned} |g(t) - g(t')| &= |f(x_t) - f(x_{t'})| \\ &\leq K\|x_t - x_{t'}\| \\ &= K\|(t - t')(y - x)\| \\ &= K|t - t'|\|y - x\| \\ &= \hat{K}|t - t'|, \text{ za } \hat{K} = K\|y - x\|. \end{aligned}$$

Dakle, g je Lipschitzova na $\langle 0,1\rangle$ pa po Teoremu 2.1.4. skupovi $\partial^\circ g(t)$ i $\partial^\circ f(x_t)\cdot(y-x)$ su kompaktni i konveksni, a budući da su podskupovi od $\mathbb{R}$, moraju biti zatvoreni intervali u $\mathbb{R}$. Zato je dovoljno sljedeću nejednakost pokazati samo za $v=\pm 1$:

$$\max\{\partial^\circ g(t)v\} \le \max\{\partial^\circ f(x_t)\cdot(y-x)v\}.$$

Po Teoremu 2.1.4. (ii) vrijedi $\max\{\partial^\circ g(t)v\} = g^\circ(t,v)$ i računamo

$$\begin{aligned}
\max\{\partial^\circ g(t)v\} &= \limsup_{s\to t,\lambda\to 0+} \frac{g(s+\lambda v)-g(s)}{\lambda} \\
&= \limsup_{s\to t,\lambda\to 0+} \frac{f(x+[s+\lambda v](y-x))-f(x+s(y-x))}{\lambda} \\
&\le \limsup_{y'\to x_t,\lambda\to 0+} \frac{f(y'+\lambda v(y-x))-f(y')}{\lambda} \\
&= f^\circ(x_t, v(y-x)) \\
&= \max\{\partial^\circ f(x_t)\cdot(v(y-x))\}.
\end{aligned}$$

□

Ova lema je potrebna za dokaz teorema srednje vrijednosti, koji u glatkoj analizi ima vrlo važnu ulogu, a ovdje ga poopćavamo za neglatke Lipschitzove funkcije.

Teorem 2.2.7 (Teorem srednje vrijednosti)**.** *Neka su $x,y\in\mathbb{R}, x\neq y$, te neka je funkcija f Lipschitzova na otvorenom skupu $U\subset\mathbb{R}^n$ tako da je $[x,y]\subset U$. Tada postoji točka $u\in\langle x,y\rangle$ takva da vrijedi*

$$f(y)-f(x)\in\partial^\circ f(u)\cdot(y-x).$$

Dokaz. Definiramo funkciju $\Theta:[0,1]\to\mathbb{R}$ tako da vrijedi

$$\Theta(t) := f(x_t)+t(f(x)-f(y)),$$

gdje je $x_t = x+t(y-x)$. Funkcija Θ je očito neprekidna i vrijedi

$$\begin{aligned}
&\Theta(0) = f(x_0) = f(x),\\
&\Theta(1) = f(x_1)+f(x)-f(y) = f(x).
\end{aligned}$$

Tada, po Bolzano-Weierstrassovom teoremu (vidi [3], str. 77.) postoji $t_o\in\langle 0,1\rangle$ tako da funkcija Θ postiže lokalni ekstrem u t_0 pa, po Teoremu 2.2.5, mora biti $0\in\partial^\circ\Theta(t_0)$. Zatim, po Teoremu 2.2.4, imamo

$$\partial^\circ\Theta(t) = \partial^\circ[f(x_t)+t(f(x)-f(y))]\subset\partial^\circ f(x_t)+[f(x)-f(y)]\partial^\circ(t),$$

a po Lemi 2.2.6. vrijedi

$$0 \in \partial^\circ f(x_t) \cdot (y - x) + [f(x) - f(y)] \cdot \partial^\circ(t),$$

a budući da je $\partial^\circ(t) = 1$ imamo

$$f(y) - f(x) \in \partial^\circ f(u)(y - x),$$

gdje je $u := x_t \in \langle 0, 1\rangle$. □

Iduća tri teorema dana su radi potpunosti i ilustracije poopćenja derivabilnosti kompozicije, produkta i kvocijenta Lipschitzovih funkcija. (Dokaz, vidi [9] str. 42–49.)

Teorem 2.2.8 (Lančano pravilo). *Neka su $h : \mathbb{R}^n \to \mathbb{R}^m$ i $g : \mathbb{R}^m \to \mathbb{R}$ funkcije takve da je svaka komponentna funkcija $h_i : \mathbb{R} \to \mathbb{R}, i = 1, \ldots, m$ lokalno Lipschitzova u $x \in \mathbb{R}^n$ i g lokalno Lipschitzova u $h(x) \in \mathbb{R}^m$. Tada je kompozicija funkcija $f = g \circ h, f : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzova u x i vrijedi*

$$\partial^\circ f(x) \subset \operatorname{conv}\left\{ \sum_{i=1}^{m} \alpha_i \xi_i : \xi_i \in \partial^\circ h_i(x) \quad \text{i} \quad \alpha \in \partial^\circ g(h(x))\right\}.$$

Jednakost vrijedi ako je zadovoljen bilo koji od sljedećih uvjeta:
(i) Funkcija g je regularna u $h(x)$, svaka h_i je regularna u x i $\alpha_i \geq 0$ za sve $i = 1, \ldots, m$. Tada je i f regularna u x.
(ii) Funkcija g je regularna u $h(x)$ i h_i je neprekidno diferencijabilna u x za svaki $i = 1, \ldots, m$.
(iii) Za slučaj $m = 1$, g je neprekidno diferencijabilna u $h(x)$.

Teorem 2.2.9 (Produkt). *Neka su f_1 i f_2 lokalno Lipschitzove funkcije u točki $x \in \mathbb{R}^n$. Tada je funkcija $f_1 f_2$ lokalno Lipschitzova u x i vrijedi*

$$\partial^\circ(f_1 f_2)(x) \subset f_2(x)\partial^\circ f_1(x) + f_1(x)\partial^\circ f_2(x).$$

Dodatno, ako $f_1(x), f_2(x) \geq 0$ i f_1, f_2 regularne u x, tada vrijedi jednakost i funkcija $f_1 f_2$ je regularna u x.

Teorem 2.2.10 (Kvocijent). *Neka su f_1 i f_2 lokalno Lipschitzove funkcije u $x \in \mathbb{R}^n$ i $f_2(x) \neq 0$. Tada je funkcija f_1/f_2 lokalno Lipscchitzova u x i vrijedi*

$$\partial^\circ\left(\frac{f_1}{f_2}\right)(x) \subset \frac{f_2(x)\partial^\circ f_1(x) - f_1(x)\partial^\circ f_2(x)}{f_2^2(x)}.$$

Nadalje, ako $f_1 \geq 0, f_2 \geq 0$ i f_1, f_2 su obje regularne u x, vrijedi jednakost i funkcija f_1/f_2 je regularna u x.

Sljedeći teorem koristan je u praksi u neglatkoj optimizaciji, a odnosi se na klasu max-funkcija.

Teorem 2.2.11. *Neka su $f_i : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzove funkcije u $x \in \mathbb{R}^n$ za svaki $i = 1, \ldots, m$. Tada je funkcija $f : \mathbb{R}^n \to \mathbb{R}$ definirana sa*

$$f(x) = \max\{f_i(x) \ : \ i = 1, \ldots, m\}$$

lokalno Lipschitzova u x i vrijedi

$$\partial^\circ f(x) \subset \operatorname{conv} \bigcup_{i \in I(x)} \partial^\circ f_i(x),$$

gdje je $I(x) = \{i \in \{1, \ldots, m\} \ : \ f_i(x) = f(x)\}$. Nadalje, ako su f_i regularne u x za svaki $i = 1, \ldots, m$ tada je f regularna u x i vrijedi jednakost umjesto inkluzije.

Dokaz. (Vidi [9] str. 48.) □

U slučaju beskonačnog skupa indeksa koristimo je sljedeći teorem:

Teorem 2.2.12. *Neka je J proizvoljan kompaktan skup indeksa (u nekom metričkom prostoru) i $\{f_j\}_{j \in J}$ skup konveksnih funkcija sa $\mathbb{R}^n$ u $\mathbb{R}$ tako da su funkcije $j \mapsto f_j(x)$ odozgo poluneprekidne. Ako definiramo funkciju $f : \mathbb{R}^n \to \mathbb{R}$ kao*

$$f(x) := \max\{f_j(x) : j \in J\} < \infty, \ x \in \mathbb{R}^n,$$

tada vrijedi

$$\partial_c f(x) = \operatorname{conv} \bigcup_{j \in J(x)} \partial_c f_j(x), \tag{2.2}$$

gdje je $J(x) = \{j \in J : f_j(x) = f(x)\}$.

Dokaz. (Vidi [4] str. 189.) □

Teorem 2.2.13 (Rademacher)**.** *Neka je $U \subset \mathbb{R}^n$ otvoren skup i $f : U \to \mathbb{R}$ Lipschitzova funkcija na U. Tada je f diferencijabilna gotovo svuda na U.*

Dokaz. (Vidi [9] str. 49.) □

Za kraj ovog poglavlja dan je teorem o računanju subdiferencijala pomoću limesa gradijenata. Budući da je prema prethodnom Lipschitzova funkcija diferencijabilna gotovo svuda, skup mjere nula na kojem nije diferencijabilna označimo sa Ω_f.

Teorem 2.2.14. *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzova u $x \in \mathbb{R}^n$. Tada*

$$\partial f(x) = \operatorname{conv}\{\xi \in \mathbb{R}^n : \exists (x_i) \subset \mathbb{R}^n \backslash \Omega_f \ \text{ tako da } x_i \to x \text{ i } \nabla f(x_i) \to \xi\}.$$

Dokaz. (Vidi [9] str. 50.) □

Poglavlje 3

Svojstvene vrijednosti simetrične matrice

U ovom poglavlju promatramo pojam svojstvene vrijednosti realne matrice. Uz pomoć nekih rezultata spektralne analize prvo računamo običnu derivaciju u smjeru i konveksni subdiferencijal. Nakon toga uvodimo još jednu generaliziranu derivaciju u smjeru, Michel-Penotovu derivaciju u smjeru i pripadni Michel-Penotov subdiferencijal. Glavni rezultat ovog poglavlja je jednakost Clarkeovog i Michel-Penotovog subdiferencijala za funkciju m-te svojstvene vrijednosti simetrične matrice.

3.1 Derivacija najveće svojstvene vrijednosti

Prisjetimo se da je svojstvena vrijednost matrice $A \in M_n(F)$ skalar $\lambda \in F$ za koji postoji netrivijalan vektor $v \in F^n$ takav da vrijedi $Av = \lambda v$. Takav vektor v nazivamo svojstvenim vektorom pridruženim svojstvenoj vrijednosti λ. Prema spektralnom teoremu, simetričnu, realnu, $n \times n$ matricu A možemo rastaviti na sljedeći način

$$A = UDU^{\tau},$$

gdje je U realna ortogonalna matrica kojoj su stupci svojstveni vektori pridruženi svojstvenim vrijednostima matrice A (vidi [1] str. 10.). Matrica D je dijagonalna matrica koja na dijagonali ima svojstvene vrijednosti matrice A. Ako prirodno definiramo funkciju
Diag : $\mathbb{R}^n \to S^n$, matricu D možemo pisati i kao $D = \mathrm{Diag}\,(\lambda_1, \ldots, \lambda_n)$.
Na prostoru realnih simetričnih matrica S^n koristimo standardni skalarni produkt matrica A i B

$$\langle\!\langle A, B \rangle\!\rangle = \mathrm{tr}(AB) = \sum_{i,j} a_{ij} b_{ij}, \quad A, B \in S^n.$$

Budući da su sve svojstvene vrijednosti realne, možemo ih poredati po veličini, m-ta najveća svojstvena vrijednost matrice A, brojeći i kratnost, označava se sa $\lambda_m(A)$.
Dakle, vrijedi

$$\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_n(A)$$

i $\lambda_m : S^n \to \mathbb{R}$ je pozitivno homogena funkcija od A.
Nadalje, definiramo funkciju

$$\sigma_m(A) = \sum_{i=1}^{m} \lambda_i,$$

koju možemo pisati i na sljedeći način

$$\begin{aligned} \sigma_m(A) &= \sup\{\sum_{i=1}^{m} Aq_i \cdot q_i : q_1, \ldots, q_m \text{ ortonormirani skup vektora u } \mathbb{R}^n\} \\ &= \sup\{\langle\!\langle UU^\tau, A \rangle\!\rangle : U \in \Omega\}, \end{aligned}$$

gdje je

$$\Omega := \{Q \in M_{n,m}(\mathbb{R}) : Q^\tau Q = I_m\},$$

skup matrica sastavljenih od m ortonormiranih n-stupaca.To je posljedica spektralnog teorema:

$$\begin{aligned} \sum_{i=1}^{m} Aq_i \cdot q_i &= \sum_{i=1}^{m} q_i^\tau A q_i \\ &= \sum_{i=1}^{m} (U^\tau q_i)^\tau D (U^\tau q_i) \\ &= \sum_{i=1}^{m} y_i^\tau D y_i, \end{aligned}$$

gdje su $y_i = U^\tau q_i$ ortonormirani. Dakle, $\sum_{i=1}^{m} Aq_i \cdot q_i \leq \sum_{i=1}^{m} \lambda_i \|y_i\|^2$, jednakost se postiže ako je y_i svojstveni vektor za λ_i, $i = 1, \ldots n$.
σ_m je konveksna funkcija jer je dobivena uzimanjem supremuma linearnih funkcija na S^n.
Dakle, λ_1 je konveksna funkcija, a λ_n je konkavna jer je dobivena razlikom linearne funkcije, $\sigma_n(A) = \mathrm{tr}A$, i konveksne funkcije, σ_{n-1}. Za svaki m, λ_m je lokalno Lipschitzova funkcija, jer je dobivena razlikom lokalno Lipschitzovih funkcija ($\lambda_m = \sigma_m - \sigma_{m-1}$).

Sada ćemo izračunati subdiferencijal najveće svojstvene vrijednosti simetrične matrice. Budući da je λ_1 konveksna funkcija možemo se poslužiti alatima konveksne analize. Funkciju λ_1 pišemo na sljedeći način

$$\lambda_1(A) = \max_{u^\tau u=1, u\in\mathbb{R}^n} u^\tau A u. \tag{3.1}$$

Računanje subdiferencijala po definiciji nije dobra ideja u ovom slučaju, pa želimo iskoristiti Teorem 2.2.12. Skup vektora $\{u \in \mathbb{R}^n : u^\tau u = 1\}$ u formuli (3.1) je kompaktan, funkcija $u \mapsto u^\tau A u$ neprekidna, a funkcija $A \mapsto u^\tau A u$ linearna, odnosno konveksna. Zadovoljeni su uvjeti Teorema 2.2.12 pa subdiferencijal funkcije λ_1 računamo pomoću formule (2.2). Potrebno je još samo izračunati subdiferencijal (u ovom slučaju gradijent) linearne funkcije $A \mapsto u^\tau A u$.

Zbog $u^\tau A u = \langle\langle uu^\tau, A\rangle\rangle$, očito je gradijent svake takve linearne funkcije dan sa jednočlanim skupom matrica ranga jedan,

$$\{uu^\tau\}.$$

Dakle, subdiferencijal[1] najveće svojstvene vrijednosti realne simetrične matrice A je

$$\partial\lambda_1(A) = \operatorname{conv}\{uu^\tau : u^\tau u = 1, Au = \lambda_1(A)u\}. \tag{3.2}$$

Odmah možemo izračunati i derivaciju u smjeru, prema Teoremu 1.2.1 (i) vrijedi

$$\lambda_1'(A, P) = \max\{\langle\langle uu^\tau, P\rangle\rangle : uu^\tau \in \partial\lambda_1(A)\}. \tag{3.3}$$

Iz (3.2) vidimo da je $\partial\lambda_1(A)$ jednočlan ako i samo ako je $\lambda_1(A)$ jednostruka svojstvena vrijednost. U tom slučaju subdiferencijal se svodi na gradijent funkcije λ_1 koji je jednak uu^τ, gdje je u svojstveni vektor pridružen svojstvenoj vrijednosti $\lambda_1(A)$.
Za najmanju svojstvenu vrijednost postupak je sličan, jer možemo promatrati negativnu funkciju početne funkcije.

Pogledajmo sada kako izgleda subdiferencijal funkcije λ_m za $1 < m < n$. Za fiksnu matricu $A \in S^n$ i fiksni indeks m definiramo dodatna dva indeksa, $\hat{m}$ i $\bar{m}$ na sljedeći način

$$\hat{m} = \min\{i : \lambda_i(A) = \lambda_m(A)\}, \quad \bar{m} = \max\{i : \lambda_i(A) = \lambda_m(A)\}.$$

Za ortogonalnu matricu U koja dijagonalizira matricu A

$$U^\tau A U = \operatorname{Diag}(\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)),$$

definiramo matricu U_m, koja je podmatrica od U, sastavljena od stupaca indeksiranih sa $\hat{m}, \hat{m}+1, \ldots, \bar{m}$.
Sljedeći teorem lako slijedi iz formula za derivacije u smjeru konveksnih funkcija σ_m i σ_{m-1} i daje nam jednostavnu formulu za računanje derivacije u smjeru za m-tu najveću svojstvenu vrijednost.

[1]Ako je $\partial_c f(x) = \partial^\circ f(x)$, onda subdiferencijal kraće označavamo sa $\partial f(x)$

Teorem 3.1.1. *Za svaku matricu $H \in S^n$ vrijedi*

$$\lambda'_m(A, H) = \lambda_{m-\hat{m}+1}(U_m^\tau H U_m). \tag{3.4}$$

Dokaz. Vidi [6]. □

Već smo vidjeli da subdiferencijal konveksne funkcije mora biti neprazan, a sada ćemo pokazati da za nekonveksnu funkciju subdiferencijal može biti i prazan skup.

Propozicija 3.1.2. *Ako je $1 < m < n$, tada funkcija λ_m nema linearne gornje ili donje ograde.*

Dokaz. Pretpostavimo suprotno, tj. da neka matrica $C \in S^n$ zadovoljava

$$\langle\langle C, G \rangle\rangle \leq \lambda_m(G), \tag{3.5}$$

za sve $G \in S^n$.
Neka je matrica U ortogonalna matrica koja dijagonalizira matricu C:

$$U^\tau C U = \mathrm{Diag}(c_1, c_2, \ldots, c_n),$$

za neki vektor c u $\mathbb{R}^n$.
Sada, za matricu $G := U(\mathrm{Diag}\ x)U^\tau$, gdje je x proizvoljan vektor iz $\mathbb{R}^n$, po pretpostavci vrijedi

$$c^\tau x \leq \lambda_m(\mathrm{Diag}\ x).$$

Ako za x odaberemo vektore standardne kanonske baze e_i, dobivamo da je $c \leq 0$ za svaki i, ali ako uzmemo $x = (-1, -1, \ldots, -1)$, dobivamo $\sum_i c \geq 1$ što vodi na kontradikciju. □

Dakle, subdiferencijal

$$\partial\lambda_m(A) = \{C \in S^n : \langle\langle C, G - A \rangle\rangle \leq \lambda_m(G) - \lambda_m(A),\ G \in S^n\} \tag{3.6}$$

je prazan skup za $A = 0$, $1 < m < n$, a slično se može pokazati i za proizvoljnu matricu $A \in S^n$. To nas motivira da promatramo drugačije derivacije u smjeru, kao na primjer Clarkeovu, definiranu u prethodnom poglavlju. Još jedna generalizirana derivacija u smjeru je Michel-Penotova derivacija u smjeru koju definiramo u idućem potpoglavlju.

3.2 Michel-Penotova derivacija u smjeru

U ovom poglavlju definiramo još jednu generaliziranu derivaciju u smjeru i generalizirani subdiferencijal koje možemo smatrati alternativom Clarkeovoj derivaciji i Clarkeovom subdiferencijalu.

Definicija 3.2.1 (Michel-Penotova derivacija u smjeru)**.** *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzova funkcija u $x \in \mathbb{R}^n$.* Michel-Penotova derivacija funkcije f u smjeru $v \in \mathbb{R}^n$ *definirana je sa*

$$f^{\diamond}(x, v) = \sup_{y \in R^n} \limsup_{t \to 0+} \frac{f(x + ty + tv) - f(x + ty)}{t}.$$

Primjer 3.2.2. *Izračunajmo derivacije u smjeru funkcije $f(x) = |x| - |\sin x|$ u točki $x = \pi$, u kojoj f nije diferencijabilna. Prvo, obična derivacija u smjeru*

$$\begin{aligned} f'(\pi, v) &= \lim_{t \to 0+} \frac{f(\pi + tv) - f(\pi)}{t} \\ &= \lim_{t \to 0+} \frac{tv - |\sin(\pi + tv)|}{t} \\ &= \lim_{t \to 0+} \frac{tv - |\sin(tv)|}{t} \end{aligned}$$

za $v < 0$ i $s := -v$ vrijedi

$$= v + v \lim_{st \to 0+} \frac{\sin st}{st}$$

a za $v \geq 0$ imamo

$$= v - v \lim_{st \to 0+} \frac{\sin st}{st}.$$

Zbog $\lim_{x \to 0} \frac{\sin x}{x} = 1$, vrijedi

$$f'(\pi, v) = \begin{cases} 2v, & v < 0, \\ 0 & v \geq 0. \end{cases}$$

Zatim računamo Clarkeovu i Michel-Penotovu derivaciju

$$\begin{aligned} f^{\circ}(\pi, v) &= \limsup_{y \to \pi, t \to 0+} \frac{f(y + tv) - f(y)}{t} \\ &= \limsup_{y \to \pi, t \to 0+} \frac{tv - |\sin(y + tv)| + |\sin y|}{t}. \end{aligned}$$

Za $v \geq 0$ i $y \leq \pi$ vrijedi

$$
\begin{aligned}
f^{\circ}(\pi, v) &= \limsup_{y\to\pi, t\to 0+} \frac{tv - \sin(y + tv) + \sin y}{t} \\
&= \limsup_{y\to\pi, t\to 0+} \frac{tv - \sin y(1 - \cos tv) - \sin tv \cos y}{t} \\
&= v - v\cos\pi \lim_{tv\to 0+} \frac{\sin tv}{tv} \\
&= 2v,
\end{aligned}
$$

dok za $v < 0$, $s := -v$ i $y \geq \pi$ vrijedi

$$
\begin{aligned}
f^{\circ}(\pi, v) &= \limsup_{y\to\pi, t\to 0+} \frac{tv + \sin(y + tv) - \sin y}{t} \\
&= \limsup_{y\to\pi, t\to 0+} \frac{tv + \sin y(\cos tv - 1) + \sin tv \cos y}{t} \\
&= v + v\cos\pi \lim_{ts\to 0+} \frac{\sin ts}{ts} \\
&= 0.
\end{aligned}
$$

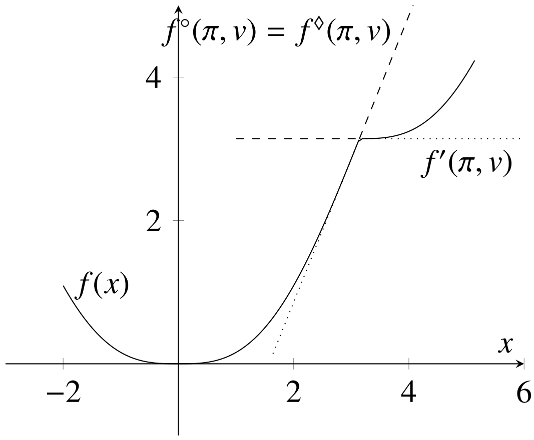
Stoga, vrijedi

$$
f^{\circ}(\pi, v) = \begin{cases} 0 & za \quad v < 0, \\ 2v & za \quad v \geq 0. \end{cases}
$$

Na sličan način izračunamo i

$$
f^{\diamond}(\pi, v) = \begin{cases} 0 & za \quad v < 0, \\ 2v & za \quad v \geq 0. \end{cases}
$$

Na slici 3.1 možemo vidjeti da je obična derivacija u smjeru puno bolja aproksimacija funkcije f u točki π, ali subdiferencijal je tada prazan skup jer obična derivacija u smjeru nije sublinearna funkcija.



Slika 3.1: Graf funkcije iz Primjera 3.2.2.

Teorem 3.2.3 (Svojstva Michel-Penotove derivacije u smjeru)**.** *Neka je f lokalno Lipschitzova s konstantom K. Tada je*
(i) $f^\diamond(x, v)$ sublinearna, Lipschitzova s konstantom K i vrijedi

$$f'(x, v) \leq f^\diamond(x, v) \leq f^\circ(x, v), v \in \mathbb{R}^n,$$

(ii) $f^\diamond(x, -v) = (-f)^\diamond(x, v)$.

Dokaz. (i) Prva nejednakost lako slijedi iz definicije Michel-Penotove derivacije za $y = v$ i iz svojstva limes superiora.
Dokažimo sada drugu nejednakost. Za fiksni $v \in \mathbb{R}^n$ neka je $\epsilon > 0$ zadan. Tada za svaki $y \in \mathbb{R}^n$ postoji $\delta(y) > 0$ tako da zbog svojstva limes superiora vrijedi (za $y_1 := x + ty$ u definiciji Clarkeove derivacaije, gdje je y_1 zamjena za oznaku y u Clarkeovoj derivaciji)

$$\frac{f(x + ty + tv) - f(x + ty)}{t} < f^\circ(x, v) + \epsilon,\ \ t \in \langle 0, \delta(y) \rangle.$$

Uzimanjem limes superiora sa obje strane dobivamo

$$\limsup_{t \to 0+} \frac{f(x + ty + tv) - f(x + ty)}{t} \leq f^\circ(x, v) + \epsilon.$$

i to vrijedi za svaki $y \in \mathbb{R}^n$, iz čega zaključujemo da vrijedi

$$f^\diamond(x, v) \leq f^\circ(x, v) + \epsilon.$$

Puštanjem $\epsilon \downarrow 0$ slijedi tražena nejednakost.
Za dokaz sublinearnosti dovoljno je provjeriti subaditivnost budući da se pozitivna homogenost dokaže isto kao i za Clarkeovu derivaciju.

Neka su $v_1, v_2 \in \mathbb{R}^n$ i $\epsilon > 0$ zadani. Tada za t dovoljno mali vrijedi

$$\frac{f(x + t(v_1 + v_2) + ty) - f(x + t(v_2 + y))}{t} \leq f^\diamond(x, v_1) + \frac{\epsilon}{2},\ y \in \mathbb{R}^n$$

i

$$\frac{f(x + tv_2 + ty) - f(x + ty)}{t} \leq f^\diamond(x, v_2) + \frac{\epsilon}{2},\ y \in \mathbb{R}^n,$$

pa zbrajanjem tih nejednakosti dobivamo

$$\frac{f(x + t(v_1 + v_2) + ty) - f(x + ty)}{t} \leq f^\diamond(x, v_1) + f^\diamond(x, v_2) + \epsilon,\ y \in \mathbb{R}^n,$$

iz čega sada, uzimanjem limes superiora i supremuma i puštanjem $\epsilon \downarrow 0$, slijedi

$$f^\diamond(x, v_1 + v_2) \leq f^\diamond(x, v_1) + f^\diamond(x, v_2).$$

Pokažimo da je $f^\diamond(x, \cdot)$ lokalno Lipschitzova; neka su $v_1, v_2 \in \mathbb{R}^n$ proizvoljni i $t > 0$ dovoljno mali, tada za svaki $y \in \mathbb{R}^n$ vrijedi

$$\begin{aligned} f(x + ty + tv_1) - f(x + ty) &= [f(x + ty + tv_2) - f(x + ty)] + [f(x + ty + tv_1) - f(x + ty + tv_2)] \\ &\leq [f(x + ty + tv_2) - f(x + ty)] + tK\|v_1 - v_2\| \end{aligned}$$

pa prema tome vrijedi

$$f^\diamond(x, v_1) \leq f^\diamond(x, v_2) + K\|v_1 - v_2\|$$

Zamjenom uloga od v_1 i v_2 dobivamo

$$f^\diamond(x, v_2) \leq f^\diamond(x, v_1) + K\|v_1 - v_2\|$$

pa zaključujemo da vrijedi

$$|f^\diamond(x, v_1) - f^\diamond(x, v_2)| \leq K\|v_1 - v_2\|.$$

(ii) Za svaki $y \in \mathbb{R}^n$ vrijedi

$$\begin{aligned} &\limsup_{t \to 0+} \frac{f(x - tv + ty) - f(x + ty)}{t} \\ &= \limsup_{t \to 0+} \frac{(-f)(x + tv + t(y - v)) - (-f)(x + t(y - v))}{t}. \end{aligned}$$

Uzimanjem supremuma po y i $y - v$ slijedi tvrdnja. □

Analogno kao i za Clarkeovu derivaciju definiramo Michel-Penotov subdiferencijal na sličan način.

Definicija 3.2.4 (Michel-Penotov subdiferencijal)**.** *Neka je $f : \mathbb{R}^n \to \mathbb{R}$ lokalno Lipschitzova funkcija u $x \in \mathbb{R}^n$. Tada je* Michel-Pentov subdiferencijal funkcije f u točki x *sljedeći skup*

$$\partial^\Diamond f(x) = \{\xi \in \mathbb{R}^n : f^\Diamond(x,v) \geq \xi \cdot v, v \in \mathbb{R}^n\}.$$

U sljedećem teoremu sažeta su neka od svojstava Michel-Penotovog subdiferencijala.

Teorem 3.2.5 (Svojstva Michel-Penotovog subdiferencijala)**.** *Neka je f lokalno Lipschitzova funkcija u x s konstantom K. Tada*
(i) $\partial^\Diamond f(x)$ je neprazan, konveksan i kompaktan skup takav da je

$$\partial^\Diamond f(x) \subset \partial^\circ f(x)$$

(ii) $f^\Diamond(x,v) = max\{\xi \cdot v : \xi \in \partial^\Diamond f(x)\},\ v \in \mathbb{R}^n$,
(iii) ako je f diferencijabilna u x onda vrijedi $\partial^\Diamond f(x) = \{\nabla f(x)\}$.

Dokaz. (i) Da je $\partial^\Diamond f(x)$ neprazan, konveksan i kompaktan pokaže se slično kao i za $\partial^\circ f(x)$, a $\partial^\Diamond f(x) \subset \partial(x)$ vrijedi zbog $f^\Diamond(x,\cdot) \leq f^\circ(x,\cdot)$.
(ii) Po definiciji od $f^\Diamond$, slično kao i za f°.
(iii) Za svaki $y, v \in \mathbb{R}^n$ vrijedi

$$\begin{aligned}\limsup_{t\to 0+} \frac{f(x+tv+ty) - f(x+ty)}{t} &\leq \limsup_{t\to 0+} \frac{f(x+t(x+y)) - f(x)}{t} \\ &+ \limsup_{t\to 0+} \frac{f(x) - f(x+ty)}{t} \\ &= f'(x, v+y) - f'(x,y) \\ &= \nabla f(x) \cdot (v+y) - \nabla f(x) \cdot y \\ &= f'(x,v) \leq f^\Diamond(x,v).\end{aligned}$$

Uzimanjem supremuma po svim $y \in \mathbb{R}^n$ slijedi tražena tvrdnja. □

Po Teoremu 3.2.3.(i) vidimo da su Clarkeova i Michel-Penotova derivacija sublinearne gornje ograde za običnu derivaciju u smjeru. Općenito, za svaku pozitivno homogenu, Lipschitzovu funkciju $p : \mathbb{R}^n \to \mathbb{R}$, možemo definirati gornju ogradu, za svaki smjer $v \in \mathbb{R}^n$, na sljedeći način

$$^*p(v) = \sup_{y\in\mathbb{R}^n}\{p(y+v) - p(y)\}. \tag{3.7}$$

Očito je $^*p$ sublinearna i najmanja od svih funkcija h koje zadovoljavaju

$$p(y+v) \leq p(y) + h(v), \quad y, v \in \mathbb{R}^n.$$

Ako je f pozitivno homogena funkcija lako se vidi iz definicije Clarkeove i Michel-Penotove derivacije u smjeru da vrijedi

$$f^\circ(0,v) = f^\diamond(0,v) = \sup_{y\in\mathbb{R}^n}\{f(y+v) - f(y)\}, \quad v \in \mathbb{R}^n. \tag{3.8}$$

Dakle, ${}^*p$ nije ništa drugo nego Clarkeova ili Michel-Penotova derivacija u smjeru funkcije p u ishodištu.
U slučaju da postoji derivacija u smjeru funkcije f, imamo

$$\begin{aligned}\limsup_{t\to 0+}\frac{f(x+tv+ty)-f(x+ty)}{t} &\le \limsup_{t\to 0+}\frac{f(x+tv+ty)-f(x)}{t}\\ &+\limsup_{t\to 0+}\frac{f(x)-f(x+ty)}{t}\\ &= f'(x,v+y)-f'(x,y).\end{aligned}$$

Uzimanjem supremuma po $y \in \mathbb{R}^n$ slijedi

$$f^\diamond(x,v) \le \sup_{y\in\mathbb{R}^n}\{f'(x,v+y) - f'(x,y)\}.$$

Za dokaz suprotne nejednakosti, neka su $v \in \mathbb{R}^n$ i $\epsilon > 0$ zadani, tada po svojstvu limes superiora, za svaki $y \in \mathbb{R}^n$ i $t > 0$ dovoljno mali, vrijedi

$$\frac{f(x+tv+ty)-f(x+ty)}{t} < f^\diamond(x,v) + \epsilon.$$

Sada, budući da limes postoji, vrijedi

$$\lim_{t\to 0+}\left(\frac{f(x+tv+ty)-f(x)}{t} + \frac{f(x)-f(x+ty)}{t}\right) \le f^\diamond(x,v) + \epsilon$$

Za $\epsilon \downarrow 0$ slijedi tvrdnja. Dakle, dokazali smo da ukoliko derivacija u smjeru postoji, Michel-Penotovu derivaciju funkcije f u smjeru v možemo računati na sljedeći način

$$f^\diamond(x,v) = \sup_{y\in\mathbb{R}^n}\{f'(x,v+y) - f'(x,y)\}. \tag{3.9}$$

Prema tome, Michel-Penotova derivacija u smjeru je najmanja sublinearna gornja ograda obične derivacije u smjeru

$$f^\diamond(x,\cdot) = {}^*(f'(x,\cdot)).$$

Ova formula će biti korisna za računanje sa funkcijama λ_m budući da su one pozitivno homogene i njihova derivacija u smjeru uvijek postoji. U suprotnom, Michel-Penotovu derivaciju morali bi računati direktno iz definicije, što često nije praktično.

3.3 Clarkeov i Michel-Penotov subdiferencijal svojstvene vrijednosti simetrične matrice

U ovom poglavlju želimo pokazati da se Clarkeov i Michel-Penotov subdiferencijal funkcije λ_m podudaraju za svaku matricu $A \in S^n$. U nastavku, subdiferencijal konveksne funkcije f u točki x kraće označavamo sa $\partial f(x)$. Slično kao i za Michel-Penotovu derivaciju, ali direktnije po definiciji, izvodi se formula za Clarkeovu derivaciju

$$f^\circ(x, v) = \limsup_{y \to x} f'(y, v), \tag{3.10}$$

koja također vrijedi samo ako derivacija u smjeru $v \in \mathbb{R}^n$ funkcije f postoji na nekom skupu Ω. Prisjetimo se sada formule za subdiferencijal funkcije λ_1

$$\partial\lambda_1(A) = conv\{uu^\tau : u^\tau u = 1, Au = \lambda_1(A)u\},$$

iz koje se lako vidi da vrijedi

$$\partial\lambda_1(0) = \{M \in S^n : M \text{ pozitivno semidefinitna, tr} \, M{=}1\}.$$

Prema tome, idući teorem nam daje laku i jasnu interpretaciju generaliziranih subdiferencijala u nuli.

Teorem 3.3.1 (Subdiferencijali u nuli)**.** *Clarkeov i Michel-Penotov subdiferencijali funkcije λ_m se podudaraju u nuli za svaki $m = 1, \ldots, n$:*

$$\partial^\diamond \lambda_m(0) = \partial^\circ \lambda_m(0) = \partial\lambda_1(0).$$

Dokaz. Prva jednakost slijedi zbog pozitivne homogenosti funkcije λ_m, odnosno iz (3.8). Za dokaz druge jednakosti potrebno je, za svaku matricu $G \in S^n$, dokazati sljedeću jednakost

$$\lambda_m^\diamond(0, G) = \lambda_1'(0, G),$$

ili, koristeći (3.8), imamo

$$\sup_{H \in S^n} \{\lambda_m(H + G) - \lambda_m(H)\} = \lambda_1(G). \tag{3.11}$$

Po Weylovoj nejednakosti ([7], Teorem 4.3.7) vrijedi

$$\lambda_m(H + G) \leq \lambda_m(H) + \lambda_1(G),$$

a da bi dokazali da se supremum zaista i postiže uzmimo ortogonalnu matricu U za koju vrijedi $U^\tau GU = \text{Diag}\,\lambda(G)$ i bilo koji $\alpha \in \mathbb{R}$ za koji vrijedi $\alpha > \lambda_1(G) - \lambda_n(G)$.
Sada definiramo matricu H na sljedeći način

$$H = U\text{Diag}\,(0, 0, \ldots, 0, \underbrace{\alpha, \alpha, \ldots, \alpha}_{m-1 \text{ članova}})U^\tau.$$

Budući da je $\lambda_m(H) = 0$, vrijedi

$$\lambda_m(H + G) - \lambda_m(H) = \lambda_m(U^\tau(H + G)U) = \lambda_m(\operatorname{Diag} d),$$

gdje je $d \in \mathbb{R}^n$ sljedeći vektor

$$(\lambda_1(G), \lambda_2(G), \ldots, \lambda_{n-m+1}(G), \lambda_{n-m+2}(G) + \alpha, \lambda_{n-m+3}(G) + \alpha, \ldots, \lambda_n(G) + \alpha).$$

Zbog svojstva od α, vrijedi $\lambda_m(\operatorname{Diag} d) = \lambda_1(G)$ i time je teorem dokazan. □

Kao što vidimo iz prethodnog teorema, jednakost ovih subdiferencijala u nuli jednostavno slijedi iz pozitivne homogenosti funkcije λ_m.

Subdiferencijali za proizvoljne realne simetrične matrice

Da bi jednakost pokazali za proizvoljnu matricu $A \in S^n$, potrebno je slijediti nekoliko jednostavnih koraka. Prvo, koristeći formulu (3.4) iz Teorema 3.1.1, računamo običnu derivaciju u smjeru, $\lambda'_m(A, \cdot)$. Nakon toga, pomoću forumle (3.9) računamo najmanju sublinearnu gornju ogradu funkcije $\lambda'_m(A, \cdot)$ što će upravo biti Michel-Penotova derivacija u smjeru. Na kraju, regulariziramo funkciju $\lambda'_m(A, \cdot)$ pomoću formule (3.10) da bi dobili Clarkeovu derivaciju.

Uočimo da je $\lambda_m(A)$ samo dio bloka jednakih svojstvenih vrijednosti. Ta činjenica nam "kvari" običnu derivaciju u smjeru jer ona tada nije sublinearna i pogodna za računanje. Vodeća svojstvena vrijednost u tom bloku je $\lambda_{\hat{m}}$ i njena derivacija u smjeru je, po Teoremu 3.1.1, upravo λ_1. Kako je ta funkcija sublinearna, od koristi će nam biti sljedeća propozicija.

Propozicija 3.3.2. *Za konačnu sublinearnu funkciju $v \mapsto \sigma(v) := f'(x, v)$ vrijedi*

$$\sigma'(0, \delta) = f'(x, \delta), \quad \delta \in \mathbb{R}^n.$$

Dokaz. Budući da je σ pozitivno homogena funkcija vrijedi $\sigma(0) = 0$, i imamo

$$\sigma'(0, \delta) = \lim_{t \to 0+} \frac{\sigma(t\delta) - \sigma(0)}{t} = \sigma(\delta) = f'(x, \delta).$$

□

Sljedeći teorem pokazuje da za Michel-Penotov subdiferencijal u fiksnoj matrici $A \in S^n$ nije bitno da li promatramo prvu, ili bilo koju drugu svojstvenu vrijednost u bloku jednakih, jer je za sve njih, Michel-Penotov subdiferencijal jednak.

Teorem 3.3.3. *Michel-Penotovi subdiferencijali funkcije λ_m u matrici $A \in S^n$ podudaraju se za svaki m za koji su svojstvene vrijednosti jednake:*

$$\partial^{\diamond}\lambda_m(A) = \partial(\lambda'_{\hat{m}}(A,\cdot))(0), \quad m = 1,\ldots,n.$$

Dokaz. Ako definiramo funkciju $\sigma(\cdot) := \lambda'_{\hat{m}}(A,\cdot)$, tada po prethodnoj propoziciji vrijedi $\sigma'(0,G) = \lambda'_{\hat{m}}(A,G)$, za svaku matricu $G \in S^n$. Prema tome, za dokaz teorema potrebno je pokazati da za svaku matricu $G \in S^n$ vrijedi

$$\lambda^{\diamond}_m(A,G) = \lambda'_{\hat{m}}(A,G).$$

Prema (3.9), gornja jednakost može se pisati na sljedeći način

$$\sup_{H\in S^n}\{\lambda'_m(A,H+G) - \lambda'_m(A,H)\} = \lambda'_{\hat{m}}(A,G).$$

Nadalje, uzmimo ortogonalnu matricu U koja dijagonalizira matricu A. Tada, po Teoremu 3.1.1, gornja jednakost ekvivalentna je s

$$\sup_{H\in S^n}\{\lambda_{m-\hat{m}+1}(U_m^\tau(H+G)U_m) - \lambda_{m-\hat{m}+1}(U_m^\tau H U_m)\} = \lambda_1(U_m^\tau G U_m).$$

Za svaku matricu $F \in S^{\bar{m}-\hat{m}+1}$ vrijedi

$$U_m^\tau(U_m F U_m^\tau)U_m = F,$$

pa prema tome, uz pogodnu supstituciju, imamo

$$\sup_{F\in S^{\bar{m}-\hat{m}+1}}\{\lambda_{m-\hat{m}+1}(F + U_m^\tau G U_m) - \lambda_{m-\hat{m}+1}(F)\} = \lambda_1(U_m^\tau G U_m),$$

pa tvrdnja sada slijedi po dokazu Teorema 3.3.1., odnosno (3.11). □

Sljedeći rezultat nam daje jasniju sliku o tome kako izgleda Michel-Penotov subdiferencijal funkcije λ_m. Radi preglednijeg zapisa, sa $E_m(A) \subset \mathbb{R}^n$ označimo svojstveni potprostor m-te najveće svojstvene vrijednosti matrice A.

Propozicija 3.3.4. *Za svaku matricu $A \in S^n$ vrijedi*

$$\partial(\lambda'_{\hat{m}}(A,\cdot))(0) = \operatorname{conv}\{xx^\tau : x \in E_m(A), \|x\| = 1\}.$$

Dokaz. Potrebno je pokazati da za svaku matricu $G \in S^n$ vrijedi

$$\lambda'_{\hat{m}}(A,G) = \max\{\langle\langle G, xx^\tau\rangle\rangle : x \in E_m(A), \|x\| = 1\},$$

što je ekvivalentno s

$$\lambda_1(U_m^\tau G U_m) = \max\{x^\tau G x : x \in E_m(A), \|x\| = 1\}.$$

Primijetimo da stupci matrice U_m čine ortonormiranu bazu potprostora $E_m(A)$, pa vektor x s desne strane gornje jednadžbe možemo pisati na sljedeći način

$$x = U_m y,\ y \in \mathbb{R}^{\hat{m}-\bar{m}+1}, \|y\| = 1,$$

pa stoga vrijedi

$$\lambda_1(U_m^\tau G U_m) = \max\{y^\tau (U_m^\tau G U_m) y : y \in \mathbb{R}^{\hat{m}-\bar{m}+1}, \|y\| = 1\}.$$

Zadnja jednakost očito je istinita, pa prema tome propozicija je dokazana. □

Sada smo spremni za glavni rezultat ovog poglavlja, naime, već smo vidjeli u Primjeru 3.2.2. da se Clarkeova i Michel-Penotova derivacija podudaraju u točki π za zadanu funkciju. Takvo nešto ne mora vrijediti za sve funkcije, ali vrijedi za funkciju λ_m:

Teorem 3.3.5. *Clarkeovi i Michel-Penotovi subdiferencijali funkcije λ_m u $A \in S^n$ su jednaki:*

$$\partial^\diamond \lambda_m(A, G) = \partial^\circ \lambda_m(A) = \operatorname{conv}\{x x^\tau : x \in E_m(A), \|x\| = 1\}, \quad m = 1, \ldots, n. \tag{3.12}$$

Dokaz. Uočimo da je potrebno dokazati samo inkluziju

$$\partial^\circ \lambda_m(A) \subset \partial^\diamond \lambda_m(A),$$

jer za obratnu inkluziju već znamo da vrijedi, a druga jednakost tada slijedi po Teoremu 3.3.3. i Propoziciji 3.3.4.
Uzmimo proizvoljnu, ali fiksnu, matricu $G \in S^n$ i indeks m. Po 3.10, postoji niz simetričnih matrica $A_k \to A$ tako da vrijedi:

$$\partial^\circ \lambda_m(A) = \lim_{k\to\infty} \lambda_m'(A_k, G).$$

Uzimanjem podniza možemo bez smanjenja općenitosti pretpostaviti da skup indeksa

$$I := \{i : \lambda_i(A_k) = \lambda_m(A_k)\}$$

ne ovisi o indeksu k, a budući da su funkcije λ_i neprekidne mora vrijediti

$$I \subset \{\hat{m}, \hat{m}+1, \ldots, \bar{m}\}.$$

Sada, za svaki k možemo odabrati ortonormiranu matricu U^k koja dijagonalizira matricu A^k, takvu da $U^k \to U$ budući da je λ_m neprekidna, a matrica U dijagonalizira matricu A. Označimo sa U_I i U_I^k redom podmatrice od U i U^k indeksirane elementima, gore definiranog, skupa I. Po Teoremu 3.1.1. vrijedi nejednakost:

$$\lambda'_m(A_k, G) \leq \lambda_1((U_I^k)^\tau G U_I^k) \to \lambda_1(U_I^\tau G U_I).$$

Neka je U_m matrica čiji su stupci indeksirani sa $\{\hat{m}, \hat{m}+1, \dots, \bar{m}\}$, pa vrijedi

$$\begin{aligned}\lambda_1(U_I^\tau G U_I) &= \max_{\|z\|=1}(U_I z)^\tau G(U_I z)\\ &\leq \max_{\|y\|=1}(U_m y)^\tau G(U_m y)\\ &= \lambda_1(U_m^\tau G U_m)\\ &= \lambda'_{\hat{m}}(A, G),\end{aligned}$$

što znači da vrijedi

$$\lambda^\circ_m(A, G) \leq \lambda'_{\hat{m}}(A, G),$$

odnosno, budući da je matrica G bila proizvoljna, imamo

$$\partial^\circ \lambda_m(A) \subset \partial(\lambda'_{\hat{m}}(A, \cdot))(0).$$

Prema Teoremu 3.3.3. slijedi tražena inkluzija i time je teorem dokazan. □

Usporedba s drugim rezultatima

Imajući na umu prethodni teorem, po formuli 2.1, Clarkeovu derivaciju u smjeru možemo pisati na sljedeći način

$$\lambda^\circ_m(A, G) = \max\{x^\tau G x : x \in E_m(A), \|x\| = 1\},$$

rezultat koji su prvi promatrali Cox i Overton ([5]). Alternativni dokaz može ići, na primjer, preko gotovo svuda diferencijabilnosti funkcije λ_m.
Cox [8] je koristio neprekidno diferencijabilne funkcije $F : \mathbb{R}^k \to S^n$. Tada promatramo kompoziciju $\lambda_m \circ F$ i želimo izračunati Clarkeov subdiferencijal. Prilično kompliciran argument daje egzaktnu formulu u slučaju da je svojstvena vrijednost najveća ili najmanja u bloku jednakih, a u suprotnom vrijedi samo inkluzija. Za $p, d \in \mathbb{R}^n$ i $A = F(p)$ rezultat glasi

$$\begin{aligned}(\lambda_{\hat{m}} \circ F)^\circ(p, d) &= (\lambda_{\bar{m}})(p, d)\\ &= \max\{x^\tau(F'(p)d)x : x \in E_m(A), \|x\| = 1\}\\ &\leq (\lambda_m \circ F)^\circ(p, d).\end{aligned}$$

Ovaj rezultat sada slijedi direktno iz formule 3.12 koristeći Teorem 2.2.8. (Lančano pravilo). Budući da su $\lambda_{\hat{m}}$ i $-\lambda_{\bar{m}}$ regularne za njih vrijedi jednakost, za općenite svojstvene vrijednosti samo inkluzija. Za Michel-Penotovu derivaciju vrijedi isti argument, uz napomenu da je za funkciju F dovoljno pretpostaviti da je diferencijabilna.

Bibliografija

[1] J. M. Borwein, A. S. Lewis, *Convex Analysis and Nonlinear Optimization*, Springer, 2006.

[2] F. H. Clarke, *Optimization and Nonsmooth Analysis*, SIAM, 1990.

[3] B. Guljaš, *Matematička analiza I i II* (web izdanje), Zagreb, 2014.

[4] J.-B. Hiriart-Urruty, C. Lemarèchal *Fundamentals of Convex Analysis*, Springer, 2004.

[5] J.-B. Hiriart-Urruty, A.S. Lewis, *The Clarke and Michel-Penot Subdifferentials of the Eigenvalues of a Symmetric Matrix*, Kluwer Academic Publishers, (1999), 13–23

[6] J.-B. Hiriart-Urruty, D. Ye, *Sensitivity analysis of all eigenvalues of a symmetric matrix*, Numer. Math. 70, (1995), 45–72

[7] R. A. Horn, C. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[8] A. S. Lewis, *Derivatives of spectral functions*, Mathematics of Operations Research 6, 1996. 576–588

[9] M. M. Mäkelä, P. Neittaanmäki, *Nonsmooth optimization*, World Scientific, Singapore, 1992.

[10] W. Schirotzek, *Nonsmooth Analysis*, Springer, 2007.

Sažetak

U ovom radu proširen je pojam diferencijabilnosti lokalno Lipschitzovih funkcija sa $\mathbb{R}^n$ u $\mathbb{R}$. Prvo promatramo derivaciju u smjeru poznatu iz klasične analize i navodimo neka njena svojstva, kao što su na primjer, ograničenost, Lipschitzovost i sublinearnost. Nakon toga definiramo subdiferencijal konveksne funkcije u točki $x \in \mathbb{R}^n$. Taj skup je neprazan, konveksan i kompaktan, a u slučaju da funkcija nije konveksna, može biti i prazan.
U nastavku promatramo općenite lokalno Lipschitzove, ne nužno konveksne funkcije i definiramo generaliziranu Clarkeovu derivaciju u smjeru i Clarkeov subdiferencijal. Navodimo neka njihova svojstva i uspoređujemo s običnom derivacijom u smjeru. Kao i u klasičnoj analizi za diferencijabilne funkcije, i za generaliziranu derivaciju u smjeru postoje pravila računanja za linearnu kombinaciju, produkt, kvocijent i kompoziciju. No, u općenitom slučaju vrijedi samo inkluzija među pripadnim subdiferencijalima. Zatim navodimo teorem o subdiferencijalu max funkcije i generalizirani teorem srednje vrijednosti.
Koristeći rezultate iz prethodnih poglavlja, računamo subdiferencijal i derivaciju najveće svojstvene vrijednosti realne simetrične matrice. Zatim uvodimo Michel-Penotovu derivaciju u smjeru i pripadni Michel-Penotov subdiferencijal i pomoću nekoliko rezultata dolazimo do zaključka da se Clarkeov i Michel-Penotov subdiferencijal podudaraju za funkciju m-te najveće svojstvene vrijednosti realne simetrične matrice.

Summary

In this master thesis we study the extended notion of diferentiability of real locally Lipschitz functions defined on space $\mathbb{R}^n$. In the first part we study the well known directional derivative of convex functions and its properties, some of which are, for instance, boundness, locally Lipschitz continuity and sublinearity. After that, the subdifferential of convex function in $x \in \mathbb{R}^n$ is defined, and we show that it is a nonempty, convex and compact set, which can be empty if the function f is not convex. Next, we study the Clarke's directional derivative and Clarke's subdifferential. We list some of their properties and compare them to the usual directional derivative.

Simlarly to clasical analysis, there are calculus rules in nonsmooth analysis for Clarke's subdifferentials, such as for linear combination of functions, product, quotient and chain rule, but in the case of Clarke subdifferential, only one inclusion can be showed. We also generalize some important results, such as mean value theorem and max-function theorem. In the last section we calculate the subdiferential and directional derivative of the largest eigenvalue of the real symmetric matrix with the help of the results from the previous chapters. Another generalized derivative is defined, called the Michel-Penot's directional derivative. After defining the Michel-Penot's subdifferential we show that Clarke's and Michel-Penot's subdifferentials coincide for mth largest eigenvalue of a real symmetric matrix.

Životopis

Rođen sam 15. travnja 1990. godine u Zagrebu. Godine 1997. upisujem Osnovnu školu Josipa Badalića u Graberju Ivanićkom, nedaleko od Ivanić Grada, a ubrzo i tečaj engleskog jezika i treniram nogomet u lokalnom klubu. U višim razredima osnovne škole pohađam natjecanja iz matematike i fizike, pri čemu ostvarujem izravan upis u srednju školu 2005. godine, te upisujem Petu gimnaziju u Zagrebu. Srednju školu završavam 2009. godine s odličnim uspjehom i oslobođen sam mature. Iste godine upisujem preddiplomski studij matematike na Prirodoslovno-matematičkom fakultetu u Zagrebu. Nakon završenog preddiplomskog studija, 2012. godine upisujem diplomski studij Financijske i poslovne matematike na istom fakultetu.