UNIVERZA NA PRIMORSKEM
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN INFORMACIJSKE TEHNOLOGIJE

Zaključna naloga

Konveksne množice in konveksne funkcije

(Convex sets and convex functions)

Ime in priimek: Bećo Merulić
Študijski program: Matematika
Mentor: prof. dr. Bojan Kuzma

Koper, oktober 2015

Ključna dokumentacijska informacija

Ime in PRIIMEK: Bećo MERULIĆ

Naslov zaključne naloge: Konveksne množice in konveksne funkcije

Kraj: Koper

Leto: 2015

Število listov: 61
Število referenc: 15

Mentor: prof. dr. Bojan Kuzma

Ključne besede: afine množice, konveksne množice, epigraf, zaprtje, relativno odprtje, efektivna domena, Jensenova neenakost, konveksne funkcije, zveznost konveksnih funkcij, subgradient, gradient, diferenciabilnost konveksnih funkcij

Math. Subj. Class. (2010): 26A51, 26B25,

Izvleček:
Poudarek te naloge je na zveznosti in diferenciabilnosti konveksnih funkcij. Najprej se spoznamo z definicijami afinih množic in afnih transformacij, potem pa z konveksnimi množicami in algebro konveksnih množic. Vzpostavimo kriterij za konveksnost. Definiramo epigraf konveksne funkcije in spoznamo da se konveksne funkcije na $\mathbb{R}^n$ lahko identificirajo z določenimi konveksnimi podmnožicami (epigrafi) v $\mathbb{R}^{n+1}$. Sprehajajoči se skozi topološko naravo konveksnih funkcij, spoznamo se z pomeno konceptov kot so zaprtje konveksne funkcije, relativna notranjost, relativna meja, efektivna domena, in z izjemno pomembnim konceptom navzdol polzveznosti, ki je v direktni zvezi z epigrafom in kot takšen bolj pomemben koncept kot običajna zveznost za konveksne funkcije. Ogledamo nekaj dejstev iz teorije dualnosti. Diferenciabilnost konveksnih funkcij je najbolj pomembno orodje pri reševanju problema ekstemov. Definiramo subgradient in gradient, ogledamo zvezo med danimi konceptimi, in dokažemo nekaj pomembnih izrekov iz teorije diferenciabilnosti, kot je recimo izrek ki pravi da za zaporedje končnih konveksnih funkcij diferenciabilnih na odprti konveksni množici C ki konvergirajo po točkah proti diferenciabilni funkciji f velja da zaporedje njunih odvodov konvergira enakomerno k f' na vsaki zaprti omejeni podmnožici množice C.

Key words documentation

Name and SURNAME: Bećo MERULIĆ

Title of final project paper: Convex sets and convex functions

Place: Koper

Year: 2015

Number of pages: 61
Number of references: 15

Mentor: Prof. Bojan Kuzma, PhD

Keywords: affine sets, convex sets, epigraph, closure, relative interior, effective domain, Jensen's inequality, convex functions, continuity of convex functions, subgradient, gradient, differentiability of convex functions

Math. Subj. Class. (2010): 26A51, 26B25

Abstract: The focus of this paper is on continuity and differentiability of convex functions. First, we get to know the definitions of affine sets and affine transformations, then those of convex sets and with algebra of convex sets. We establish a criterion for convexity. We define epigrapf of a convex functions and learn that the convex functions on $\mathbb{R}^n$ can be identified with certain convex subsets (epigraphs) in $\mathbb{R}^{n+1}$. Walking through the topological nature of the convex functions we become familiar with the importance of concepts such as closure of the convex function, the relative interior, the relative boundary, the effective domain, and of the very important concept of lower semi-continuity, which is in direct connection with the epigraph and as such is more important concept than normal continuity for convex functions. We learn some facts from the theory of duality. Differentiability of convex functions is the most important tool in tackling problems of finding extremes. We define subgradient and gradient, learn the connection between those two concepts, and prove some important theorems from the theory of differentiability, such as the theorem stating that for a sequence of finite convex functions, differentiable on an open convex set C, which converge point-wise to a differentiable functions f, the sequence of their derivatives converges uniformly to f' on each closed bounded subset of C.

Zahvala

Najprej se iskreno zahvaljujem svom mentorju dr. Bojanu Kuzmi za strokovno pomoč, usmerjanje in posredovanje bogatega znanja, ne le pri nastajanju te diplomske naloge, ampak v celem študijskem obdobju, brez izjem.

Zahvalo dolgujem tudi vsem asistentom in profesorjem zaposlenim na Fakulteti za matematiko, naravoslovje in informacijske tehnologije, ki so pripomogli k mojemu napredku v izobraževanju.

Največja zahvala pa vsekakor gre moji materi, ki mi je finančno in moralno stala ob strani skozi ves študij, za neskončno razumevanje in potrpežljivost, še posebej v trenutkih, ko sem bil v dvomih.

Kazalo vsebine

Seznam kratic

tj.	to je
npr.	na primer
oz.	oziroma

1 Uvod

"Zdi se mi, da je pojem konveksne funkcije prav tako pomemben, kot je pozitivna funkcija ali naraščajoča funkcija. Če se ne motim pri tem, pojem bi moral najti svoje mesto v osnovnimi ekspozicijami teorije realnih funkcij."
(J. L. W. V. Jensen)

Konveksne funkcije so zelo pomembne v številnih matematičnih področjih. Še posebej so pomembne v uporabni matematiki, pri študiranju optimizacijskih problemov, ker jih odlikujejo številne priročne lastnosti v zvezi z ekstremimi. Npr. (strogo) konveksna funkcija na zaprti množici ne more imeti več kot enega minimuma. Konveksne funkcije ohranijo takšne lastnosti tudi v neskončno-dimenzionalnih primerih, seveda pri primernih dodatnih predpostavkah, in so posledično najbolj razumljive funkcije v variacijskem računu. V teoriji verjetnosti, aplikacija konveksne funkcije na pričakovano vrednost slučajne spremenljivke je vedno manjša ali enaka pričakovani vrednosti konveksne funkcije slučajne spremenljivke. Ta neenakost je znana kot Jensenova neenakost, in je temelj veliko pomembnih neenakosti.

Namen te naloge je da bralca spozna z teorijo konveksnih množic in konveksnih funkcij, pri čemer centralno vlogo ima aplikacija na probleme ekstremov. Naloga predvsem temelji na knjigi Convex Analysis(Princeton University Press, 1970), avtorja R.T.Rockafellara. Vložen je bil trud, da se čim več pove a se pri tem ne presega meja relativno elementarnega tehniškega nivoja. Kar se tiče tehniških predpogojev, kljub temu da se od bralca ne zahteva poznavanje bolj globokih področjih abstraktne matematike, se vsekakor pričakuje dobro poznavanje linearne algebre in elementarne realne analize (konvergence zaporedij, zveznosti funkcij, odprtih in zaprtih množic, kompaktnosti, in podobno). Naloga je, kolikor je to bilo možno, urejena po principu subjekta snovi, npr. vsa dejstva o relativnem odprtju in zaprtju so zbrana v en del. Poleg tega je poudarek na logičnem razvoju snovi.

Najprej se bralec spozna z konveksnimi množicami in konveksnimi funkcijami. Poudarek je na vzpostavljanju kriterija za konveksnost. Številni primeri pomagajo pri razumevanju osnovnih konceptov. Bistvenega pomena je dejstvo da se konveksne funkcije na $\mathbb{R}^n$ lahko identificirajo z določenimi konveksnimi podmnožicami (epigrafi) v $\mathbb{R}^{n+1}$, dokler se konveksne množice v $\mathbb{R}^n$ lahko identificirajo z nekaterimi konveksnimi funkcijami (indikatorske funkcije) v $\mathbb{R}^n$. Zaradi tega se lahko "sprehajamo" med analitičnim in geometrijskim pristopom. Navadne funkcije razumemo geometrijsko v smislu njunih grafov, konveksne pa, nasproti temu, v smislu njunih epigrafov.

Ko se spoznamo z algebraičnimi lastnostimi, privzamemo bolj topološki pristop obravnavanju problema in sicer s pomočjo znanih pojmov kot so notranjost, zaprtje, in zveznost. Konveksne funkcije imajo res presentljivo topološko naravo, kar izsledi iz ene intuitivne predstave: če premični segment v konveksni množici C ima eno krajišče v notranjosti C in drugo krajišče na meji C, potem so vse točke med krajiščimi v notranjosti C. Predstavljeni so pojmi kot so relativna notranjost, relativna meja, efektivna domena in navzdol polzveznost. Slednji je za konveksne funkcije bolj pomemben kot običajna zveznost, ker je v direktni zvezi z epigrafom: funkcija je navzdol polzvezna če in samo če je njen epigraf zaprt. Vsako funkcijo lahko spremenimo v navzdol polzvezno, tako da redefiniramo njeno vrednost v natančno določenih točkah na robu njene efektivne domene. Po tej poti pridemo do koncepta zaprtja konveksne funkcije, ki temelji na operaciji zaprtja epigrafov kot podmnožice $\mathbb{R}^{n+1}$, kadar so funkcije pravilne. Ogledamo nekaj dejstev iz teorije dualnosti, od katerih je najbolj pomembno to da je zaprta konveksna množica presek vseh polprostorov ki jo vsebujejo.

Na koncu obravnavamo diferencialno teorijo, ki je osnovno orodje pri reševanju problema ekstremov. Spoznamo elementarno teorijo levih in desnih odvodov zaprte pravilne konveksne funkcije. Utrdimo zvezo med gradientimi in subgradientimi, in pokažemo da navadna preslikava gradient obstaja skoraj povsod in da je zvezna.

Simbol $\mathbb{R}$ nam bo vedno pomenil množico realnih števil, $\mathbb{R}^n$ pa običajni vektorski prostor realnih n-teric: $x = (\varepsilon_1, \ldots, \varepsilon_n)$. Razen če eksplicitno napovemo drugače, bomo predvsem delali v prostoru $\mathbb{R}^n$. Notranji produkt dveh vektorjev $x = (\varepsilon_1, \ldots, \varepsilon_n)$ in $x^* = (\varepsilon_1^*, \ldots, \varepsilon_n^*)$ v $\mathbb{R}^n$ se izraža z:

$$\langle x, x^* \rangle = \varepsilon_1 \varepsilon_1^* + \ldots + \varepsilon_n \varepsilon_n^*.$$

Simbol A bomo uporabljali takor za $m \times n$ realno matriko A, takor za pripadajočo linearno transformacijo $x \to Ax$ iz $\mathbb{R}^n$ v $\mathbb{R}^m$. Transponirano matriko in pripadajočo adjungirano linearno transformacijo iz $\mathbb{R}^m$ v $\mathbb{R}^n$ bomo označili z A^*, tako da imamo identiteto:

$$\langle Ax, y^* \rangle = \langle x, A^* y^* \rangle.$$

2 Afine množice, konveksne množice in stožci

2.1 Afine množice

Definicija 2.1. Naj bosta x in y različni točki v $\mathbb{R}^n$. Množica točk oblike

$$\{(1-\lambda)x + \lambda y = x + \lambda(y-x), \qquad \lambda \in \mathbb{R}\}$$

se imenuje *premica skozi x in y.*

Definicija 2.2. Podmnožica M v $\mathbb{R}^n$ se imenuje *afina množica* če je $(1-\lambda)x+\lambda y \in M$ za vsak $x \in M$, $y \in M$, $\lambda \in \mathbb{R}$.

Primer 2.3. Prazna množica $\emptyset$ in prostor $\mathbb{R}^n$ sta trivialni primeri afinih množic. Prav tako je množica M, ki sestoji iz le ene točke, po definiciji afina množica. V splošnem afina množica vsebuje, skupaj z poljubnimi dvema različnima točkama, celotno premico skozi dve točki.

Bolj formalna geometrija afinih množic sledi iz dejstev linearne algebre o podprostorih v $\mathbb{R}^n$. Natančna korespondenca med afinimi množicami in podprostorimi bo bolj podrobno opisana z naslednjima dvema trditvama.

Izrek 2.4. *Podprostori v $\mathbb{R}^n$ so natanko tiste afine množice ki vsebujejo koordinatno izhodišče.*

Dokaz. Vsak podprostor vsebuje 0 in, ker je zaprt za seštevanje in množenje s skalarjem, je v posebnem afina množica. Po drugi strani, naj bo M afina množica ki vsebuje 0. Za vsak $x \in M$ in $\lambda \in \mathbb{R}$, imamo:

$$\lambda x = (1-\lambda)0 + \lambda x \in M,$$

tako da je M zaprt za množenje s skalarjem. Če velja $x \in M$ in $y \in M$, imamo:

$$\frac{1}{2}(x+y) = \frac{1}{2}x + (1-\frac{1}{2})y \in M,$$

in odtod

$$x + y = 2(\frac{1}{2}(x+y)) \in M.$$

Torej, M je tudi zaprta za seštevanje in je podprostor. □

Definicija 2.5. Za $M \subset \mathbb{R}^n$ in $a \in \mathbb{R}^n$, je *premik* množice M za vektor a množica:

$$M + a = \{x + a \big| x \in M\}.$$

Enostavno lahko vidimo, da je premik afine množice zopet afina množica.

Definicija 2.6. Afina množica M je *vzporedni premik* afine množice L, če je $M = L+a$ za nek vektor a.

Očitno je "M je vzporedni premik L" ekvivalenčna relacija na kolekciji afinih podmnožic $\mathbb{R}^n$.

Primer 2.7. Premica je v običajnem smislu lahko paralelna ravnini, vendar v zgoraj definiranem ni. Sklepamo, da je slednja definicija paralelnosti bolj restriktivna kot običajna.

Izrek 2.8. *Vsaka neprazna afina množica M je vzporedni premik natančno določenem podprostoru L, ki je podan z:*

$$L = M - M = \{x - y \big| x \in M, y \in M\}.$$

Dokaz. Pokažimo najprej, da M ne more biti vzporedni premik dvema različnima podprostorima. Podprostora L_1 in L_2, ki bi bila vzporedna premaknjeva M, bi tedaj bila vzporedna med sabo, kar bi pomenilo $L_2 = L_1 + a$ za nek a in, ker je $0 \in L_2$, še $-a \in L_1$, s tem tudi $a \in L_1$. Tedaj $L_1 \supset L_1 + a = L_2$! Podoben argument nas prinese do sklepa $L_2 \supset L_1$, kar pomeni $L_1 = L_2$. To zagotavlja enoličnost. Opazimo, da je za poljuben $y \in M$, $L := M - y = M + (-y)$ premik M ki vsebuje 0, ki je po 2.4 podprostor vzporeden M. Zaradi enoličnosti je $L = M - y$, neodvisno od tega kateri $y \in M$ izberemo, torej imamo $L = M - M$. □

Definicija 2.9. *Dimenzija* neprazne afine množice je dimenzija njenega paralelnega podprostora. Dimenzija $\emptyset$ je -1. Afine množice dimenzije 0, 1, ali 2 imenujemo *točke*, *premice*, in *ravnine*, respektivno.

Definicija 2.10. Afina množica dimenzije $n - 1$ v $\mathbb{R}^n$ se imenuje *afina hiperravnina.*

Hiperravnine so zelo pomembne, ker imajo v n-dimenzionalni geometriji vlogo podobno vlogi točk v navadni geometriji. Hiperravnine in druge afine množice si lahko predstavljamo z linearnimi funkcijami in linearnimi enačbami. To se lahko sklepa iz teorije ortogonalnosti v $\mathbb{R}^n$. Spomnimo se, $x \perp y$ pomeni $\langle x, y\rangle = 0$. Za dani podprostor L v $\mathbb{R}^n$, se množica vektorjev x takšnih da je $x \perp L$ oz. $x \perp y$ za vsak $y \in L$, imenuje *ortogonalni komplement* L, označujemo ga z $L^\perp$. Velja, da je $L^\perp$ tudi podprostor, in

$$dimL + dimL^\perp = n.$$

Ortogonalni komplement $(L^\perp)^\perp$ podprostora $L^\perp$ je L. Res,

$$(L^\perp)^\perp = \{z \big| \langle z, y\rangle = 0 \quad \forall y \in L^\perp\} =$$

$$= \{z \big| \langle z, y\rangle = 0 \quad \forall y : \langle x, y\rangle = 0 \quad \forall y \in L\} = L.$$

Če vektorji $b_1, \ldots, b_m$ sestavljajo bazo za L, potem je $x \perp L$ ekvivalentno pogoju da je $x \perp b_1, \ldots, x \perp b_m$. V posebnem primeru velja, da so $(n-1)$-dimenzionalni podprostori $\mathbb{R}^n$ ortogonalni komplementi enodimenzionalnih podprostorov L z bazo sestavljeno iz posameznega neničelnega vektorja b. Se pravi, $(n-1)$-dimenzionalni podprostori so množice oblike $\{x \big| x \perp b\}$, kjer je $b \in \mathbb{R}^n \backslash \{0\}$. Hiperravnine so njihovi premiki. Ampak

$$\{x \big| x \perp b\} + a = \{x + a \big| \langle x, b\rangle = 0\} = \{y \big| \langle y - a, b\rangle = 0\} = \{y \big| \langle y, b\rangle = \beta\},$$

kjer je $\beta = \langle a, b\rangle$. Po tej poti smo prišli do naslednje karakterizacije afine hiperravnine.

Izrek 2.11. *Za dan $\beta \in \mathbb{R}$ in neničelni $b \in \mathbb{R}^n$, množica*

$$H = \{x \big| \langle x, b\rangle = \beta\}$$

je afina hiperravnina v $\mathbb{R}^n$. Še več, vsaka afina hiperravnina se lahko predstavi na takšen način, in sicer z b in β določenima do skupnega neničelnega večkratnika.

Pravimo, da je takšen vektor b *normalen* hiperravnini H. Vsaki drugi vektor, ki je normalen hiperravnini H, je nujno pozitivni ali negativni skalarni večkratnik vektorja b.

Naslednji izrek nam poda karakterizacijo afinih podmnožic v $\mathbb{R}^n$ v smislu rešitev sistemov linearnih enačb n spremenljivk.

Izrek 2.12. *Za dan $b \in \mathbb{R}^m$ in $m \times n$ realno matriko B, je množica*

$$M = \{x \in \mathbb{R}^n \big| Bx = b\}$$

afina v $\mathbb{R}^n$. Še več, vsako afino množico lahko predstavimo na tak način.

Dokaz. Če je $x \in M$, $y \in M$, $\lambda \in \mathbb{R}$, potem za $z = (1-\lambda)x + \lambda y$ velja

$$Bz = (1-\lambda)Bx + \lambda By = (1-\lambda)b + \lambda b = b,$$

torej $z \in M$. Sklepamo, da je M afina množica. Po drugi strani, če začnemo s poljubno neprazno afino množico M različno od $\mathbb{R}^n$, bo za podprostor $L = M - M$, in bazo $b_1, \ldots, b_m$ podprostora $L^\perp$ veljalo

$$L = (L^\perp)^\perp = \{x \big| x \perp b_1, \ldots, x \perp b_m\}$$

$$= \{x \big| \langle x, b_i \rangle = 0, \quad i = 1, \ldots, m\} = \{x \big| Bx = 0\},$$

kjer je B $m \times n$ matrika z vrsticami $b_1, \ldots, b_m$. Zaradi paralelnosti M in L, obstaja $a \in \mathbb{R}^n$ takšen da

$$M = L + a = \{x \big| B(x - a) = 0\} = \{x \big| Bx = b\},$$

kjer je $b = Ba$. Končno, afine množice $\emptyset$ in $\mathbb{R}^n$ lahko predstavimo v dani obliki če privzamemo za B $m \times n$ ničelno matriko, npr. z $b = 0$ za $\mathbb{R}^n$ in $b \neq 0$ za $\emptyset$. $\square$

Opazimo, da imamo v 2.12

$$M = \{x \big| \langle x, b_i \rangle = \beta_i, i = 1, \ldots, m\} = \bigcap_{i=1}^{m} H_i,$$

kjer je b_i i-ta vrstica B, β_i i-ta komponenta b, in

$$H_i = \{x \big| \langle x, b_i \rangle = \beta_i\}.$$

Vsaka H_i je hiperravnina ($b_i \neq 0$), prazna množica ($b_i = 0, \beta_i \neq 0$), ali $\mathbb{R}^n$ ($b_i = 0, \beta_i = 0$). Prazno množico si lahko predstavljamo kot presek dveh različnih paralelnih hiperravnin, medtem ko si $\mathbb{R}^n$ lahko predstavljamo kot presek prazne kolekcije hiperravnin na $\mathbb{R}^n$. Torej smo ugotovili, da drži naslednja posledica zgornjega izreka:

Posledica 2.13. *Vsaka afina podmnožica v $\mathbb{R}^n$ je presek končne kolekcije hiperravnin.*

Afina množica M iz 2.12 se lahko izrazi s pomočjo vektorjev $b'_1, \ldots, b'_n$, ki formirajo stolpce B, na naslednji način

$$M = \{x = (\varepsilon_1, \ldots, \varepsilon_n) \big| \varepsilon_1 b'_1 + \cdots + \varepsilon_n b'_n = b\}.$$

Namreč, ker je

$$x = (\varepsilon_1, \ldots, \varepsilon_n) = \varepsilon_1(1, 0, \ldots, 0) + \cdots + \varepsilon_n(0, \ldots, 0, 1), \text{sledi}$$

$$b = Bx = B\varepsilon_1(1, 0, \ldots, 0) + \cdots + B\varepsilon_n(0, \ldots, 0, 1) =$$

$$\varepsilon_1 B(1, 0, \ldots, 0) + \cdots + \varepsilon_n B(0, \ldots, 0, 1) = \varepsilon_1 b'_1 + \cdots + \varepsilon_n b'_n.$$

Očitno, je presek poljubne kolekcije afinih množic zopet afina množica. Opazimo, da je za poljubno afino množico $S \subset \mathbb{R}^n$, presek kolekcije afinih množic M takšnih da velja $M \supset S$, afina množica, ki vsebuje S. Še več, omenjeni presek je natančno določena najmanjša afina množica ki vsebuje S.

Definicija 2.14. Za dano množico $S \subset \mathbb{R}^n$, najmanjšo natančno določeno afino množico, ki jo vsebuje, imenujemo *afina ogrinjača* in jo označimo z aff$S$.

Definicija 2.15. Množica $m+1$ točk $\{b_o, b_1, \ldots, b_m\}$ je *afino neodvisna* če je aff$\{b_o, b_1, \ldots, b_m\}$ m-dimenzionalna množica.

Očitno,

$$\text{aff}\{b_o, b_1, \ldots, b_m\} = L + b_0,$$

kjer je

$$L = \text{aff}\{0, b_1 - b_0, \ldots, b_m - b_0\}.$$

Po 2.4, L je enak najmanjšem podprostoru ki vsebuje $b_1 - b_0, \ldots, b_m - b_0$, dimenzije m, če in samo če so $b_1 - b_0, \ldots, b_m - b_0$ linearno neodvisni. Se pravi, $b_0, \ldots, b_m$ so afino neodvisni če in samo če so $b_1 - b_0, \ldots, b_m - b_0$ linearno neodvisni.

Primer 2.16. Poljubna afino neodvisna množica $m+1$ točk v $\mathbb{R}^n$ se lahko razširi na afino neodvisno množico $n+1$ točk: namreč, poljubna m-dimenzionalna afina množica M se lahko izraža kot afina ogrinjača $m+1$ točk (premikanjem točk ki pripadajo bazi podprostora paralelnega M).

Če je $M = \text{aff}\{b_1, \ldots, b_m\}$, so vektorji v podprostoru L vzporedno premaknjenem M linearne kombinacije $b_1 - b_0, \ldots, b_m - b_0$. Vektorji v M so torej tisti ki se lahko izražajo v obliki:

$$x = \lambda_1(b_1 - b_0) + \cdots + \lambda_m(b_m - b_0) + b_0$$

oz. v obliki

$$x = \lambda_0 b_0 + \lambda_1 b_1 + \cdots + \lambda_m b_m, \quad \lambda_0 + \lambda_1 + \cdots + \lambda_m = 1.$$

Koeficienti v takšnem izrazu za x so natačno določeni če in samo če so $b_0, b_1, \ldots, b_m$ afino neodvisni. V tem primeru pravimo, da $\lambda_0, \lambda_1, \ldots, \lambda_m$, kot parametri, definirajo *baricentrični koordinatni sistem* za M.

Definicija 2.17. Pravimo, da je preslikava $T : x \longrightarrow Tx$ iz $\mathbb{R}^n$ v $\mathbb{R}^m$ *afina transformacija* če drži

$$T((1-\lambda)x + \lambda y) = (1-\lambda)Tx + \lambda Ty$$

za vsak $x, y \in \mathbb{R}^n$, in $\lambda \in \mathbb{R}$.

Izrek 2.18. *Afine transformacije iz $\mathbb{R}^n$ v $\mathbb{R}^m$ so preslikave T oblike $Tx = Ax + a$, kjer je A linearna transformacija in $a \in \mathbb{R}^m$.*

Dokaz. Če je T afina, naj bo $a = T0$ in $Ax = Tx - a$. Potem je A afina transformacija in $A0 = T0 - T0 = 0$. Linearnost A sledi iz

$$A(\lambda x) = A((1-\lambda)0 + \lambda x) = T((1-\lambda)0 + \lambda x) - T0 = (1-\lambda)T0 + \lambda Tx - T0 =$$

$$\lambda Tx - \lambda T0 = \lambda(Tx - T0) = \lambda Ax, \text{in}$$

$$A(x+y) = A(2(\frac{1}{2}x + (1-\frac{1}{2})y)) = 2(A(\frac{1}{2}x + (1-\frac{1}{2})y)) = 2T(\frac{1}{2}x + (1-\frac{1}{2})y)) - 2T0 =$$

$$2T(\frac{1}{2}x) - T0 + 2T(1-\frac{1}{2})y) - T0 = Tx - T0 + Ty - T0 = Ax + Ay.$$

Po drugi strani, če je $Tx = Ax + a$, kjer je A linearna, imamo

$$T((1-\lambda)x+\lambda y) = (1-\lambda)Ax+\lambda Ay+a = (1-\lambda)Ax+\lambda Ay+(1-\lambda)a+\lambda a = (1-\lambda)Tx+\lambda Ty.$$

Sledi, da je T afina. □

Napovejmo še, da je inverz afine transformacije, če obstaja, sam afina transformacija.

Izrek 2.19. *Naj bosta $\{b_0, b_1, \ldots, b_m\}$ in $\{b'_0, b'_1, \ldots, b'_m\}$ afino neodvisni množici v $\mathbb{R}^n$. Potem obstaja injektivna afina transformacija T prostora $\mathbb{R}^n$ nase, takšna da je $Tb_i = b'_i$ za $i = 0, \ldots, m$. Če je $m = n$, T je enolična določena.*

Dokaz. Po potrebi lahko zvečamo dani množici, dokler ne pridemo do slučaja $m = n$, seveda brez škode po splošnost. Iz linearne algebre je znano, da obstaja natančno določena injektivna linearna transformacija A prostora $\mathbb{R}^n$ nase, ki preslika bazo $\{b_1 - b_0, \ldots, b_n - b_0\}$ prostora $\mathbb{R}^n$ v bazo $\{b'_1 - b'_0, \ldots, b'_n - b'_0\}$. Afina transformacija ki jo iščemo je tedaj dana z $Tx = Ax + a$, kjer je $a = b'_0 - Ab_0$. □

2.2 Konveksne množice in stožci

Definicija 2.20. Podmnožica C v $\mathbb{R}^n$ je *konveksna* če je $(1-\lambda)x+\lambda y \in C$ za vsak $x \in C$, $y \in C$, in $0 \leq \lambda \leq 1$.

Primer 2.21. Vse afine množice so konveksne. Kocke v $\mathbb{R}^3$ so konveksne vendar niso afine množice. Prazna množica $\emptyset$ in $\mathbb{R}^n$ sta obe afini in prav tako konveksni množici.

Definicija 2.22. Za $x, y \in \mathbb{R}^n$ imenujemo množico $[x,y] := \{(1-\lambda)x+\lambda y) \big| 0 \leq \lambda \leq 1\}$ *daljica* med x in y.

Primer 2.23. Daljice so konveksne vendar niso afine množice, razen če sestoje iz ene točke.

Definicija 2.24. Za vsak neničelni $b \in \mathbb{R}^n$ in vsak $\beta \in \mathbb{R}$, množice

$$\{x \big| \langle x, b\rangle \leq \beta\}, \qquad \{x \big| \langle x, b\rangle \geq \beta\},$$

imenujemo *zaprti polprostori.* Množice

$$\{x \big| \langle x, b\rangle < \beta\}, \qquad \{x \big| \langle x, b\rangle > \beta\},$$

imenujemo *odprti polprostori.*

Vse štiri množice so očitno neprazne in konveksne.

Primer 2.25. Hiperravnine so afine in konveksne množice, polprostori le konveksne.

Izrek 2.26. *Presek poljubne kolekcije konveksnih množic je konveksna množica.*

Dokaz. Elementarno. □

Posledica 2.27. *Naj bodo $b_i \in \mathbb{R}^n$ in $\beta_i \in \mathbb{R}$ za $i \in \mathbb{I}$, kjer je $\mathbb{I}$ poljubna indeksna množica. Potem je množica*

$$C = \{x \in \mathbb{R}^n \big| \langle x, b_i\rangle \leq \beta_i, \forall i \in \mathbb{I}\}$$

konveksna.

Dokaz. Naj bo $C_i = \{x \big| \langle x, b_i\rangle \leq \beta_i\}$. Potem za C_i velja da je ali zaprt polprostor ali $\mathbb{R}^n$ ali $\emptyset$(če je recimo $b_i = 0$ in $B_i < 0$) in zaradi tega je $C = \cap_{i\in\mathbb{I}} C_i$ konveksna množica, kot presek konveksnih množic. □

Posledica bi bila še vedno veljavna, če bi nekatere med neenakostimi "$\leq$" nadomestili z "$\geq$", "$<$", "$>$", ali "$=$". Na podlagi te ugotovitve sklepamo, da je za poljuben sistem linearnih neenakosti in enačb v n spremenljik, sistem rešitev C konveskna množica v $\mathbb{R}^n$.

Definicija 2.28. Množica, ki se lahko izrazi kot presek *končno* mnogo zaprtih polprostorov v $\mathbb{R}^n$ se imenuje *poliedrična* konveksna množica.

Iskaže se, da se poliedrične množice se obnašajo precej boljše kot splošne konveksne množice.

Definicija 2.29. Vsota vektorjev

$$\lambda_1 x_1 + \cdots + \lambda_m x_m$$

se imenuje *konveksna kombinacija* $x_1, \ldots, x_m$ če so vsi koeficienti λ_i nenegativni in je $\lambda_1 + \cdots + \lambda_n = 1$.

Primer 2.30. Če se m delcev z masami $\alpha_1, \ldots, \alpha_m$ nahaja v točkah $x_1, \ldots, x_m$ v prostoru $\mathbb{R}^3$, je težišče sistema točka $\lambda_1 x_1 + \cdots + \lambda_m x_m$, kjer je

$$\lambda_i = \frac{\alpha_i}{\alpha_1 + \cdots + \alpha_m}.$$

Izrek 2.31. *Podmnožica množice $\mathbb{R}^n$ je konveksna če in samo če vsebuje vse konveksne kombinacije njenih elementov.*

Dokaz. Množica C je po definiciji konveksna če in samo če je $\lambda_1 x_1 + \lambda_2 x_2 \in C$ za $x_1 \in C, x_2 \in C, \lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_1 + \lambda_2 = 1$. Pokažimo, da to pomeni da je C zaprta za konveksne kombinacije $m > 2$ njenih elementov. (Torej, iz ekvivalentnosti konveksnosti in dejstva da množica vsebuje vse konveksne kombinacije njenih elementov za slučaj ko je $m = 2$ bo sledila ekvivalentnost teh pogojev za poljuben $m > 2$) Vzemimo $m > 2$ in recimo da je C zaprt za konveksne kombinacije z manj kot m vektorjev. Za konveksno kombinacijo $x = \lambda_1 x_1 + \cdots + \lambda_m x_m \in C$, vsaj en od skalarjev λ_i je različen od 1 (v nasprotnem bi bilo $\lambda_1 + \cdots + \lambda_m = m \neq 1$); naj bo tisti λ_i in

$$y = \lambda'_2 x_2 + \cdots + \lambda'_m x_m, \quad \lambda'_i = \frac{\lambda_i}{1 - \lambda_1}.$$

Potem $\lambda_i \geq 0$ za $i = 2, \ldots, m$, in

$$\lambda'_2 + \cdots + \lambda'_m = \frac{\lambda_2 + \cdots + \lambda_m}{\lambda_2 + \cdots + \lambda_m} = 1.$$

Sledi, da je y konveksna kombinacija $m - 1$ elementov C in $y \in C$, po indukciji. Ker je $x = (1 - \lambda_1) y + \lambda_1 x_1$, sledi $x \in C$. □

Definicija 2.32. Presek vseh konveksnih množic ki vsebujejo dano množico $S \subset \mathbb{R}^n$ se imenuje *konveksna ogrinjača* množice S (oznaka conv$S$).

Po 2.26, conv$S$ je konveksna množica.

Izrek 2.33. *Za poljuben $S \subset \mathbb{R}^n$, conv$S$ sestoji iz vseh konveksnih kombinacij elementov iz S.*

Dokaz. Elementi množice S pripadajo conv$S$, in torej tudi konveksne kombinacije elementov iz S, po 2.31. Po drugi strani, naj bosta $x = \lambda_1 x_1 + \cdots + \lambda_m x_m$ in $y = \mu_1 y_1 + \cdots + \mu_r y_r$ dve konveksni kombinaciji, kjer $x_i, y_j \in S$. Vektor

$$(1-\lambda)x + \lambda y = (1-\lambda)\lambda_1 x_1 + \cdots + (1-\lambda)\lambda_m x_m + \lambda\mu_1 y_1 + \cdots + \lambda\mu_r y_r,$$

kjer je $0 \leq \lambda \leq 1$, je konveksna kombinacija elementov S, saj je

$$(1-\lambda)(\lambda_1 + \cdots + \lambda_m) + \lambda(\mu_1 + \cdots + \mu_r) = (1-\lambda)1 + \lambda 1 = 1.$$

Sklepamo, da je množica konveksnih kombinacij elementov S konveksna. Ker vsebuje S sledi, da mora sovpadati z najmanjšo množico z to lastnostjo, namreč conv$S$. □

Posledica 2.34. *Konveksna ogrinjača končne podmnožice $\{b_0, \ldots, b_m\}$ množice $\mathbb{R}^n$ sestoji iz vseh vektorjev oblike $\lambda_0 b_0 + \cdots + \lambda_m b_m$, kjer je $\lambda_0 \geq 0, \ldots, \lambda_m \geq 0$, in $\lambda_0 + \cdots + \lambda_m = 1$.*

Dokaz. Vsako konveksno kombinacijo elementov iz $\{b_0, \ldots, b_m\}$ lahko izrazimo kot konveksno kombinacijo $b_0, \ldots, b_m$ tako da damo ničelne koeficiente pred vektorje b_i ki nam niso potrebni. □

Definicija 2.35. Množica ki je konveksna ogrinjača končno mnogo točk se imenuje *politop*.

Primer 2.36. Polieder in politop sta različna pojma. Politop je omejeni polieder.

Definicija 2.37. Če je množica $\{b_0, b_1, \ldots, b_m\}$ afino neodvisna, imenujemo njeno konveksno ogrinjačo *m-dimenzionalni simpleks*, $b_0, b_1, \ldots, b_m$ pa imenujemo *oglišča* simpleksa.

Dimenzijo konveksne množice C definiramo kot dimenzijo afine ogrinjače množice C.

Izrek 2.38. *Dimenzija konveksne množice C je maksimalna dimenzija različnih simpleksov vsebovanih v C.*

Dokaz. Konveksna ogrinjača poljubne podmnožice v C je vsebovana v C. Torej, maksimalna dimenzija različnih simpleksov vsebovanih v C je največji m takšen da C vsebuje afino neodvisno množico $m+1$ elementov. Recimo, da je $\{b_0, b_1, \ldots, b_m\}$ takšna množica z maksimalnim m, in da je M njena afina ogrinjača. Potem je dim$M = m$ in $M \subset$ aff$C$. Še več, $C \subset M$, ker bi drugače za nek $b \in C \backslash M$ množica $m+2$ elementov $b_0, b_1, \ldots, b_m, b$ bila afino neodvisna, kar bi bilo v nasprotju z maksimalnostjo m. Ker je aff$C$ najmanjša afina množica ki vsebuje C sledi, da je aff$C = M$ in odtod dim$C = m$. □

Definicija 2.39. Podmnožica K v $\mathbb{R}^n$ se imenuje *stožec* če je zaprta za množenje s pozitivnim skalarjem, oz. če je $\lambda x \in K$ za $x \in K$, $\lambda > 0$.

Stožec je unija pol-premic ki izvirajo izkoordinatnega izhodišča. Koordinatno izhodišče je lahko, vendar ne nujno, del te množice.

Definicija 2.40. *Konveksni stožec* je stožec ki je konveksna množica.

Primer 2.41. Polprostori $\mathbb{R}^n$ so konveksni stožci. Isto velja za odprte in zaprte polprostore glede na hiperravnino skozi koordinatno izhodišče.

Najbolj pomembni primeri konveksnih stožcev sta naslednji dve množici.

Definicija 2.42. *Nenegativni ortant* množice $\mathbb{R}^n$ je

$$\{x = (\varepsilon_1, \ldots, \varepsilon_n) \in \mathbb{R}^n \big| \varepsilon_1 \geq 0, \ldots, \varepsilon_n \geq 0\}.$$

Pozitivni ortant množice $\mathbb{R}^n$ je

$$\{x = (\varepsilon_1, \ldots, \varepsilon_n) \in \mathbb{R}^n \big| \varepsilon_1 > 0, \ldots, \varepsilon_n > 0\}.$$

Izrek 2.43. *Presek poljubne kolekcije konveksnih stožcev je konveksni stožec.*

Dokaz. Elementarno. □

Posledica 2.44. *Naj bo $b_i \in \mathbb{R}^n$ za $i \in \mathbb{I}$, kjer je $\mathbb{I}$ poljubna indeksna množica. Potem je*

$$K = \{x \in \mathbb{R}^n \big| \langle x, b_i \rangle \leq 0, i \in \mathbb{I}\}$$

konveksni stožec.

Dokaz. Kot v 2.27. □

Izrek 2.45. *Podmnožica množice $\mathbb{R}^n$ je konveksni stožec če in samo če je zaprta za seštevanje in množenje s pozitivnim skalarjem.*

Dokaz. Naj bo K stožec, in $x, y \in K$. Če je K konveksen stožec, potem je $z = \frac{1}{2}(x + y) \in K$ in $x + y = 2z \in K$. Po drugi strani, če je K zaprt za seštevanje, in $0 < \lambda < 1$, vektorji $(1-\lambda)x$ in λy pripadata K, in odtod $(1-\lambda)x+\lambda y \in K$. Sklepamo, da je K konveksen če in samo če je zaprt za seštevanje. □

Posledica 2.46. *Podmnožica množice $\mathbb{R}^n$ je konveksni stožec če in samo če vsebuje vse pozitivne linearne kombinacije njenih elementov.*

Posledica 2.47. *Naj bo S poljubna podmnožica množice $\mathbb{R}^n$ in K množica vseh pozitivnih linearnih kombinacij elementov iz S. Potem je K najmanjši konveksni stožec ki vsebuje S.*

Dokaz. Očitno je K zaprta za seštevanje in množenje s pozitivnim skalarjem. Prav tako pa $K \supset S$. Vsaki konveksni stožec ki vsebuje S mora, po drugi strani, vsebovati K. □

Posledica 2.48. *Naj bo C konveksna množica in*

$$K = \{\lambda x \big| \lambda > 0, x \in C\}.$$

Potem je K najmanjši konveksni stožec ki vsebuje C.

Dokaz. Vsaka pozitivna linearna kombinacija elementov C je po 2.47 pozitivni skalarni večkratnik konveksne kombinacije elementov C in je zaradi tega element množice K. □

2.2.1 Algebra konveksnih množic

Veliko zanimivih algebraičnih operacij ohrani razred konveksnih množic. Npr. če je C konveksna podmnožica v $\mathbb{R}^n$, potem sta konveksni tudi vsaka translacija $C + a$ in vsaka množica $\lambda C = \{\lambda x \big| x \in C\}$. Slednjo množico običajno imenujemo skalarni večkratnik množice C.

Definicija 2.49. Rečemo, da je konveksna množica *simetrična*, če velja $C = -C = \{-x \big| x \in C\}$.

Simetrična množica nujno vsebuje koordinatno izhodišče.

Izrek 2.50. *Če sta C_1 in C_2 konveksni množici v $\mathbb{R}^n$, tedaj je konveksna tudi*

$$C_1 + C_2 = \{x_1 + x_2 \big| x_1 \in C_1, x_2 \in C_2\}.$$

Dokaz. Naj bosta x, y točki v $C_1 + C_2$. Potem obstajajo vektorji $x_1 \in C_1$, $y_1 \in C_1$, $x_2 \in C_2$, $y_2 \in C_2$, takšni da je

$$x = x_1 + x_2, \qquad y = y_1 + y_2.$$

Za $0 < \lambda < 1$ imamo

$$(1-\lambda)x + \lambda y = [(1-\lambda)x_1 + \lambda y_1] + [(1-\lambda)x_2 + \lambda y_2],$$

iz zaradi konveksnosti C_1 in C_2

$$(1-\lambda)x_1 + \lambda y_1 \in C_1, \qquad (1-\lambda)x_2 + \lambda y_2 \in C_2.$$

Odtod sklepamo, da velja $(1-\lambda)x + \lambda y \in C_1 + C_2$. □

Primer 2.51. Naj bo $C_1 \subset \mathbb{R}^n$ poljubna konveksna množica in $C_2 \subset \mathbb{R}^n$ nenegativni ortant. Za $x = (x_1, \ldots, x_n)$ in $y = (y_1, \ldots, y_n)$ definirajmo relaciji na $\mathbb{R}^n$:

$$x \geq y \Leftrightarrow x_1 \geq y_1, \ldots, x_n \geq y_n$$

$$x \leq y \Leftrightarrow x_1 \leq y_1, \ldots, x_n \leq y_n.$$

Potem je

$$C_1 + C_2 = \{x_1 + x_2 \big| x_1 \in C_1, x_2 \geq 0\} = \{x \big| \exists x_1 \in C_1, x_1 \leq x\},$$

slednja množica je konveksna po 2.50 kadar je C_1 konveksna množica.

Opazimo, da če so $C_1, \ldots, C_m$ konveksne množice, potem je to tudi linearna kombinacija

$$C = \lambda_1 C_1 + \cdots + \lambda_m C_m.$$

Definicija 2.52. Linearna kombinacija množic $C_1, \ldots, C_m$, $C = \lambda_1 C_1 + \cdots + \lambda_m C_m$, se imenuje *konveksna kombinacija* $C_1, \ldots, C_m$ kadar velja $\lambda_1 \geq 0, \ldots, \lambda_m \geq 0$ in $\lambda_1 + \cdots + \lambda_m = 1$.

Lahko si geometrično predstavljamo C kot "mešanico" $C_1, \ldots, C_m$. Npr. če sta C_1 in C_2 trikotnik in krožni disk v $\mathbb{R}^2$, ko λ teče od 0 proti 1, $C = (1-\lambda)C_1 + \lambda C_2$ spreminja svojo obliko iz trikotnika v trikotnik z zaobljenimi koti. Zaradi geometrične intuicije se splača če si predstavljamo $C_1 + C_2$ kot unijo vseh premikov $x_1 + C_2$ ko x_1 varira po C_1.

Spomnimo se, da brez predpostavke konveksnosti velja $C_1 + C_2 = C_2 + C_1$, $(C_1 + C_2) + C_3 = C_1 + (C_2 + C_3)$, $\lambda_1(\lambda_2 C) = (\lambda_1 \lambda_2)C$ in $\lambda(C_1 + C_2) = \lambda C_1 + \lambda C_2$ za množice C_1, C_2 in C_3. Konveksna množica, ki sestoji le iz enega elementa, 0, je identiteta za seštevanje. Inverzi za seštevanje ne obstajajo za množice ki vsebujejo več kot eno točko, največ kar lahko povemo je da je v splošnem $0 \in [C + (-C)]$ ko $C \neq \emptyset$.

Na srečo, obstaja vsaj eno pomembno dejstvo iz algebre množic ki je odvisno od konveksnosti, kot si bomo ogledali v naslednjem izreku. Izpolnjevanje tega je prav zaprav ekvivalentno konveksnosti množice C, ker pomeni da je $\lambda C + (1-\lambda)C$ vsebovan v C ko je $0 \leq \lambda \leq 1$.

Izrek 2.53. *Če je C konveksna množica in $\lambda_1 \geq 0, \lambda_2 \geq 0$, potem velja*

$$(\lambda_1 + \lambda_2)C = \lambda_1 C + \lambda_2 C.$$

Dokaz. Inkluzija $(\lambda_1 + \lambda_2)C \subset \lambda_1 C + \lambda_2 C$ očitno velja tudi brez predpostavke konveksnosti. Po drugi strani, zaradi konveksnosti velja

$$C \supset \frac{\lambda_1}{\lambda_1 + \lambda_2} C + \frac{\lambda_2}{\lambda_1 + \lambda_2} C,$$

če je $\lambda_1 + \lambda_2 > 0$, ker nam konveksnost množice C poda $(1-\lambda)C + \lambda C \subset C$ za $0 < \lambda < 1$. Če je pa $\lambda_1 = 0$ ali $\lambda_2 = 0$, imamo trivialen slučaj. □

Za poljubni dve konveksni množici C_1 in C_2 v $\mathbb{R}^n$, obstaja natančno določena največja konveksna množica vsebovana u C_1 in C_2, namreč $C_1 \cap C_2$, in natančno določena najmanjša konveksna množica ki vsebuje C_1 in C_2, namreč $\mathrm{conv}(C_1 \cup C_2)$. Očitno je, da to ne velja samo za dve množici, ampak za poljubno družino množic $\{C_i \big| i \in \mathbb{I}\}$. Povedeno drugače, kolekcija vseh konveksnih podmnožic v $\mathbb{R}^n$ je polna mreža za naravno delno urejenost tj. inkluzijo.

Definicija 2.54. Za poljubno linearno transformacijo A iz $\mathbb{R}^n$ v $\mathbb{R}^m$, definiramo

$$AC = \{Ax \big| x \in C\} \text{za } C \subset \mathbb{R}^n$$

$$A^{-1}D = \{x \big| Ax \in D\} \text{za } D \subset \mathbb{R}^m,$$

in jih imenujemo *slika* C od A in *praslika* D pod A, respektivno.

Izrek 2.55. *Naj bo A linearna transformacija iz $\mathbb{R}^n$ v $\mathbb{R}^m$. Potem za vsako konveksno množico C v $\mathbb{R}^n$ velja, da je AC konveksna množica v $\mathbb{R}^m$ in prav tako je $A^{-1}D$ konveksna množica v $\mathbb{R}^n$ za vsako konveksno množico D v $\mathbb{R}^m$.*

Dokaz. Elementarno. □

Posledica 2.56. *Ortogonalna projekcija konveksne množice C na podprostor L je konveksna množica.*

Dokaz. Ortogonalna projekcija, ki preslika na L je linearna transformacija ki vsaki točki x priredi natančno določeni $y \in L$ tako da $(x-y) \perp L$. □

Izrek 2.57. *Naj bosta C in D konveksni množici v $\mathbb{R}^m$ in $\mathbb{R}^n$, respektivno. Tedaj je*

$$C \oplus D := \{x = (y,z) \big| y \in C, z \in D\}$$

konveksna množica v $\mathbb{R}^{m+n}$.

Dokaz. Trivialno. □

Definicija 2.58. Množica $C \oplus D$ se imenuje *direktna vsota* C in D. Enako ime uporabljamo za vsoto $C + D$, $C \subset \mathbb{R}^n$, $D \subset \mathbb{R}^n$, če se vsak vektor $x \in C + D$ lahko izraža na natančno določen način v obliki $x = y + z$, pri čemer sta $y \in C$ in $z \in D$.

3 Konveksne funkcije

Definicija 3.1. Naj bo f funkcija katere vrednosti so realne ali $+\infty$ ali $-\infty$, in katere domena je podmnožica S v $\mathbb{R}^n$. Množica

$$\{(x,\mu)\big|x \in S, \mu \in \mathbb{R}, \mu \geq f(x)\}$$

se imenuje *epigraf* funkcije f, in se označa z epi$f$.

Pojem epigrafa za konveksne funkcije je analogen pojmu grafa za navadne funkcije, tj. jo popolnoma karakterizira.

Definicija 3.2. Funkcija f je *končna* če lahko zavzame le končno mnogo vrednosti.

Definicija 3.3. Funkcija f je *konveksna* funkcija na $S \subset \mathbb{R}^n$ če je epi$f$ konveksna podmnožica $\mathbb{R}^{n+1}$. Funkcija f na S je *konkavna* če je funkcija $-f$ konveksna na S.

Opazimo, da s tem novim orodjem lahko definiramo afino funkcija kot funkcijo katera je končna, konveksna in konkavna.

Izrek 3.4. *Vsaka afina funkcija je konveksna.*

Dokaz. Opazimo, da je definicija konveksnosti funkcije f na $S \subset \mathbb{R}^n$ ekvivalentna pogoju, da za poljubna x in y iz S in $0 \leq \lambda \leq 1$ velja:

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

kot si bomo še enkrat ogledali mal kasneje. Po drugi strani, funkcija f je afina če velja

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

Torej, vsaka afina funkcija je konveksna. $\square$

Izrek 3.5. *Vsaka norma je konveksna funkcija.*

Dokaz. Naj bo $||x||_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ za $1 \leq p < \infty$ poljubna norma na $\mathbb{R}^n$ različna od supremum norme. Potem velja

$$||\lambda x + (1-\lambda)y||_p = (\sum_{i=1}^n |\lambda x_i + (1-\lambda)y_i|^p)^{\frac{1}{p}} \leq$$

$$(\sum_{i=1}^{n}|\lambda x_i|^p)^{\frac{1}{p}} + (\sum_{i=1}^{n}|(1-\lambda)y_i|^p)^{\frac{1}{p}} =$$
$$||\lambda x||_p + ||(1-\lambda)y||_p,$$

zaradi trikotniške neenakosti, tj. norma je konveksna.
Analogno temu, za supremum normo $||x||_\infty = \max_k |x_k|$ velja

$$||\lambda x+(1-\lambda)y||_\infty = \max_k|\lambda x_k+(1-\lambda)y_k| \leq \max_k|\lambda x_k|+\max_k|(1-\lambda)y_k| = ||\lambda x||_\infty+||(1-\lambda)y||_\infty,$$

tj. tudi supremum norma je konveksna. Drugače pa lahko samo opazimo da je norma pozitivno homogena in subaditivna, in zaradi tega konveksna funkcija, kot si bomo utrdili kasneje. □

Definicija 3.6. *Efektivna domena* konveksne funkcije f na S, oznaka dom$f$, je projekcija epigrafa funkcije f na $\mathbb{R}^n$:

$$\text{dom} f = \{x \big| \exists \mu, (x,\mu) \in \text{epi} f\} = \{x \big| f(x) < +\infty\}.$$

Efektivna domena je po 2.55 konveksna množica v $\mathbb{R}^n$, kot slika konveksne množice epi$f$ pod linearno transformacijo.

Definicija 3.7. Dimenzija efektivne domene se imenuje *dimenzija* funkcije f.

Definicija 3.8. Konveksna funkcija f je *prava* če je njen epigraf neprazen in če ne vsebuje vertikalnih premic, oz. če je $f(x) < +\infty$ za vsaj en x in $f(x) > -\infty$ za vsak x.

Se pravi, f je prava če in samo če je konveksna množica $C = \text{dom} f$ neprazna in če je omejitev funkcije f na C končna.

Definicija 3.9. Konveksna funkcija je *neprava* če ni prava.

Primer 3.10. Funkcija f na $\mathbb{R}$ definirana z

$$f(x) = \begin{cases} -\infty & \text{če } |x| < 1 \\ 0 & \text{če } |x| = 0 \\ +\infty & \text{če } |x| > 1 \end{cases}$$

je neprava konveksna funkcija.

Po definiciji, f je konveksna na S če in samo če je njen epigraf $\{(x,\mu) \big| x \in S, \mu \in \mathbb{R}, \mu \geq f(x)\}$ konveksna množica v $\mathbb{R}^n$, kar velja če in samo če je

$$(1-\lambda)(x,\mu) + \lambda(y,\nu) = ((1-\lambda)x + \lambda y, (1-\lambda)\mu + \lambda\nu) \in \text{epi} f$$

za vsaka $(x,\mu),(y,\nu) \in \text{epi}f$ in vsak $0 \leq \lambda \leq 1$. Z drugimi besedami, f je konveksna na S če in samo če je $(1-\lambda)x + \lambda y \in S$ in

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)\mu + \lambda\nu,$$

za vsaka $x, y \in S, f(x) \leq \mu \in \mathbb{R}, f(y) \leq \nu \in \mathbb{R}, 0 \leq \lambda \leq 1$. Naslednji dve verziji zgornjega premisleka sta zelo pomembni.

Izrek 3.11. *Naj bo f funkcija iz C v $(-\infty,+\infty]$, kjer je C konveksna množica (npr. $C = \mathbb{R}^n$). Potem je f konveksna na C če in samo če je*

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y), \quad 0 < \lambda < 1,$$

za vsak x in y v C.

Veliko avtorjev privzame neenakost v 3.11 kot *definicijo* konveksnosti funckije f ki preslika konveksno množico C na $(-\infty,+\infty]$. Ker smo hoteli poudariti geometrično razumevanje konveksnih funkcij, smo v tej nalogi izbrali 3.3 kot definicijo konveksnosti.

Izrek 3.12. *Naj bo f funkcija iz $\mathbb{R}^n$ v $[-\infty,+\infty]$. Potem je f konveksna če in samo če je*

$$f((1-\lambda)x + \lambda y) < (1-\lambda)\alpha + \lambda\beta, \quad 0 < \lambda < 1,$$

kadarkoli je $f(x) < \alpha, f(y) < \beta$.

Dokaz. Če je f konveksna velja

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) \leq (1-\lambda)\alpha + \lambda\beta, \quad 0 < \lambda < 1,$$

za $f(x) < \alpha, f(y) < \beta$. Po drugi strani, recimo da za vsak α takšen da je $f(x) < \alpha$, in za vsak β takšen da je $f(y) < \beta$ velja

$$f((1-\lambda)x + \lambda y) < (1-\lambda)\alpha + \lambda\beta, \quad 0 < \lambda < 1.$$

Potem očitno ne moremo imeti

$$f((1-\lambda)x + \lambda y) > (1-\lambda)f(x) + \lambda f(y),$$

za neke $x, y, 0 < \lambda < 1$, saj bi tedaj lahko poiskali $\alpha_0 > f(x)$ in $\beta_0 > f(y)$ da velja

$$f((1-\lambda)x + \lambda y) = (1-\lambda)\alpha + \lambda\beta,$$

in bi nas to privelo v protislovje. Se pravi, za vsak x in vsak y, velja

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y), \quad 0 < \lambda < 1,$$

tj. funkcija f je konveksna. □

Smo prišli do izjemno pomembne neenakosti, ki joj pravimo Jensenova neenakost.

Izrek 3.13 (Jensenova neekanost). *Naj bo f funkcija iz $\mathbb{R}^n$ v $(-\infty, +\infty]$. Potem je f konveksna če in samo če za vsak $m \geq 2$ velja*

$$f(\lambda_1 x_1 + \cdots + \lambda_m x_m) \leq \lambda_1 f(x_1) + \cdots + \lambda_m f(x_m)$$

kadarkoli je $\lambda_1 \geq 0, \ldots, \lambda_m \geq 0$, $\lambda_1 + \cdots + \lambda_m = 1$, in $x_1, \ldots, x_m \in \mathbb{R}^n$.

Dokaz. Dokaz bo snovil na principu matematične indukcije. Jensenova neenakost trivialno velja za $m = 2$, in za $m = 1$: $f(\lambda_1 x_1) = f(1x_1) = f(x_1) \leq 1f(x_1) = \lambda_1 f(x_1)$. Denimo, da neenakost velja za $m = k$. Pokažimo, da bo tedaj veljavna za $m = k + 1$. Naj bodo $x_1, \ldots, x_k, x_{k+1} \in \mathbb{R}^n$ in $\lambda_1 \geq 0, \ldots, \lambda_k \geq 0, \lambda_{k+1} \geq 0$ takšni skalarji da velja $\lambda_1 + \cdots + \lambda_k + \lambda_{k+1} = 1$. Vsaj en od skalarjev $\lambda_1, \ldots, \lambda_{k+1}$ je različen od 1 (v nasprotnem imamo trivialen slučaj). Brez škode za splošnost, privzemimo da je $\lambda_{k+1} < 1$ in definirajmo

$$u = \frac{\lambda_1}{1 - \lambda_{k+1}} x_1 + \cdots + \frac{\lambda_k}{1 - \lambda_{k+1}} x_k.$$

Potem imamo

$$\frac{\lambda_1}{1 - \lambda_{k+1}} + \cdots + \frac{\lambda_k}{1 - \lambda_{k+1}} = 1$$
$$\lambda_1 x_1 + \cdots + \lambda_k x_k + \lambda_{k+1} x_{k+1} = (1 - \lambda_{k+1}) u + \lambda_{k+1} x_{k+1}$$

Kjer je f konveksna, sledi

$$f((1 - \lambda_{k+1}) u + \lambda_{k+1} x_{k+1}) \leq (1 - \lambda_{k+1}) f(u) + \lambda_{k+1} f(x_{k+1})$$

in, po indukcijski hipotezi,

$$f(u) \leq \frac{\lambda_1}{1 - \lambda_{k+1}} f(x_1) + \cdots + \frac{\lambda_k}{1 - \lambda_{k+1}} f(x_k)$$

Odtod končno sledi

$$f(\lambda_1 x_1 + \cdots + \lambda_{k+1} x_{k+1}) \leq \lambda_1 f(x_1) + \cdots + \lambda_{k+1} f(x_{k+1}).$$

Torej, neenakost velja za $m = k + 1$ in po indukciji za vsak pozitivni m. □

Izrek 3.14. *Zvezno odvedljiva funkcija $f : \mathbb{R} \to \mathbb{R}$ je konveksna natanko tedaj ko je $f(y) \geq f(x) + f'(x)(y - x)$ za poljubna x in y iz domene funkcije, oz. ko je graf funkcije f nad njeno pripadajočo tangento.*

Dokaz. Recimo, da je funkcija f konveksna in naj bosta x in y iz domene funkcije. Kot interval v $\mathbb{R}$, domena funkcija f je konveksna množica, sklepamo da za vsak $0 < \lambda \leq 1$ domena funkcije f vsebuje $x + \lambda(y - x)$. Konveksnost funkcije f nam zagotavlja

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

Ko delimo obe strani z λ dobimo

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

oz.

$$f(y) \geq f(x) + \frac{f(x+\lambda(y-x)) + f(x)}{\lambda}.$$

Pustimo $\lambda \to 0$ in ko izračunamo limite končno sklepamo

$$f(y) \geq f(x) + \lim_{\lambda \to 0} \frac{f(x+\lambda(y-x)) + f(x)}{\lambda} = f(x) + f'(x)(y-x).$$

Po drugi strani, recimo da za poljubna x in z iz domene funkcije f (interval) velja $f(y) \geq f(x) + f'(x)(y-x)$. Naj bosta $x \neq y$ in $0 \leq \lambda \leq 1$. Naj bo $z = \lambda x + (1-\lambda)y$. Velja

$$f(x) \geq f(z) + f'(z)(x-z) \quad f(y) \geq f(z) + f'(z)(y-z).$$

Zmnožimo prvo neenakost z λ, drugo z $\lambda - 1$ in jih seštejmo. Dobimo

$$\lambda x + (1-\lambda)y \geq (1-\lambda+\lambda) f(z) + f'(z)(\lambda x - \lambda z + y - \lambda y - z + \lambda z)$$

oz.

$$\lambda x + (1-\lambda)y \geq f(z) + f'(z)(\lambda x + y - \lambda y - \lambda x - y + \lambda y) = f(z),$$

tj. funkcija je konveksna. □

Izrek 3.15. *Naj bo f dvakrat zvezno odvedljiva funkcija na odprtem intervalu (α, β), z realnimi vrednostimi. Potem je f konveksna če in samo če je f'' nenegativna povsod na (α, β).*

Dokaz. Denimo najprej, da je f'' nenegativna povsod na (α, β). Potem je f' nepadajoča na (α, β). Za $\alpha < x < y < \beta$, $0 < \lambda < 1$, bodi $z = (1-\lambda)x + \lambda y$. Po izreku o povprečni vrednosti $\exists \varepsilon \in [x, z]$:

$$f(z) - f(x) = \int_x^z f'(t)dt = f'(\varepsilon)(z-x) \leq f'(z)(z-x).$$

Analogno temu

$$f(y) - f(z) = \int_z^y f'(t)dt \geq f'(z)(y-z)$$

Ker je $z - x = \lambda(y-x)$ in $y - z = (1-\lambda)(y-x)$, imamo

$$f(z) \leq f(x) + \lambda f'(z)(y-x)$$

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

Ko zmnožmo neenakosti z $(1-\lambda)$ in λ, respektivno, in nato jih seštejemo, dobimo

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

Leva strana je prav zaprav $f(z) = f((1-\lambda)x + \lambda y)$, in to dokazuje konveksnost f na (α, β) po 3.11. Po drugi strani, recimo da f'' ni nenegativna na (α, β). Potem bo f'' negativna na podintervalu (α', β') zaradi zveznosti. Po premisleku paralelnem že danem, na (α', β') bi tedaj imeli

$$f(z) - f(x) > f'(z)(z-x),$$

$$f(y) - f(z) < f'(z)(y-z),$$

in torej

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

v tem primeru f ne bi bila konveksna na (α, β), oz. bi prišli do protislovja. □

Primer 3.16. Konveksnost naslednjih funkcij na $\mathbb{R}$ je posledica prejšnjega izreka.

1. $f(x) = e^{\alpha x}$, kjer je $-\infty < \alpha < \infty$;
2. $f(x) = x^p$ če $x \geq 0$ oz. $f(x) = \infty$ če $x < 0$, kjer je $1 \leq p < \infty$;
3. $f(x) = -x^p$ če $x \geq 0$ oz. $f(x) = \infty$ če $x < 0$, kjer je $0 \leq p \leq 1$;
4. $f(x) = x^p$ če $x > 0$ oz. $f(x) = \infty$ če $x \leq 0$, kjer je $-\infty < p \leq 0$;
5. $f(x) = (\alpha^2 - x^2)^{-1/2}$ če $|x| < \alpha$ oz. $f(x) = \infty$, če $|x| \geq \alpha$, kjer je $-\infty < p \leq 0$;
6. $f(x) = -log(x)$ če $x > 0$ oz. $f(x) = \infty$ če $x \leq 0$.

Primer 3.17. Naj bo

$$f(x) = \begin{cases} -logx & \text{če } x > 0, \\ +\infty & \text{če } x \leq 0. \end{cases}$$

Tako definirana funkcija je konveksna. Naj bo $\lambda_1 x_1 + \cdots + \lambda_m x_m$ konveksna kombinacija pozitivnih števil $x_1, \ldots, x_m$. Potem velja

$$-log(\lambda_1 x_1 + \cdots + \lambda_m x_m) \leq -\lambda_1 log x_1 - \cdots - \lambda_m log x_m$$

po Jensenovi neenakosti. Ko zmnožimo z (-1) in izračunamo eksponencial dobimo

$$\lambda_1 x_1 + \cdots + \lambda_m x_m \geq x_1^{\lambda_1} \ldots x_m^{\lambda_m}.$$

Za $\lambda_1 = \ldots = \lambda_m = \frac{1}{m}$,

$$\frac{x_1 + \cdots + x_m}{m} \geq (x_1 \ldots x_m)^{\frac{1}{m}}.$$

To pa je znana neenakost med aritmetično in geometrično sredino.

Izrek 3.18. *Naj bo f dvakrat zvezno odvedljiva funkcija na odprti množici $C \subset \mathbb{R}^n$, tako da je C konveksna in da f zavzame realne vrednosti. Potem je f konveksna na C če in samo če je njena Hessejeva matrika*

$$Q_x = (q_{ij}(x)), \qquad q_{ij}(x) = \frac{\partial^2 f}{\partial \varepsilon_i, \partial \varepsilon_j}(\varepsilon_1, \ldots, \varepsilon_n),$$

pozitivno semi-definitna za vsak $x \in C$.

Dokaz. Konveksnost f na C je ekvivalentna konveksnosti omejitve f na vsaki daljici v C, kar je enako kot konveksnost funkcije $g(\lambda) = f(y + \lambda z)$ na odprtem realnem intervalu $\{\lambda \big| y + \lambda z \in C\}$ za vsak $y \in C$ in $z \in \mathbb{R}^n$. Direkten izračun pokaže, da je

$$g'(\lambda) = \frac{\partial f}{\partial \lambda} f(y + \lambda z) = f'(y + \lambda z)z.$$

in odtod

$$g''(\lambda) = \frac{\partial f}{\partial \lambda} f'(y + \lambda z)z = f''(y + \lambda z)z^2.$$

Se pravi, za $x = y + \lambda z$, $g''(\lambda) = \langle z, Q_x z\rangle$, kjer je Q_x ustrezna Hessejeva matrika. Po 3.15, g je konveksna za vsak $y \in C$ in $z \in \mathbb{R}^n$ če in samo če je $\langle x, Q_x z\rangle \geq 0$ za vsak $x \in C$ in $z \in \mathbb{R}^n$. □

Obstajajo nekatere zelo uporabne korespondence med konveksnimi množicami in konveksnimi funkcijami. Najbolj enostavna vsaki množici C v $\mathbb{R}^n$ priredi indikatorsko funkcijo $\delta(\cdot|C)$.

Definicija 3.19. *Indikatorsko funkcijo $\delta(\cdot|C)$* množice C v $\mathbb{R}^n$ definirano kot

$$\delta(x|C) = \begin{cases} 0 & \text{če } x \in C \\ +\infty & \text{če } x \notin C. \end{cases}$$

Epigraf indikatorske funkcije je polcilinder z prečnim prerezom C.

Lema 3.20. *Množica C je konveksna natanko tedaj ko je njena pripadajoča indikatorska funkcija konveksna na $\mathbb{R}^n$.*

Dokaz. Če je C konveksna, potem je $\lambda x + (1-\lambda)y \in C$ za $0 \leq \lambda \leq 1$ in $x, y \in C$, torej

$$\delta(\lambda x + (1-\lambda)y|C) = 0.$$

Prav tako je

$$\lambda\delta(x|C) + (1-\lambda)\delta(y|C) = 0.$$

Če velja $x, y \notin C$, $x \in C, y \notin C$ ali $x \notin C, y \in C$ potem velja

$$\delta(\lambda x + (1-\lambda)y|C) = +\infty = \lambda\delta(x|C) + (1-\lambda)\delta(y|C).$$

Se pravi, indikatorska funkcija je konveksna. Po drugi strani, če je indikatorska funckija konveksna, potem za $0 \leq \lambda \leq 1$ in $x, y \in C$ imamo

$$\delta(\lambda x + (1-\lambda)y|C) = \lambda\delta(x|C) + (1-\lambda)\delta(y|C) = 0,$$

kar pomeni da je $\lambda x + (1-\lambda)y \in C$, tj. C je konveksna. □

Definicija 3.21. *Podporna funkcija* $\delta^*(\cdot|C)$ konveksne množice C v $\mathbb{R}^n$ je definirana z

$$\delta^*(x|C) = \sup\{\langle x, y\rangle \big| y \in C\}.$$

Lema 3.22. *Podporna funkcija konveksne množice C v $\mathbb{R}^n$ je konveksna.*

Dokaz. Opazimo, da je podporna funkcija pozitivno homogena in subaditivna. Kot si bomo ogledali v 3.37, to pomeni da je prav tako konveksna. □

Definicija 3.23. *Funkcional Minkowskega* konveksne množice C definiramo kot

$$\gamma(x|C) = \inf\{\lambda \geq 0 \big| x \in \lambda C\}, \quad C \neq \emptyset.$$

Lema 3.24. *Funkcional Minkowskega je konveksna funkcija.*

Dokaz. Opazimo, da je funkcional Minkowskega pozitivno homogena in subaditivna funkcija. Kot smo že omenili v prejšnjem dokazu, si bomo ogledali v 3.37 zakaj to pomeni da je prav tako konveksna funkcija. □

Definicija 3.25. *Evklidska razdalja* $d(\cdot, C)$ se definira kot

$$d(x, C) = \inf\{|x-y| \big| y \in C\}. \tag{3.1}$$

Lema 3.26. *Evklidska razdalja je konveksna funkcija.*

Dokaz. Najprej opazimo, da je

$$d(\lambda x_1 + (1-\lambda)x_2, C) = \inf\{|\lambda x_1 + (1-\lambda)x_2 - y| \big| y \in C\}.$$

Po drugi strani je

$$\lambda d(x_1, C) + (1-\lambda)d(x_2, C) = \lambda \inf\{|x_1 - y| \big| y \in C\} + (1-\lambda)\inf\{|x_2 - y| \big| y \in C\}.$$

Očitno,

$$\inf\{|\lambda x_1 + (1-\lambda)x_2 - y| \big| y \in C\} \leq \inf\{|\lambda(x_1 - y)| \big| y \in C\} + \inf\{|(1-\lambda)(x_2 - y)| \big| y \in C\},$$

torej je evklidska razdalja konveksna. □

Izrek 3.27. *Za poljubno konveksno funkcijo f in $\alpha \in [-\infty, +\infty]$, sta množici $\{x \big| f(x) < \alpha\}$ in $\{x \big| f(x) \leq \alpha\}$ konveksni.*

Dokaz. Množica $\{x \big| f(x) \leq \alpha\}$ je projekcija $\mathbb{R}^n$ na presek epi$f$ in hiperravnine $\{(x, \mu) \big| \mu = \alpha\}$ v $\mathbb{R}^{n+1}$, torej je $\{x \big| f(x) \leq \alpha\}$ konveksna. Prva množica je očitno konveksna po 3.12. Privzamemo $\beta = \alpha$ in dobimo da je funkcija f konveksna če in samo če je $f((1-\lambda)x + \lambda y) < (1-\lambda)\alpha + \lambda\alpha = \alpha$ za $f(x) < \alpha$ in $f(y) < \alpha$. Torej, ker za x in y iz množice $\{x \big| f(x) < \alpha\}$ velja $f(x) < \alpha$ in $f(y) < \alpha$, in ker je dana funkcija f konveksna, velja tudi $f((1-\lambda)x + \lambda y) < \alpha$ oz. $(1-\lambda)x + \lambda y \in \{x \big| f(x) < \alpha\}$ tj. smo prišli do konveksnosti prve množice. □

Posledica 3.28. *Naj bo f_i konveksna funkcija na $\mathbb{R}^n$ in α_i realno število za vsak $i \in \mathbb{I}$, kjer je $\mathbb{I}$ poljubna indeksna množica. Potem je*

$$C = \{x \big| f_i(x) \leq \alpha_i, \quad \forall i \in \mathbb{I}\}$$

konveksna množica.

Dokaz. Množica C je presek množic $C_i = \{x \big| f_i(x) \leq \alpha_i\}$, ki so po 3.27 konveksne. Kot smo že utrdili, presek konveksnih množic je konveksna množica. □

Definicija 3.29. Funkcija f na $\mathbb{R}^n$ je *pozitivno homogena* če za vsak x velja

$$f(\lambda x) = \lambda f(x), \quad 0 < \lambda < \infty.$$

Lema 3.30. *Pozitivna homogenost funkcije f je ekvivalentna pogoju, da je epigraf funkcije f stožec v $\mathbb{R}^n$.*

Dokaz. Če je funkcija pozitivno homogena, potem za poljuben $\lambda > 0$ velja

$$(x, t) \in \text{epi}f \Leftrightarrow f(x) \leq t \Leftrightarrow \lambda f(x) = f(\lambda x) \leq \lambda t \Leftrightarrow (ax, at) \in \text{epi}f.$$

Po drugi strani, recimo da je epigraf stožec. To pomeni da za poljuben $\lambda > 0$ velja: če je $(x, t) \in \text{epi}f$ potem je $(\lambda x, \lambda t) \in \text{epi}f$. Očitno, $(x, f(x)) \in \text{epi}f$, tako da je $(\lambda x, \lambda f(x)) \in \text{epi}f$, kar pomeni $f(\lambda x) \leq \lambda f(x)$. Po tej poti sklepamo, da $(\lambda x, f(\lambda x)) \in \text{epi}f$. Torej, če je epigraf stožec, potem je $(x, \frac{f(\lambda x)}{\lambda}) \in \text{epi}f$ oz. $f(x) \leq \frac{f(\lambda x)}{\lambda}$. Končno, iz dve neenakosti sledi $f(x) \leq \frac{f(\lambda x)}{\lambda} \leq f(x)$, tj. funkcija je pozitivno homogena. □

Primer 3.31. Funkcija $g(x) = |x|$ ni linearna, je pa pozitivno homogena in konveksna.

Izrek 3.32. *Pozitivno homogena funkcija f iz $\mathbb{R}^n$ na $(-\infty, +\infty]$ je konveksna če in samo če je*

$$f(x + y) \leq f(x) + f(y)$$

za vsak $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$.

Dokaz. Dokaz je zelo enostaven, namreč ko upoštevamo trikotniško neenakost in homogenost imamo za poljubna $x, y \in \mathbb{R}^n$ in $0 \leq \lambda \leq 1$

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

Se pravi, subaditivnost in konveksnost pozitivno homogene funkcije sta ekvivalentni. □

Opazimo, da iz tega izreka sledi konveksnost norme ki je sama pozitivno homogena in subaditivna funkcija.

Posledica 3.33. *Če je f pozitivno homogena prava konveksna funkija, potem velja*

$$f(\lambda_1 x_1 + \cdots + \lambda_m x_m) \leq \lambda_1 f(x_1) + \cdots + \lambda_m f(x_m)$$

kadarkoli je $\lambda_1 > 0, \ldots, \lambda_m > 0$.

Posledica 3.34. *Če je f pozitivno homogena prava konveksna funkcija, potem je*

$$f(-x) \geq -f(x)$$

za vsak x.

Dokaz. $f(x) + f(-x) \geq f(x - x) = f(0) \geq 0$ (zadnja neenakost sledi iz pozitivne homogenosti). □

Izrek 3.35. *Naj bo f pozitivno homogena prava konveksna funkcija in L dan podprostor z bazo sestavljeno iz vektorjev $b_1, \ldots, b_m$. Funkcija f je linearna na L če in samo če je $f(-x) = -f(x)$ za vsak $x \in L$, kar velja če je $f(b_i) = -f(b_i)$ za vse bazne vektorje.*

Dokaz. Pri danih pogojih velja $f(\lambda_i b_i) = \lambda_i f(b_i)$ za vsak $\lambda_i \in \mathbb{R}$: namreč, pozitivna homogenost nam zagotavlja ta pogoj za vsak $\lambda_i > 0$, ker je $f(-x) = -f(x)$ je tudi $f(-\lambda_i x) = -\lambda_i f(x)$ in pogoj velja tudi za vsak $\lambda_i < 0$, v primeru ko je $\lambda_i = 0$ je $f(0) = 0$, saj je $f(0) = -f(0)$. Za poljuben $x = \lambda_i b_i + \cdots + \lambda_m b_m \in L$ velja

$$f(\lambda_1 b_1) + \cdots + f(\lambda_m b_m) \geq f(x) = -f(-x)$$

$$\geq -(f(-\lambda_1 b_1) + \cdots + f(-\lambda_m b_m)) = f(\lambda_1 b_1) + \cdots + f(\lambda_m b_m)$$

po 3.32 in 3.34. Odtop pa

$$f(x) = f(\lambda_1 b_1) + \cdots + f(\lambda_m b_m) = \lambda_1 f(b_1) + \cdots + \lambda_m f(b_m).$$

Torej, f je linearna na L, in v posebnem drži $f(-x) = -f(x)$ za vsak $x \in L$. □

Izrek 3.36. *Naj bosta f_1 in f_2 pravi konveksni funkciji na $\mathbb{R}^n$. Tedaj je $f_1 + f_2$ konveksna funkcija.*

Dokaz. Če sta f_1 in f_2 pravi konveksni funkciji na $\mathbb{R}^n$, potem je po 3.11

$$f_1((1-\lambda)x + \lambda y) \leq (1-\lambda)f_1(x) + \lambda f_1(y)$$

$$f_2((1-\lambda)x + \lambda y) \leq (1-\lambda)f_2(x) + \lambda f_2(y).$$

Torej,

$$(f_1 + f_2)((1-\lambda)x + \lambda y) \leq (1-\lambda)f_1(x) + \lambda f_1(y) + (1-\lambda)f_2(x) + \lambda f_2(y)$$

oz.

$$(f_1 + f_2)((1-\lambda)x + \lambda y) \leq (1-\lambda)(f_1 + f_2)(x) + \lambda(f_1 + f_2)(y)$$

tj. $f_1 + f_2$ je konveksna funkcija. □

Izrek 3.37. *Supremum po točkah poljubne kolekcije konveksnih funkcij je konveksna funkcija.*

Dokaz. Spomnimo se, da je presek poljubne kolekcije konveksnih množic konveksna množica. Bodi $f(x) = \sup_{i \in \mathbb{I}} f_i(x)$. Trdimo, da je

$$\operatorname{epi}(f) = \bigcap_{i \in \mathbb{I}} \operatorname{epi}(f_i).$$

Ker so f_i konveksne funkcije, so $\operatorname{epi}(f_i)$ konveksne množice, torej je po 2.26 tudi njihov presek konveksna množica, se pravi $\operatorname{epi}(f)$ je konveksna množica in odtod je f konveksna funkcija. □

Iz tega izreka sledi konveksnost podporne funkcije $\delta^*(\cdot|C)$ množice C v $\mathbb{R}^n$, ker je podporna funkcija po definiciji supremum po točkah določene kolekcije linearnih funkcij, namreč funkcij $\langle \cdot, y \rangle$ ko y zavzame različne vrednosti po množici C.

4 Relativna odprtja, zaprtja, podporne in konjugirane funkcije konveksnih funkcij

Poleg bolj znanih topoloških definicij kot so evklidska razdalja med dvema točkama, evklidska metrika, evklidska enotna krogla, krogla sa polmerom $\varepsilon > 0$ in centrem a, odprtost in zaprtost, notranjost in zaprtje množice, si bomo v tem poglavju ogledali še nekatere topološke definicije in lastnosti bolj značilne in specifične za konveksne množice.

Definicija 4.1. Za poljuben $a \in \mathbb{R}^n$ in $\varepsilon > 0$, (odprta) krogla s polmerom ε in centrom a je

$$B_\varepsilon(a) = \{x \big| d(x,a) < \varepsilon\}.$$

Če je $d(x,a) \leq \varepsilon$, potem pravimo, da je $B_\varepsilon(a)$ zaprta krogla s polmerom ε in centrom a.

Spomnimo se, da notranjost in zaprtje množice C v $\mathbb{R}^n$ lahko podamo z

$$\mathrm{cl}C = \bigcap \{x \big| \exists y \in C, d(x,y) \leq \varepsilon\} \quad \text{za} \quad \varepsilon > 0$$

$$\mathrm{int}C = \{x \big| \exists \varepsilon > 0 : \quad B_\varepsilon(x) \subset C\}.$$

Spomnimo se, da lahko razumemo notranjost množice C kot unijo vseh odprtih množic ki so vsebovane v C, namreč kot največjo odprto množico vsebovano v množici C, zaprtje množice C pa kot najmanjšo zaprto množico ki vsebuje množico C, oz. kot presek vseh zaprtih množic ki vsebujejo množico C.

Definicija 4.2. *Relativna notranjost* konveksne množice C v $\mathbb{R}^n$ je množica

$$\mathrm{ri}C = \{x \in \mathrm{aff}C \big| \exists \varepsilon > 0, B_\varepsilon(x) \cap (\mathrm{aff}C) \subset C\}.$$

Relativno notranjost označujemo ri$C$ in definiramo kot notranjost ki jo dobimo ko C posmatramo kot podmnožico njene afine ogrinjače aff$C$; ri$C$ sestoji iz točk $x \in \mathrm{aff}C$ za katere obstaja $\varepsilon > 0$, tako da je $y \in C$ kadarkoli je $y \in \mathrm{aff}C$ in $d(x,y) \leq \varepsilon$. Očitno, velja $\mathrm{ri}C \subset C \subset \mathrm{cl}C$.

Definicija 4.3. Razliko cl$C \backslash$ri$C$ imenujemo *relativna meja* množice C. Pravimo, da je množica C *relativno odprta* če je ri$C = C$.

Primer 4.4. Za n-dimenzionalno konveksno množico velja po definiciji aff$C = \mathbb{R}^n$ in torej ri$C =$ int$C$.

Primer 4.5. Inkluzija $C_1 \supset C_2$ ne pomeni nujno da je ri$C_1 \supset$ riC_2. Npr. če je C_1 kocka v $\mathbb{R}^3$ in C_2 ena izmed stranic kocke C_1, sta riC_1 in riC_2 neprazni in disjunktni: aff$C_1 = \mathbb{R}^3$ in aff$C_2 = \mathbb{R}^3 \times 0$. Ker je ri$C_1 = \{x \in \mathbb{R}^3 \big| \exists \varepsilon > 0 : B_\varepsilon(x) \cap \mathbb{R}^3 \subset C_1\}$, in ri$C_2 = \{x \in \mathbb{R}^3 \times 0 \big| \exists \varepsilon > 0 : B_\varepsilon(x) \cap \mathbb{R}^3 \times 0 \subset C_2\}$ sta očitno neprazni in disjunktni!

Primer 4.6. Afine množice so po definiciji relativno odprte, poleg tega pa tudi zaprte.

Zaprtja in relativne notranjosti se ohranijo z premikanjem. Bolj splošno, poljubna injektivna afina transformacija prostora $\mathbb{R}^n$ nase jih ohrani. Namreč, injektivna afina transformacija ohranja afine ogrinjače in je zvezna v obe smeri, ker so komponente slike poljubnega vektorja, linearne oz. afine funkcije komponent začetnega vektorja. Se iskaže, da to lahko uporabljamo kot orodje pri dokazovanju, na naslednji nači. Recimo da je C m-dimenzionalna konveksna množica v $\mathbb{R}^n$. Potem obstaja injektivna afina transformacija T množice $\mathbb{R}^n$ nase, ki preslika aff$C$ na podprostor

$$L = \{x = (\varepsilon_1, \ldots, \varepsilon_m, \varepsilon_{m+1}, \ldots, \varepsilon_n) \big| \varepsilon_{m+1} = 0, \ldots, \varepsilon_n = 0\} = \mathbb{R}^m + 0_{n-m}$$

L si lahko predstavljamo kot kopijo prostora $\mathbb{R}^m$. Na ta način si lahko olajšamo delo tako da pri dokazovanju splošne konveksne množice reduciramo na konveksne množice ki imajo celi prostop kot svoj afina ogrinjača.

Od bistvenega pomena je naslednja lastnost zaprtja in relativne notranjosti konveksnih množic.

Izrek 4.7. *Če je C konveksna množica v $\mathbb{R}^n$, in $x \in$ ri$C$, $y \in$ cl$C$. Potem za $0 \leq \lambda < 1$ velja, da $(1-\lambda)x + \lambda y$ pripada ri$C$, in torej v posebnem C.*

Dokaz. Lahko si olajšajmo delo tako, da privzamemo da je C n-dimenzionalna množica in v posebnem ri$C =$ int$C$. Naj bo $\lambda \in [0, 1)$. Pokažimo, da je za nek $\varepsilon > 0$, $B_\varepsilon((1-\lambda)x + \lambda y) \subset C$. Ker je $y \in$ cl$C$, za vsak $\varepsilon > 0$ velja $\exists z \in C \cap B_\varepsilon(y)$, torej $|y - z| < \varepsilon$ in zato $y \in B_\varepsilon(z)$ ter seveda $z \in C$, odtod pa končno sledi $y \in C + B_\varepsilon$. Torej za vsak $\varepsilon > 0$

$$B_\varepsilon((1-\lambda)x + \lambda y) \subset (1-\lambda)x + \lambda(C + \varepsilon B) + B_\varepsilon$$

$$= (1-\lambda)[x + \varepsilon(1+\lambda)(1-\lambda)^{-1}B] + \lambda C.$$

(Smo uporabili dejstvo, da za množico C in kroglo B velja $\lambda(C + B) = \lambda C + \lambda B$. Saj je po eni strani $\lambda(C + B) = \lambda \cup \{y + B \big| y \in C\} = \cup\{\lambda(y + B) \big| y \in C\}$, in po drugi strani $\lambda C + \lambda B = \cup\{\lambda y + \lambda B \big| y \in C\}$, je očitno zakaj to velja.) Ker smo predpostavili da je $x \in \mathrm{ri}C$, in torej v $\mathrm{int}C$ za n-dimenzionalno množico C, je za dovolj majhen ε, slednja množica vsebovana v $(1-\lambda)C + \lambda C = C$. □

Izrek 4.8. *Naj bo C poljubna konveksna množica v $\mathbb{R}^n$. Potem sta clC in riC konveksni množici v $\mathbb{R}^n$ ki imata isto afino ogrinjačo, in s tem in isto dimenzijo, kot C.*

Dokaz. Kot linearna kombinacija konveksnih množic, je množica $C + \varepsilon B$ konveksna za poljuben $\varepsilon > 0$. Ker je presek kolekcije takšnih množic $\mathrm{cl}C$, sklepamo da je $\mathrm{cl}C$ konveksna množica. Dalje, je $\mathrm{cl}C \subset \mathrm{aff}C$, in ker je $\mathrm{aff}C$ afina množica ki vsebuje C, $\mathrm{aff}(\mathrm{cl}C)$ pa najmanjša afina množica, ki vsebuje $\mathrm{cl}C$), sledi $\mathrm{aff}(\mathrm{cl}C) \subset \mathrm{aff}C$. Se pravi, $\mathrm{aff}(\mathrm{cl}C)$ in $\mathrm{aff}(C)$ sovpadata. Po 4.7 sledi, da je $\mathrm{ri}C$ konveksna množica. Zadošča še pokazati, da ko je C n-dimenzionalna za $n > 0$, je $\mathrm{int}C \neq \emptyset$. Po 2.38, n-dimenzionalna konveksna množica vsebuje n-dimenzionalni simpleks. Pokažimo, da takšen simpleks S ima neprazno notranjost. Denimo, da so oglišča simpleksa S vektorji $x_0, \ldots, x_n$, se pravi

$$S = \{(\varepsilon_1, \ldots, \varepsilon_n) \big| \varepsilon_j \geq 0, \varepsilon_1 + \cdots + \varepsilon_n \leq 1\}.$$

Tedaj obstaja afina bijekcija T tako da velja: $Tx_0 = 0, Tx_1 = e_1 = (1, 0, \ldots, 0), \ldots, Tx_n = e_n = (0, \ldots, 0, 1)$, torej je $T(S) = conv\{0, e_1, \ldots, e_n\}$. Sledi, da je za $\varepsilon = \frac{1}{2^n}$, $B_\varepsilon(\varepsilon, \ldots, \varepsilon) \subset T(S)$. Torej, dani simpleks ima neprazno notranjost, kar pomeni $\mathrm{int}C \neq \emptyset$. □

Izrek 4.9. *Za poljubno konveksno množico C v $\mathbb{R}^n$, $cl(riC) = clC$ in $ri(clC) = riC$.*

Dokaz. Ker je $\mathrm{ri}C \subset C$, trivialno velja da je $\mathrm{cl}(\mathrm{ri}C) \subset \mathrm{cl}C$, ker vemo da v splošnem $A \subset B$ pomeni $cl(A) \subset cl(B)$. Po drugi strani, za poljuben $y \in \mathrm{cl}C$ in poljuben $x \in \mathrm{ri}C$, je odprta daljica med x in y je v celoti vsebovana v $\mathrm{ri}C$, razen morda y, po 4.7. Torej $y \in \mathrm{cl}(\mathrm{ri}C)$. Ugotovili smo, da je $\mathrm{cl}(\mathrm{ri}C) = \mathrm{cl}C$. Vemo, da je $C \subset \mathrm{cl}C$, in da afini ogrinjači C in $\mathrm{cl}C$ sovpadata. Odtod sledi da je $\mathrm{ri}C \subset \mathrm{ri}(\mathrm{cl}C)$. Po drugi strani, naj bo $z \in \mathrm{ri}(\mathrm{cl}C)$. Pokažimo, da je $z \in \mathrm{ri}C$. Naj bo x poljubna točka v $\mathrm{ri}C$. Oglejmo si premico skozi x in z (če je $x = z$, trivialno, nimamo premico vendar točko). Naj bo $\mu > 0$ takšen da je $1 - \mu$ dovolj majhen, potem je točka

$$y = (1-\mu)x + \mu z = z + (1-\mu)(x - z)$$

na tej premici, in je vsebovana v $\mathrm{ri}(\mathrm{cl}C)$ in posledično v $\mathrm{cl}C$. Za takšen y, lahko z izražamo v obliki $(1-\lambda)x + \lambda y$ za $0 < \lambda < 1$. Po 4.7, $z \in \mathrm{ri}C$. □

Naslednja karakterizacija relativnih notranjosti je pogosto zelo uporabna. V prihodnjih poglavjih jo bomo uporabljali pri dokazovanju bolj kompleksnih izrekov.

Izrek 4.10. *Za neprazno konveksno množico C v $\mathbb{R}^n$ velja da je $z \in riC$ če in samo če za vsak $x \in C$ obstaja $\mu > 1$ tako da $(1-\mu)x + \mu z$ pripada C.*

Dokaz. V bistvu, ta pogoj pravi da se vsaka daljica v C ki ima z kot eno izmed krajišč lahko podaljša skozi z, brez da bi šli ven iz C. Če je $z \in \text{ri}C$, to seveda drži. Obratno, recimo da z zadošča tem pogoju. Po 4.8, $\text{ri}C \neq \emptyset$. Sledi, da $\exists x \in \text{ri}C$. Naj bo y pripadajoča točka $(1-\mu)x + \mu z \in C$, kjer je $\mu > 1$. Potem je $z = (1-\lambda)x + \lambda y$, kjer je $0 < \lambda = \mu^{-1} < 1$. Končno, po 4.7 je $z \in \text{ri}C$. □

Posledica 4.11. *Naj bo C konveksna množica v $\mathbb{R}^n$. Potem je $z \in intC$ če in samo če za vsak $y \in \mathbb{R}^n$ obstaja $\varepsilon > 0$ tako da je $z + \varepsilon y \in C$.*

Prejšnji štiri izreki, skupaj z številnimi posledicami, so med najbolj pomembnimi topološkimi dejstvi konveksnih množic in funkcij. Zdaj pa bomo obravnavali zaprtje konveksnih funkcij. Spomnimo se, da v linearni algebri linearnost funkcije pomeni zveznost funkcije. Kljub temu, da so stvari z konveksnimi funkcijami bolj zapletene, konveknost funkcije še vedno pomeni veliko topoloških lastnosti, kar temelji na uporabi teorije zaprtja in relativne notranjosti konveksnih množic na epigrafe konveksnih funkcij. Pokazali bomo, da je spodnja polzveznost, kot jo bomo definirali, izjemno pomembna za konveksne funkcije, in da s pomočjo enostavne operacije zaprtje od vsake pravilne konveksne funkcije lahko naredimo navzdol polzvezno funkcijo, tako da redefiniramo njeno vrednost v točkah na relativni meji njene efektivne domene.

Definicija 4.12. Pravimo, da je funkcija $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ *navzdol polzvezna* v točki x v S če je $f(x) \leq \lim_{i\to\infty} f(x_i)$, za vsako zaporedje $x_1, x_2, \dots$ v S takšno da x_i konvergira k x in da limita zaporedja $f(x_1), f(x_2), \dots$ obstaja v $[-\infty, +\infty]$. Ta pogoj je ekvivalenten naslednjemu:

$$f(x) \leq \liminf_{y\to x} f(y) = \lim_{\varepsilon\downarrow 0}(\inf\{f(y)\big||y-x| \leq \varepsilon\}).$$

Analogno temu, pravimo da je f *navzgor polzvezna* v točki x če velja:

$$f(x) \geq \limsup_{y\to x} f(y) = \lim_{\varepsilon\downarrow 0}(\sup\{f(y)\big||y-x| \leq \varepsilon\})$$

kar je ekvivalentno pogoju

$$\{x\big|f(x) > \lambda)\text{je odprta za vsak}\lambda \in \mathbb{R}.$$

Primer 4.13. Funkcija je lahko navzgor ali navzdol polzvezna, brez da bi bila levo ali desno zvezna. Naj bo

$$f(x) = \begin{cases} 1 & \text{če } x < 1, \\ 2 & \text{če } x = 1, \\ \frac{1}{2} & \text{če } x > 1. \end{cases}$$

Dana funkcija je navzgor polzvezna v $x = 1$ čeprav ni niti levo niti desno zvezna. Limita po levi strani je enaka 1, limita po desni strani je $\frac{1}{2}$, obe pa sta manjši od funkcijske vrednosti, tj. od 2.

Izrek 4.14. *Naj bo f poljubna funkcija iz $\mathbb{R}^n$ v $[-\infty, +\infty]$. Naslednje trditve so ekvivalentne:*

1. *f je navzdol polzvezna povsod na $\mathbb{R}^n$;*
2. *$\{x \big| f(x) \leq \alpha\}$ je zaprta za vsak $\alpha \in \mathbb{R}$;*
3. *epi$f$ je zaprta množica v $\mathbb{R}^{n+1}$.*

Dokaz. Pokažimo najprej, da iz 1. sledi 2. Bodi $(x_n, y_n) \in \text{epi}f$ zaporedje, ki konvergira proti (x, y). Tedaj $y_n \geq f(x_n)$ za vsak n, in po definiciji polzveznosti v točki x tudi $f(x_n) \geq f(x)$, torej $y_n \geq f(x_n) \geq f(x)$ za vsak n in je v limiti $y \geq f(x)$. Navzdol polzveznost v x se lahko izraža v obliki pogoja da je $\mu \geq f(x)$ kadarkoli je $\mu = \lim \mu_i$ in $x = \lim x_i$ za zaporedja $\mu_1, \mu_2, \ldots,$ in $x_1, x_2, \ldots,$ takšna da velja $\mu_i \geq f(x_i)$ za vsak i. Ampak ta pogoj je ekvilenten 3.! (Če privzamemo $\alpha = \mu = \mu_1 = \mu_2 = \ldots$, potem tudi 2. velja.) Po drugi strani, recimo da velja 2. in naj zaporedje x_i konvergira k x in $f(x_i)$ k μ. Za vsak realni $\alpha > \mu$, $f(x_i)$ mora nujno biti manjše kot α, in zaradi tega

$$x \in \text{cl}\{y \big| f(y) \leq \alpha\} = \{y \big| f(y) \leq \alpha\}.$$

Sledi $f(x) \leq \mu$ tj. iz 2. sledi 1. □

Definicija 4.15. Za poljubno funkcijo f na $\mathbb{R}^n$, obstaja največja navzdol polzvezna funkcija majorizirana z f, in sicer funkcija katere epigraf je zaprtje epigrafa funkcije f v prostoru $\mathbb{R}^n$. Tej funkciji rečemo *navzdol polzvezna ogrinjača* funkcije f.

Definicija 4.16. *Zaprtje* konveksne funkcije f definiramo kot navzdol polzvezno ogrinjačo funkcije f če f nikjer ni enaka $-\infty$. Če za nek x velja $f(x) = -\infty$, potem zaprtje funkcije f definiramo kot konstantno funkcijo $-\infty$.

Zaprtje konveksne funkcije f je konveksna funkcija, za katero običajno uporabljamo oznako cl$f$.

Primer 4.17. Za pravo konveksno funkcijo f imamo po definiciji $\text{epi}(\text{cl}f) = \text{cl}(\text{epi}f)$.

Definicija 4.18. Pravimo, da je konveksna funkcija f *zaprta* če je $\text{cl}f = f$.

Primer 4.19. Zaprtost funkcije in navzdol polzveznost sta prav zaprav ekvivalentni za pravo konveksno funkcijo. Nasproti temu, sta edine nepravilni zaprti konveksni funkciji konstantni funkciji $+\infty$ in $-\infty$.

Primer 4.20. Konveksna funkcija je lahko zaprta četudi njena efektivna domena ni zaprta, recimo $f(x) = \frac{1}{x}$ kadar je $x > 0$ in $f(x) = \infty$ kadar je $x \leq 0$, definirana na $\mathbb{R}$.

Primer 4.21. Naj bo f konveksna funkcija na $\mathbb{R}$, tako da bo $f(x) = 0$ za $x > 0$ in $f(x) = \infty$ za $x \leq 0$. v tem primeru cl$f$ sovpada povsod z f razen v koordinatnem izhodišču, v katerem zavzame vrednost 0 namesto $+\infty$.

Izrek 4.22. *Če je f neprava konveksna funkcija, potem je $f(x) = -\infty$ za vsak $x \in$ ri(dom$f$). Sklepamo, da je neprava konveksna funkcija nujno neskončna razen morda v točkah na relativni meji njene efektivne domene.*

Dokaz. Če efektivna domena f vsebuje vsaj eno točko, potem po definiciji neprave konveksne funkcije vsebuje tudi točke v katerih f zavzame vrednost $-\infty$. Naj bo u takšna točka in $x \in \text{ri}(\text{dom}f)$. Po 4.10 obstaja $\mu > 1$ tako da je $y = (1-\mu)u + \mu x \in \text{dom}f$. Imamo $x = (1-\lambda)u + \lambda y$, kjer je $0 < \lambda = \mu^{-1} < 1$. Po 3.12

$$f(x) = f((1-\lambda)u + \lambda y) < (1-\lambda)\alpha + \lambda\beta$$

za poljuben $\alpha > f(u)$ in $\beta > f(y)$. Ker je $f(u) = -\infty$ in $f(y) = f((1-\mu)u + \mu x) = (1-\mu)f(u) + \mu f(x) < +\infty$, potem je $f(x)$ nujno enaka $-\infty$. □

Posledica 4.23. *Navzdol polzvezna neprava konveksna funkcija ne more imeti končne vrednosti.*

Dokaz. Ker je f neprava konveksna funkcija, velja $f(x) = -\infty$ za vsak $x \in \text{ri}(\text{dom}f)$. Zaradi navzdol polzveznosti, mora množica točk x kjer je $f(x) = -\infty$ vsebovati cl(ri(dom$f$)), saj je za $f(x) = -\infty$ navzdol polzveznost ekvivalentna zaprtosti funkcije (cl$f = f$). Odtod sledi $\text{cl}(\text{ri}(\text{dom}f)) = \text{cl}(\text{dom}f) \supset \text{dom}f$ po 4.9. □

Posledica 4.24. *Za nepravo konveksno funkcijo f velja, da je cl$f$ zaprta neprava konveksna funkcija ki sovpada z f na ri(dom$f$).*

Posledica 4.25. *Naj bo f konveksna funkcija katere efektivna domena je relativno odprta. V tem primeru velja $f(x) > -\infty$ za vsak x ali $f(x) = +\infty$ za vsak x.*

Najbolj pomembna topološka lastnost konveksnih množic v prostoru $\mathbb{R}^n$ je ta, da obstaja močna zveza med zaprtjem in relativno notranjostjo konveksnih množic. Preden pridemo do tega opazimo, da je notranjost epigrafa funkcije f pomembna pri analizi zaprtja funkcije f, saj zaprtje pravilne konveksne funkcije f temelji na zaprtju epigrafa funkcije f.

Lema 4.26. *Za poljubno konveksno funkcijo f, ri(epi$f$) sestoji iz parov (x,μ) takšnih da $x \in$ ri(dom$f$) in $f(x) < \mu < \infty$.*

Dokaz. Zadošča pokazati, da velja

$$\text{int}(\text{epi}f) = \{(x,\mu) \big| x \in \text{int}(\text{dom}f), f(x) < \mu < \infty\}.$$

Očitno, vsebovanost "$\subset$" velja. Se pravi, samo še vsebovanost "$\supset$" potrebuje verifikacijo. Naj bo $\bar{x} \in \text{int}(\text{dom} f)$ in $m\bar{u}$ realno število takšno da je $m\bar{u} > f(\bar{x})$. Bodi $a_1, \dots, a_r$ točki v dom$f$ takšne da je $\bar{x} \in \text{int}P$, kjer je $P = \text{conv}\{a_1, \dots, a_r\}$ in naj bo $\alpha = \max\{f(a_i)\big| i = 1, \dots, r\}$. Poljuben $x \in P$ lahko izražamo kot konveksno kombinacijo

$$x = \lambda_1 a_1 + \cdots + \lambda_r a_r, \quad \lambda_i \geq 0, \quad \lambda_1 + \cdots + \lambda_r = 1,$$

in zaradi tega imamo po Jensenovi neenakosti

$$f(x) \leq \lambda_1 f(a_1) + \cdots + \lambda_r f(a_r) \leq (\lambda_1 + \cdots + \lambda_r)\alpha = \alpha.$$

Potem je odprta množica $\{(x, \mu)\big| x \in \text{int}P, \alpha < \mu < \infty\}$ vsebovana v epi$f$. Za vsak $\mu > \alpha$ imamo $(\bar{x}, \mu) \in \text{int}(\text{epi} f)$ in sledi, da $(\bar{x}, \bar{\mu})$ lahko posmatramo kot točko v relativni notranjosti vertikalnega premičnega segmenta v epi$f$ ki sreča int(epi$f$). To pa po 4.7 pomeni $(\bar{x}, \bar{\mu}) \in \text{int}(\text{epi} f)$. □

Smo prišli do najbolj pomembnega izreka o cl$f$.

Izrek 4.27. *Recimo, da je f prava konveksna funkcija na $\mathbb{R}^n$. Potem velja, da je clf zaprta prava konveksna funkcija. Še več, clf sovpada z f razen morda v točkah na relativni meji domf.*

Lema 4.28. *Naj bo C konveksna množica in M afina množica ki vsebuje točko iz riC. Potem velja*

$$cl(M \cap C) = M \cap clC.$$

Dokaz. Najprej ugotovimo, da za konveksne množice C_i v $\mathbb{R}^n$, $i \in \mathbb{I}$, kjer je $\mathbb{I}$ indeksna množica, takšne da je presek množice riC_i neprazen, velja

$$\text{cl}\{C_i\big| i \in \mathbb{I}\} = \cap\{\text{cl}C_i\big| i \in \mathbb{I}\},$$

Dokaz ni precej zahteven. Najprej fiksiramo x v preseku množic riC_i. Za poljuben y v preseku množic clC_i velja, da je za $0 \leq \lambda < 1$ vektor $(1 - \lambda)x + \lambda y \in \text{ri}C_i$, po 4.7. Ko $\lambda \uparrow 1$, velja da je y limita tega vektorja. Sledi,

$$\cap_i \text{cl}C_i \subset \text{cl} \cap_i \text{ri}C_i \subset \text{cl} \cap_i C_i \subset \cap_i \text{cl}C_i.$$

Ker za afino množico velja $M = \text{cl}M$, smo s tem končali dokaz leme. □

Dokaz. Ker je epi(cl$f$) = cl(epi$f$) in ker je epi$f$ konveksna po 4.8 sklepamo, da je epi(cl$f$) zaprta konveksna množica v $\mathbb{R}^{n+1}$ in da je cl$f$ navzdol polzvezna konveksna funkcija. Lastnosti cl$f$, in njeno zaprtost, v posebnem, nam bo zagotavljala zadnja trditev v izreku po 4.23, zaradi končnosti f na množici dom$f$. Za poljuben $x \in$ ri(dom$f$), vertikalna premica $M = \{(x,\mu) \big| \mu \in \mathbb{R}\}$ seka ri(epi$f$) po 4.26. Torej

$$M \cap \mathrm{cl}(\mathrm{epi}f) = \mathrm{cl}(M \cap \mathrm{epi}f) = M \cap \mathrm{epi}f,$$

po 4.28. To pa pomeni, da je, razen morda v točkah na relativni meji dom$f$ $(\mathrm{cl}f)(x) = f(x)$. Po drugi strani, recimo da $x \notin \mathrm{cl}(\mathrm{dom}f)$. Iz "lim inf" formule za cl$f$ velja

$$\mathrm{cl}(\mathrm{dom}f) \supset \mathrm{dom}(\mathrm{cl}f) \supset \mathrm{dom}f,$$

in zato $(\mathrm{cl}f)(x) = \infty = f(x)$. □

Posledica 4.29. *Če je f prava konveksna funkcija, potem se dom(cl$f$) in dom$f$ razlikujeta največ v nekaterih dodatnih točkah na relativni meji dom$f$. V posebnem, dom(cl$f$) in dom$f$ imata enako zaprtje in relativno notranjost, ter enako dimenzijo.*

Posledica 4.30. *Naj bo f prava konveksna funkcija. Če je dom$f$ afina množica, potem je f zaprta.*

Dokaz. v tem primeru sploh nimamo točk na relativni meji dom$f$, in cl$f$ povsod sovpada z f. □

Konveksna funkcija f je vedno navzdol polzvezna razen morda na točkah na relativni meji dom$f$, po 4.27 in 4.22.

Izrek 4.31. *Naj bo f prava konveksna funkcija in $x \in$ ri(dom$f$). Potem je za vsak y*

$$(clf)(y) = \lim_{\lambda \uparrow 1} f((1-\lambda)x + \lambda y)$$

Izrek velja tudi če je f neprava konveksna funkcija in $y \in$ cl(dom$f$).

Dokaz. Ker je f prava konveksna funkcija, je cl$f$ navzdol polzvezna in cl$f \leq f$ (sovpadata razen morda v točkah na relativni meji dom$f$), imamo

$$(\mathrm{cl}f)(y) \leq \liminf_{\lambda \uparrow 1} f((1-\lambda)x + \lambda y)$$

Torej samo še moramo pokazati, da velja

$$(\mathrm{cl}f)(y) \geq \limsup_{\lambda \uparrow 1} f((1-\lambda)x + \lambda y).$$

Naj bo β realno število in sicer takšno da velja $\beta \geq (\mathrm{cl}f)(y)$, in naj bo $\alpha > f(x)$. Potem je $(y,\beta) \in \mathrm{epi}(\mathrm{cl}f) = \mathrm{cl}(\mathrm{epi}f)$ in $(x,\alpha) \in \mathrm{ri}(\mathrm{epi}f)$, po 4.26. Torej je po 4.7

$$(1-\lambda)(x,\alpha) + \lambda(y,\beta) \in \mathrm{ri}(\mathrm{epi}f), \quad 0 \leq \lambda < 1.$$

Sledi, da je

$$f((1-\lambda)x+\lambda y) < (1-\lambda)\alpha + \beta, \quad 0 \leq \lambda < 1.$$

Končno, pridemo do sklepa

$$\limsup_{\lambda\uparrow 1} f((1-\lambda)x+\lambda y) \leq \limsup_{\lambda\uparrow 1}[(1-\lambda)\alpha + \lambda\beta] = \beta.$$

V slučaju da f ni prava in da je $y \in \mathrm{cl}(\mathrm{dom} f)$, velja $f((1-\lambda)x+\lambda y) = -\infty$ za $0 \leq \lambda < 1$ po 4.7 in 4.22. □

Posledica 4.32. *Naj bo f zaprta konveksna funkcija. Potem velja*

$$f(y) = \lim_{\lambda\uparrow 1} f((1-\lambda)x+\lambda y)$$

za vsak $x \in domf$ in vsak y.

Dokaz. Najprej definirajmo $\varphi(\lambda) = f((1-\lambda)x+\lambda y)$. Ker je

$$\varphi((1-u)\lambda_1 + u\lambda_2) = f((1-(1-u)\lambda_1 - u\lambda_2)x + ((1-u)\lambda_1 + u\lambda_2)y)$$

$$= ((1-(1-u)\lambda_1 - u\lambda_2)f(x) + ((1-u)\lambda_1 + u\lambda_2)f(y) =$$

$$= (1-u)f((1-\lambda_1)x + \lambda_1 y)) + uf((1-\lambda_2)x + \lambda_2 y) = (1-u)\varphi(\lambda_1) + u\varphi(\lambda_2),$$

je tako definiran φ prava konveksna funkcija na $\mathbb{R}$ za katero velja $\varphi(0) = f(x) < \infty$ in $\varphi(1) = f(y)$. Ker je $\{\lambda \big| \varphi(\lambda) \leq \alpha\}$ inverzna slika zaprte množice $\{z \big| f(z) \leq \alpha\}$ pod zvezno transformacijo $\lambda \to (1-\lambda)x + \lambda y = z$, po 4.14 sledi, da je φ navzdol polzvezna. Efektivna domena φ je določeni interval. Če notranje točke tega intervala ležijo med 0 in 1, potem po 4.31 lahko sklepamo da je limita $\varphi(\lambda)$ ko $\lambda \uparrow 1$ enaka $(\mathrm{cl}\varphi)(1) = \varphi(1)$. V vseh drugeh primerih sta limita in $\varphi(1)$ oba trivialno $+\infty$. □

Za potrebe dokazovanja v naslednjih poglavjih, prav tako kot zaradi bolj globokega in jasnega razumevanja konveksnosti, si bomo ogledali nekaj osnovnih dejstev iz teorije dualnosti. Bomo spoznali kako lahko vsaki funkciji priredimo konjugirano funkcijo.

Definicija 4.33. Hiperravnina H *separira* množici C_1 in C_2 v $\mathbb{R}^n$ če je množica C_1 vsebovana v enem zaprtem polprostoru asociranem z H in je množica C_2 vsebovana v drugem zaprtem polprostoru asociranem z H. Če množici C_1 in C_2 nista obe vsebovani v samem H, rečemo da H separira *pravilno* množici C_1 in C_2. Če $\exists\varepsilon > 0$ takšen da je množica $C_1 + \varepsilon B$ vsebovana v enem odprtem polprostoru asociranem z H in je množica $C_2 + \varepsilon B$ vsebovana v drugem odprtem polprostoru asociranem z H, potem rečemo da H separira *močno* množici C_1 in C_2.

Lema 4.34. *Zaprta konveksna množica C je presek zaprtih polprostorov ki jo vsebujejo.*

Dokaz. Le skica. Najprej pokažemo, da je

$$\inf\{|x_1 - x_2| \big| x_1 \in C_1, x_2 \in C_2\} > 0$$

potreben in zadoščen pogoj za obstoj hiperravnine ki separira močno množici C_1 in C_2 v $\mathbb{R}^n$. Da bi to naredili, moramo dokazati da je potreben in zadoščen pogoj za obstoj hiperravnine ki separira močno množici C_1 in C_2 v $\mathbb{R}^n$ da je presek množic riC_1 in riC_2 prazen. Več podrobnosti o tem si bralec lahko najde na str. 97 in str. 98 v [15]. Recimo, da je $\emptyset \neq C \neq \mathbb{R}^n$. Za poljuben $a \notin C$, množici $C_1 = \{a\}$ in $C_2 = C$ zadoščata pogoju $\inf\{|x_1 - x_2| \big| x_1 \in C_1, x_2 \in C_2\} > 0$, in torej obstaja hiperravnina ki separira močno množici C_1 in C_2 v $\mathbb{R}^n$. En od zaprtih polprostorov ki sta asocirani z to hiperravnino vsebuje C vendar ne a. Se pravi, presek zaprtih polprostorov ki vsebujejo C vsebuje le točke vsebovane v C, tj. C je presek zaprtih polprostorov ki jo vsebujejo. □

Spomnimo se, da se hiperravnine v $\mathbb{R}^{n+1}$ lahko predstavljajo s pomočjo linearnih funkcij na $\mathbb{R}^{n+1}$, te se pa lahko predstavljajo v obliki

$$(x, \mu) \to \langle x, b\rangle + \mu\beta_0, \quad b \in \mathbb{R}^n, \quad \beta_0 \in \mathbb{R}.$$

Ker neničelne linearne funkcije, ki so med seboj skalarni večkratniki, določajo iste hiperravnine, nas zanimata samo primera ko je $\beta_0 = 0$ in $\beta_0 = -1$. V prvem primeru so hiperravnine so oblike

$$\{(x, \mu) \big| \langle x, b\rangle = \beta\}, \quad 0 \neq b \in \mathbb{R}^n, \quad \beta \in \mathbb{R}.$$

Takšne hiperravnine bomo imenovali *vertikalne.* V drugem primeru so hiperravnine oblike

$$\{(x, \mu) \big| \langle x, b\rangle - \mu = \beta\}, \quad b \in \mathbb{R}^n, \quad \beta \in \mathbb{R}.$$

Te pa so grafi afinih funkcij $h(x) = \langle x, b\rangle - \beta$ na $\mathbb{R}^n$. Sklepamo, da so vsi zaprti polprostori v $\mathbb{R}^{n+1}$ oblike

1. $\{(x, \mu) \big| \langle x, b\rangle \leq \beta\} = \{(x, \mu) \big| h(x) \leq 0\}, b \neq 0,$
2. $\{(x, \mu) \big| \mu \geq \langle x, b\rangle - \beta\} = \text{epi}h,$
3. $\{(x, \mu) \big| \mu \leq \langle x, b\rangle = \beta\}$

Ta tri tipa zaprtih polprostorov bomo imenovali *vertikalni, zgornji* in *spodnji*, respektivno.

Izrek 4.35. *Zaprta konveksna funkcija f je supremum po točkah kolekcije afinih funkcij h takšnih da velja $h \leq f$.*

Dokaz. Privzemimo, da je f prava funkcija, sicer imamo trivialen slučaj. Ker je epi$f$ zaprta konveksna množica, je po 4.34 presek zaprtih polprostorov v $\mathbb{R}^{n+1}$, ki jo vsebujejo. Spodnji polprostori ne morejo vsebovati epi$f$, torej v preseku imamo samo vertikalne in zgornje zaprte polprostore. Ne morejo vsi polprostori v preseku biti vertikalni, saj bi to pomenilo da je epi$f$ unija vertikalnih premic v $\mathbb{R}^{n+1}$, kar bi bilo v nasprotju s pogojem da je funkcija prava. Zgornje zaprti polprostori ki vsebujejo epi$f$ so epigrafi afinih funkcij h takšnih da $h \leq f$. Njun presek je epigraf supremuma po točkah takšnih funkcij h. Za dokaz izreka bo dovolj dokazati, da je presek vertikalnih in zgornje zaprtih polprostorov ki vsebujejo epi$f$ enak preseku samo zgornje zaprtih polprostorov ki vsebujejo epi$f$. Denimo, da je

$$V = \{(x, \mu) \big| 0 \geq \langle x, b_1 \rangle - \beta_i = h_1(x)\}$$

vertikalni polprostor ki vsebuje epi$f$, in da je (x_0, μ_0) točka, ki ni v V. Zadošča pokazati, da obstaja afina funkcija h takšna da je $h \leq f$ in $\mu_0 < h(x_0)$. Že vemo da obstaja vsaj ena afina funkcija h_2 takšna da je epi$h_2 \supset$ epi$f$, tj. $h_2 \leq f$. Za vsak $x \in$ dom$f$ imamo $h_1(x) \leq 0$ in $h_2(x) \leq f(x)$, in sicer

$$\lambda h_1(x) + h_2(x) \leq f(x), \quad \forall \lambda > 0.$$

Ko $x \notin$ dom$f$, zgornja enakost velja trivialno, ker je v tem primeru $f(x) = +\infty$. Torej, če fiksiramo $\lambda \geq 0$ in definiramo h kot

$$h(x) = \lambda h_1(x) + h_2(x) = \langle x, \lambda b_1 + b_2 \rangle - (\lambda \beta_1 + \beta_2)$$

bomo imeli $h \leq f$. Ker je $h_1(x_0) > 0$, dovolj veliki λ nam bo zagotavljal $h(x_0) > \mu_0$, kot smo hoteli. □

Naj bo f poljubna konveksna funkcija na $\mathbb{R}^n$. Po 4.35, f si lahko opišemo tako da opišemo množico F^* ki sestoji iz vseh parov (x^*, μ^*) v $\mathbb{R}^{n+1}$ takšnih da je afina funkcija $h(x) = \langle x, x^* \rangle - \mu^*$ majorizirana z f. Potem za vsak x imamo $h(x) \leq f(x)$ če in samo če je

$$\mu^* \geq \sup\{\langle x, x^* \rangle - f(x) \big| x \in \mathbb{R}^n\}.$$

Torej, F^* je epigraf funkcije f^* na $\mathbb{R}^n$, definirane kot

$$f^*(x^*) = \sup_x \{\langle x, x^* \rangle - f(x)\} = -\inf_x \{f(x) - \langle x, x^* \rangle\}.$$

Torej velja

$$f(x) \geq \langle x, x^* \rangle - f^*(x^*) \quad \forall x$$

in

$$\forall \varepsilon > 0 : \exists x : f(x) \ngeq \langle x, x^* \rangle - (f^*(x^* - \varepsilon)).$$

Definicija 4.36. Zgoraj definirano funkcijo f^* imenujemo *konjugirana* funkcija funkcije f.

Se iskaže, da je f^* supremum po točkah zaporedja afinih funkcij $g(x^*) = \langle x, x^* \rangle - \mu$, takšnih da (x, μ) pripada množici $F = \text{epi}f$, in je po 3.37 kot takšna funkcija konveksna, in v bistvu zaprta konveksna funkcija. Naj bo $F^* = \text{epi}f^*$. Ker je f supremum po točkah afinih funkcij $h(x) = \langle x, x^* \rangle - \mu^*$ takšnih da je $(x^*, \mu^*) \in F^*$, imamo

$$f(x) = \sup_{x^*}\{\langle x, x^* \rangle - f^*(x^*)\} = -\inf_{x^*}\{f^*(x^*) - \langle x, x^* \rangle\},$$

kar pomeni, da je f^{**}, tj. funkcija ki je konjugirana funkciji f^* enaka $\text{cl}f$.

Primer 4.37. Konstantni funkciji $-\infty$ in $+\infty$ sta konjugirani med sabo. Ker sta te dve edini nepravilni zaprti konveksni funkciji, morajo vsi drugi konjugirani pari biti pravilni.

Zgoraj omenjene ugotovitve lahko podamo v obliki izreka.

Izrek 4.38. *Naj bo f konveksna funkcija. Konjugirana funkcija funkciji f je potem zaprta konveksna funkcija, prava če in samo če je f prava. Lahko povemo še več, in sicer da velja $(clf) = f^*$ in $f^{**} = clf$.*

Primer 4.39. Naj bo funkcija f zaprta prava konveksna funkcija $e^x, x \in \mathbb{R}$. Direkten izračun pokaže, da je funkciji f konjugirana funkcija

$$f^*(x^*) = \begin{cases} x^* log x^* - x^* & \text{če } x^* > 0, \\ 0 & \text{če } x^* = 0, \\ +\infty & \text{če } x^* < 0. \end{cases}$$

Primer 4.40. Naj bosta p in q takšni da velja $\frac{1}{p} + \frac{1}{q} = 1$. Naj bo f funkcija podana z $f(x) = \frac{1}{p}|x|^p$ za $1 < p < +\infty$. Potem je $f^*(x^*) = \frac{1}{q}|x^*|^q$.

Spomnimo se, da *podporno funkcijo* $\delta^*(\cdot|C)$ množice C definiramo kot

$$\delta^*(x^*|C) = \sup\{\langle x, x^* \rangle \big| x \in C\}.$$

Podporna funkcija množice C opisuje vse zaprte polprostore ki vsebujejo C. Res,

$$C \subset \{x \big| \langle x, x^* \rangle \leq \beta\} \quad \text{če in samo če } \beta \geq \delta^*(x^*|C).$$

Izrek 4.41. *Naj bo C konveksna množica. Potem je $x \in clC$ če in samo če je $\langle x, x^* \rangle \leq \delta^*(x^*|C)$ za vsak vektor x^*. Po drugi strani, $x \in riC$ če in samo če velja isti pogoj, vendar z strogo neenakostjo za vsak x^* takšen da je $-\delta^*(-x^*|C) \neq \delta^*(x^*|C)$. Dalje, $x \in intC$ če in samo če je $\langle x, x^* \rangle < \delta^*(x^*|C)$ za vsak $x^* \neq 0$. Končno, če privzamemo $C \neq \emptyset$, velja $x \in affC$ če in samo če je $\langle x, x^* \rangle = \delta^*(x^*|C)$ za vsak x^* takšen da je $-\delta^*(-x^*|C) = \delta^*(x^*|C)$.*

Dokaz. Le del. Najprej pokažimo, da je poljubno podmnožico S v $\mathbb{R}^n$ cl(conv$S$) presek vseh zaprtih polprostorov ki vsebujejo S. To pa je posledica dejstva da je zaprta konveksna množica C presek vseh zaprtih polprostorov ki jo vsebujejo, in smo te že vgotovili. Potem pokažimo, da za poljubno konveksno množico C velja da je $x \in C$ točka na relativni meji C če in samo če obstaja linearna funkcija h ki ni konstantna na C, takšna da h doseže maksimum na C v x. Iz teh dveh dejstev, respektivno, nam bosta sledila karakterizaciji cl$C$ in ri$C$. Slučaj, ko je ri$C$ prav zaprav int$C$ je slučaj ko C ni vsebovana niti v eni hiperravnini, tj. ko je $-\delta^*(-x^*|C) \neq \delta^*(x^*|C)$ za poljuben $x^* \neq 0$. To nam zagotavlja karakterizacijo int$C$. Karakterizacija aff$C$ nam pove v bistvu pove da je najmanjša afina množica ki vsebuje C enaka kot presek vseh hiperravnin ki vsebujejo C. □

Posledica 4.42. *Naj bosta C_1 in C_2 konveksni množici v $\mathbb{R}^n$. Potem velja cl$C_1 \subset$ clC_2 če in samo če je $\delta^*(\cdot|C_1) \leq \delta^*(\cdot|C_2)$.*

Dokaz. Po prejšnji trditvi je $x \in \mathrm{cl}C_1$ če in samo če je $\langle x, x^*\rangle \leq \delta^*(x^*|C_1)$ za vsak vektor x^*. Podobno, $x \in \mathrm{cl}C_2$ če in samo če je $\langle x, x^*\rangle \leq \delta^*(x^*|C_2)$ za vsak vektor x^*. Če je $\mathrm{cl}C_1 \subsetneqq \mathrm{cl}C_2$, potem $\exists x \in \mathrm{cl}C_2, x \notin \mathrm{cl}C_1$ in sicer velja $\delta^*(\cdot|C_2) \geq \langle x, x^*\rangle$ in $\delta^*(\cdot|C_1) < \langle x, x^*\rangle$. Torej velja $\delta^*(\cdot|C_1) \leq \delta^*(\cdot|C_2)$. Po drugi strani, če je $\delta^*(\cdot|C_1) \leq \delta^*(\cdot|C_2)$, potem imamo da je $\delta^*(\cdot|C_2) \geq \langle x, x^*\rangle$ za vsak $x \in \mathrm{cl}C_1 inx^*$, se pravi $\mathrm{cl}C_1 \subset \mathrm{cl}C_2$. □

Iz tega sledi, da si lahko zaprto konveksno množico C predstavljamo kot množico rešitev sistema neenakosti danih z njeno podporno funkcijo

$$C = \{x \big| \langle x, x^*\rangle \leq \delta^*(x^*|C), \forall x^*\}.$$

Torej podporna funkcija popolnoma določa množico. Odtod sledi, da obstaja pomembna injektivna korespondenca med zaprtimi konveksnimi množicami v $\mathbb{R}^n$ in nekateri funkcijami na $\mathbb{R}^n$. Ta korespondenca ima nekatere presenetljive lasnosti. Npr. podporna funkcija vsote dveh nepraznih konveksnih množic C_1 in C_2 je dana z

$$\delta^*(x^*|C_1 + C_2) = \sup\{\langle x_1 + x_2, x^*\rangle \big| x_1 \in C_1, x_2 \in C_2\}$$

$$= \sup\{\langle x_1, x^*\rangle \big| x_1 \in C_1\} + \sup\{\langle x_2, x^*\rangle \big| x_2 \in C_2\}$$

$$= \delta^*(x^*|C_1) + \delta^*(x^*|C_2)$$

. Korespondenca med podpornimi funkcijami se lahko posmatra kod poseben primer konjugacije, če imamo v mislih trivialno injektivno korespondenco med konveksnimi množicami C in indikatorskimi funkcijami $\delta(\cdot|C)$. Po definiciji je funkcija konjugirana $\delta(\cdot|C)$ podana z

$$\sup_{x \in \mathbb{R}^n} \{\langle x, x^*\rangle - \delta(x|C)\} = \sup_{x \in C} \langle x, x^*\rangle = \delta^*(x^*|C).$$

Potem $\delta^*(x^*|C)$ zadošča

$$(\delta^*(\cdot|C))^* = cl\delta(\cdot|C) = \delta(\cdot|clC),$$

po 4.38.

Izrek 4.43. *Indikatorska funkcija in podporna funkcija zaprte konveksne množice sta konjugirani med sabo. Funkcije katere so podporne funkcije nepraznih konveksnih množic so zaprte pravilne konveksne funkcije ki so pozitivno homogene.*

Dokaz. Po zgornjem premisleku in 4.38. Zadošča pokazati, da zaprta prava konveksna funkcija f ne more imeti vrednosti razen 0 in $+\infty$ če in samo če je njen konjugat pozitivno homogen. Prva lastnost funkcije f je ekvivalentna pogoju $f(x) = \lambda f(x)$ za vsak x in $\lambda > 0$. Druga lastnost je ekvivalentna pogoju $f^*(x^*) = \lambda f^*(\lambda^{-1}x^*) = (f^*\lambda)(x^*)$ za vsak x^* in $\lambda > 0$. Ampak

$$(\lambda f)^*(x^*) = \sup_x \{\langle x, x^*\rangle - \lambda f(x)\}$$
$$= \sup_x \{\lambda(\langle x, \lambda^{-1}x^*\rangle - f(x))\} = \lambda f^*(\lambda^{-1}x^*).$$

Torej je za vsak $\lambda > 0$ $f = \lambda f$ če in samo če je $f^* = f^*\lambda$ za vsak $\lambda > 0$, kadar je f zaprta konveksna funkcija. □

V posebnem, 4.43 nam pove da je $\delta^*(x^*|C)$ navzdol polzvezna funkcija x^* in

$$\delta^*(x_1^* + x_2^*|C) \leq \delta^*(x_1^*|C) + \delta^*(x_2^*|C), \quad \forall x_1^*, x_2^*.$$

Posledica 4.44. *Naj bo f poljubna pozitivno homogena funkcija različna od $+\infty$. Potem je clf podporna funkcija nekatere zaprte konveksne množice C, namreč*

$$C = \{x^* \big| \forall x, \langle x, x^*\rangle \leq f(x)\}.$$

Dokaz. Obstajata dve možnosti: $\mathrm{cl}f$ je zaprta prava pozitivno homogena konveksna funkcija, ali je $\mathrm{cl}f$ konstantna funkcija $-\infty$, kot podporna funkcija $\emptyset$. Torej za nekatero zaprto konveksno množico C velja $\mathrm{cl}f{=}\delta^*(\cdot|C)$. Po definiciji sledi, da je $f^* = (\mathrm{cl}f)^* = \delta(\cdot|C)$ in $C = \{x^* \big| f^*(x^*) \leq 0\}$. Ampak $f^*(x^*) \leq 0$ če in samo če je $\langle x, x^*\rangle - f(x) \leq 0$ za vsak x. □

Posledica 4.45. *Podporne funkcije nepraznih omejenih konveksnih množic so končne pozitivno homogene konveksne funkcije.*

Dokaz. Po 4.30, končna konveksna funkcija je nujno zaprta. Samo moramo opaziti da je konveksna množica C omejena če in samo če je $\delta^*(x^*|C) < +\infty$ za vsak x^*. Res, množica C v $\mathbb{R}^n$ je omejena če in samo če je vsebovana v nekateri kocki, in to velja natančno tedaj ko je vsaka linearna funkcija omejena navzgor na C. □

5 Zveznost konveksnih funkcij

Smo si ogledali kako lahko s pomočjo operacije zaprtja od konveksne funkcije naredimo navzdol polzvezno funkcijo. Zdaj pa bomo študirali situacije ko je funkcija f podana kod navzgor polzvezna. V tem primeru je $\mathrm{cl}f$ zvezna, in pa sama funkcija f, vse dokler sovpada z $\mathrm{cl}f$. Si bomo ogledali nekaj lastnosti enakomerne zveznosti in ekvizveznosti konveksnih funkcij.

Definicija 5.1. Funkcija f na $\mathbb{R}^n$ je *zvezna relativno podmnožici* S v $\mathbb{R}^n$ če je zožitev f na S zvezna funkcija, tj. če za $x \in S$, velja $\lim\limits_{y \in S, y \longrightarrow x} f(y) = f(x)$.

Izrek 5.2. *Konveksna funkcija f na $\mathbb{R}^n$ je zvezna relativno katerikoli relatvno odprti konveksni množici C v njeni efektivni domeni, in v posebnem primeru relativno $ri(domf)$.*

Dokaz. Naj bo g funkcija ki sovpada s f na C vendar je $+\infty$ povsod drugje. Vemo, da ima takšna funkcija C kot svojo efektivno domeno. Po potrebi lahko zamenjamo f z g, in obravnavamo samo primer ko je $C = \mathrm{dom}f$. Lahko tudi privzamemo, brez izgube splošnosti, da je C n-dimenzionalna množica in zato odprta, torej ne samo relativno odprta. Če je f neprava, po 4.22 zavzame vrednost $-\infty$ na C, in zveznost sledi trivialno. Recimo, da je f prava, oz. končna na C. Imamo $(\mathrm{cl}f)(x) = f(x)$ za vsak $x \in C$ (4.27), torej je f navzdol polzvezna na C. Da bi pokazali zveznost, zadošča dokazati da so vse množice $\{x \big| f(x) \geq \alpha\}$ zaprte, in to bo pomenilo da je f navzgor polzvezna povsod po 4.14, ki nam pove da zaprtost množic $\{x \big| f(x) \leq \alpha\}$ pomeni navzdol polzveznost, torej iz zaprtosti množic $\{x \big| f(x) \geq \alpha\}$ analogno temu sledi navzgor polzveznost. Ker je $C = \mathrm{dom}f$ odprta, po 4.26 velja $\mathrm{int}(\mathrm{epi}f) = \{(x, \mu) \big| \mu > f(x)\}$. Torej za vsak $\alpha \in \mathbb{R}$, $\{x \big| f(x) < \alpha\}$ je projekcija odprteg konveksneg preseka $\mathrm{int}(\mathrm{epi}f)$ in polprostora $\{(x, \mu) \big| \mu < \alpha\}$ v $\mathbb{R}^{n+1}$ na $\mathbb{R}^n$, kar pomeni da je $\{x \big| f(x) < \alpha\}$ odprta in njen komplement $\{x \big| f(x) \geq \alpha\}$ je zaprt. $\square$

Posledica 5.3. *Vsaka konveksna funkcija ki je končna na celem $\mathbb{R}^n$ je zvezna.*

Definicija 5.4. Funkcija f ki preslika množico $S \subset \mathbb{R}^n$ v $\mathbb{R}$, se imenuje *Lipschitzova* relativno množici S če obstaja $\alpha \geq 0$ tako da

$$|f(y) - f(x)| \leq \alpha |y - x|, \quad \forall y \in S, \quad \forall x \in S.$$

Izrek 5.5. *Naj bo f prava konveksna funkcija. Naj bo S poljubna zaprta omejena podmnožica $ri(domf)$. Potem je f Lipschitzova relativno S.*

Dokaz. Denimo, brez škode za splošnost, da je dom$f$ n-dimenzionalna v $\mathbb{R}^n$, tako da je S v notranjosti dom$f$. Naj bo B zaprta evklidska enotna krogla. Za vsak $\varepsilon > 0$, $S+\varepsilon B$ je zaprta omejena množica (slika kompaktne množice $S \times B$ pod zvezno transformacijo $(x, u) \longrightarrow x + \varepsilon u$). Ker ima zaporedje množic

$$(S + \varepsilon B) \cap (\mathbb{R}^n \backslash \mathrm{int}(\mathrm{dom}f)), \quad \varepsilon > 0$$

prazen presek, ena od množic v tem zaporedju je prazna. Sledi, da za nek $\varepsilon > 0$, velja

$$S + \varepsilon B \subset \mathrm{int}(\mathrm{dom}f).$$

Po 5.2, f je zvezna na $S + \varepsilon B$. Ker je $S + \varepsilon B$ zaprta omejena množica, sledi da je f omejena na $S + \varepsilon B$. Naj bosta α_1 in α_2 njena spodnja in zgornja meja, respektivno. Naj bosta x in y poljubni dve različni točki v S in naj bo

$$z = y + \varepsilon \frac{y - x}{|y - x|}$$

Potem je $z \in S + \varepsilon B$ in

$$y = (1 - \lambda)x + \lambda z, \qquad \lambda = (y - x)/(\varepsilon + |y - x|).$$

Iz konveksnosti funkcije f imamo

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

in odtod

$$f(y) - f(x) \leq \lambda(\alpha_2 - \alpha_1) \leq \alpha|y - x|, \quad \alpha = \frac{\alpha_2 - \alpha_1}{\varepsilon}.$$

To neenakost velja za vsak $x, y \in S$, tj. f je Lipschitzova relativno S. □

Končna konveksna funkcija f na $\mathbb{R}^n$ je enakomerno zvezna, še več - Lipschitzova relativno vsaki omejeni množici po prejšnjem izreku, ampak f ni nujno enakomerno zvezna ali Lipschitzova relativno celotnem $\mathbb{R}^n$. Funkcija $f = x^2$ je enakomerno zvezna in Lipschitzova relativno vsaki omejeni množici v $\mathbb{R}$, vendar ni enakomerno zvezna ni Lipschitzova relativno celotnem $\mathbb{R}$.

Definicija 5.6. Naj bo $\{f_i \big| i \in \mathbb{I}\}$ kolekcija funkcij z realnimi vrednostimi na podmnožici S v $\mathbb{R}^n$. Pravimo, da je $\{f_i \big| i \in \mathbb{I}\}$ *enakomerno Lipschitzova* relativno S če $\exists \alpha \geq 0$ da

$$|f_i(y) - f_i(x)| \leq \alpha|y - x|, \quad \forall y \in S, \quad \forall x \in S, \quad \forall i \in \mathbb{I}.$$

Recimo, da je ta pogoj izpolnjen. Tedaj velja, da je kolekcija *enakomerno zvezna* relativno S, oz. da $\forall \varepsilon > 0 \quad \exists \delta > 0$ tako da velja

$$|f_i(y) - f_i(x)| \leq \varepsilon, \quad \forall y \in S, \quad \forall x \in S, \quad \forall i \in \mathbb{I}, \quad |y - x| < \delta.$$

Definicija 5.7. Pravimo, da je $\{f_i \big| i \in \mathbb{I}\}$ *omejena po točkak* na S če je množica realnih števil $\{f_i(x), i \in \mathbb{I}\}$ omejena za vsak $x \in S$. Če obstajata realni števili α_1 in α_2 da

$$\alpha_1 \leq f_i(x) \leq \alpha_2, \quad \forall x \in S, \quad \forall i \in \mathbb{I},$$

rečemo da je $\{f_i \big| i \in \mathbb{I}\}$ *enakomerno omejena* na S.

Izrek 5.8. *Za poljubno relativno odprto konveksno množico C, naj bo $\{f_i \big| i \in \mathbb{I}\}$ poljubna kolekcija konveksnih funkcij ki so končne in omejene po točkah na C. Naj bo S poljubna zaprta omejena podmnožica množice C. Potem je $\{f_i \big| i \in \mathbb{I}\}$ enakomerno zvezna na S in ekvi-Lipschitzova relativno S. Zaključek bo še vedno veljaven če namesto omejenosti po točkah privzamemo dve predpostavke:*

1. *Obstaja podmnožica C' množice C takšna da je conv(clC') $\supset C$ in za vsak $x \in C'$ velja da je množica* $\sup\{f_i(x) \big| i \in \mathbb{I}\}$ *končna.*
2. *Obstaja vsaj en $x \in C$ takšen da je* $\inf\{f_i(x) \big| i \in \mathbb{I}\}$ *končna množica.*

Dokaz. Brez škode za splošnost, naj bo C odprta. Naj veljata 1. in 2., bomo dokazali da je $\{f_i \big| i \in \mathbb{I}\}$ enakomerno omejena na vsaki zaprti omejeni podmnožici C. Ekvi-Lipschitzov pogoj bo nato sledil po dokazu 5.5, ker je Lipschitzova konstanta α v tem dokazu odvisna samo od α_1 in α_2. Naj bo

$$f(x) = \sup\{f_i(x) \big| i \in \mathbb{I}\}.$$

To je po 3.37 konveksna funkcija. Ker je $\mathrm{cl}C' \subset \mathrm{cl}(\mathrm{dom}f)$ in zaradi tega še $\mathrm{conv}(\mathrm{cl}C') \subset \mathrm{cl}(\mathrm{dom}f)$ in $C \subset \mathrm{cl}(\mathrm{dom}f)$, imamo po 1.

$$\mathrm{dom}f \supset \mathrm{int}(\mathrm{cl}(\mathrm{dom}f)) \supset \mathrm{int}C = C,$$

ker je dom$f$ konveksna množica, po 4.9. Torej, po 5.2 je f zvezna relativno množici C. V posebnem, f je omejena navzgor na vsaki zaprti omejeni podmnožici C, oz. $\{f_i \big| i \in \mathbb{I}\}$ je enakomerno omejena navzgor na vsaki zaprti omejeni podmnožici C. Za dokaz da je $\{f_i \big| i \in \mathbb{I}\}$ tudi enakomerno omejena navzdol na vsaki zaprti omejeni podmnožici C, bo dovolj zgraditi funkcijo g, zvezno in z realnimi vrednostimi, da

$$f_i(x) \geq g(x), \quad \forall x \in C, \quad \forall i \in \mathbb{I}.$$

Uporabljajoči 2., izberimo poljuben $\bar{x} \in C$ da

$$-\infty < \beta_1 = \inf\{f_i(\bar{x}) \big| i \in \mathbb{I}\}.$$

Izberimo $\varepsilon > 0$ da $\bar{x} + \varepsilon B \subset C$, kjer je B evklidska enotna krogla, in naj bo β_2 pozitivna zgornja meja f na $\bar{x} + \varepsilon B$. Za vsak $x \in C, x \neq \bar{x}$, imamo $\bar{x} = (1-\lambda)z + \lambda x$ za

$$z = \bar{x} + \frac{\varepsilon(\bar{x} - x)}{|\bar{x} - x|}$$

$$\lambda = \frac{\varepsilon}{(\varepsilon + |\bar{x} - x|)}.$$

Ker je $0 < \lambda < 1$ in $|z - \bar{x}| = \varepsilon$, imamo (za poljuben $i \in \mathbb{I}$)

$$\beta_1 \leq f_i(\bar{x}) \leq (1-\lambda)f_i(z) + \lambda f_i(x) \leq (1-\lambda)\beta_2 + \lambda f_i(x) \leq \beta_2 + \lambda f_i(x)$$

in posledično

$$f_i(x) \geq \frac{(\beta_1 - \beta_2)}{\lambda} = \frac{(\varepsilon + |\bar{x} - x|)(\beta_1 - \beta_2)}{\varepsilon}.$$

Količina na desni strani je zvezno odvisna od x. Neenakost velja za vsak $x \in C$ in vsak $i \in \mathbb{I}$. □

Izrek 5.9. *Naj bo $f_1, f_2, \ldots,$ zaporedje končnih konveksnih funkcij na relativno odprti konveksni množici C. Recimo, da zaporedje konvergira po točkah na gosti podmnožici C, oz. da obstaja $C' \subset C$ tako da je cl$C' \supset C$ in, za vsak $x \in C'$, limita $f_1(x), f_2(x), \ldots,$ obstaja in je končna. Limita potem obstaja za vsak $x \in C$ in funkcija f za katero je*

$$f(x) = \lim_{i \to \infty} f_i(x)$$

je končna in konveksna na C. Zaporedje $f_1, f_2, \ldots,$ konvergira enakomerno proti f na vsaki kompaktni podmnožici množice C.

Dokaz. Naj bo C odprta, brez izgube splošnosti. Kolekcija funkcij $\{f_i \big| x = 1, 2, \ldots\}$ je omejena po točkah na C' in je po 5.8 ekvi-Lipschitzova na vsaki zaprti omejeni podmnožici C. Naj bo S poljubna kompaktna podmnožica C, in S' kompaktna podmnožica C da int$S' \supset S$ (eksistenco takšne množice S' nam zagotavlja argument s začetka dokaza 5.5). Tedaj $\exists \alpha > 0$

$$|f_i(y) - f_i(x)| \leq \alpha|y - x|, \quad \forall y \in S', \quad \forall x \in S', \quad \forall i.$$

Ker je C' gosta množica v S, za vsak $\varepsilon > 0$, obstaja končna podmnožica C'_0 v $C' \cap S'$ takšna da je vsaka točka v S, na razdalji manjši od $\frac{\varepsilon}{3\alpha}$ od vsaj ene točke v C'_0. Ker je C'_0 končna in funkcije f_i konvergirajo po točkah na C'_0, obstaja število i_0 takšno da

$$|f_i(z) - f_i(z)| \leq \frac{\varepsilon}{3} \quad \forall i \geq i_0, \quad \forall j \geq i_0, \quad \forall z \in C'_0.$$

Za dan $x \in S$, naj bo z ena od točk iz C'_0 takšna da $|z - x| \leq \frac{\varepsilon}{3\alpha}$. Za vsaka $i, j \geq i_0$

$$|f_i(x) - f_j(x)| \leq |f_i(x) - f_i(z)| + |f_i(z) - f_j(z)| + |f_j(z) - f_j(x)|$$

$$\leq \alpha|x - z| + \frac{\varepsilon}{3} + \alpha|z - x| \leq \varepsilon.$$

Torej, za dani $\varepsilon > 0, \exists i_0$, da

$$|f_i(x) - f_j(x)| \leq \varepsilon, \quad \forall i \geq i_0 \quad \forall j \geq j_0, \quad \forall x \in S$$

Sledi, da za vsak $x \in S$, realna števila $f_1(x), f_2(x), \dots,$ formirajo Cauchyjevo zaporedje, tako da limita $f(x)$ obstaja in je končna. Poleg tega, $\forall \varepsilon > 0 \quad \exists i_0 \in \mathbb{N}$:

$$|f_i(x) - f(x)| = \lim_{j \to \infty} |f_i(x) - f_j(x)| \leq \varepsilon, \quad \forall x \in S, \quad \forall i \geq i_0.$$

Sklepamo, da funkcije f_i konvergirajo enakomerno k f na S. Ker je S poljubna zaprta omejena podmnožica množice C, sklepamo da f obstaja povsod na C. Neenakost

$$f_i((1-\lambda)x + \lambda y) \leq (1-\lambda)f_i(x) + \lambda f_i(y)$$

se ohrani za vsaka $x, y \in C$ in $\lambda \in [0,1]$ ko $i \longrightarrow \infty$, saj so vse funkcije f_i konveksne, in ker konvergirajo enakomerno proti f imamo

$$\lim_{i \to \infty} f_i((1-\lambda)x + \lambda y) \leq (1-\lambda) \lim_{i \to \infty} f_i(x) + \lambda \lim_{i \to \infty} f_i(y)$$

tj.

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

oz. je f konveksna na množici C. □

Posledica 5.10. *Naj bo f končna konveksna funkcija, in $f_1, f_2, \dots,$ zaporedje končnih konveksnih funkcij na relativno odprti konveksni množici C, tako da velja*

$$\limsup_{i \to \infty} f_i(x) \leq f(x) \quad \forall x \in C.$$

Potem, za vsako zaprto omejeno podmnožico S v C in $\varepsilon > 0$, obstaja i_0 da velja

$$f_i(x) \leq f(x) + \varepsilon, \quad \forall i \geq i_0, \quad \forall x \in S.$$

Dokaz. Naj bo $g_i(x) = \max\{f_i(x), f(x)\}$. Zaporedje konveksnih (3.37) končnih funkcij g_i konvergira po točkah k f na C, in torej konvergira enakomerno k f na S. Iz enakomerne konvergence pa sledi da velja 5.10. □

Izrek 5.11. *Naj bo $f_1, f_2, \dots,$ zaporedje končnih konveksnih funkcij na relativno odprti konveksni množici C. Naj bo zaporedje realnih števil $f_1(x), f_2(x), \dots,$ omejeno za vsak $x \in C'$, kjer je C' gosta podmnožica v C. Potem lahko izberemo podzaporedje $f_1, f_2, \dots,$ ki konvergira enakomerno na zaprtih omejenih podmnožicah C k neki končni konveksni funkciji f.*

Oglejmo si eno lemo preden začnemo dokazovati zgornji izrek.

Lema 5.12. *Če je C' poljubna podmnožica množice $\mathbb{R}^n$, obstaja števna podmnožica C'' množice C' da velja $clC'' \supset C'$.*

Dokaz. Le skica. Naj bo Q_1 kolekcija vseh zaprtih (evklidskih) krogel v $\mathbb{R}^n$ katerih centri imajo racionalne koordinate in katerih polmeri so racionalni. Naj bo Q podkolekcija ki sestoji iz vseh krogel v Q_1 ki imajo neprazen presek z C'. Do C'' pridemo z izbiranjem točk iz $D \cap C'$ za vsak $D \in Q$. $\square$

Uporabimo zdaj dejstvo iz leme za dokazovanje izreka.

Dokaz. Torej najprej uporabimo zgornjo lemo na podmnožico C' množice C, takšno da je $\text{cl}C' \supset C$ in $\{f_i(x_i)\big|i = 1, 2, \ldots\}$ omejena za vsak $x \in C'$. Po tej poti bomo prišli do množice C'' ki ima lastnosti iz prejšnje leme. Vse kar moramo pokazati je, da obstaja podzaporedje zaporedja $f_1, f_2, \ldots$ ki konvergira po točkah na C''. Uredimo elemente C'' v zaporedje $x_1, x_2, \ldots$, potem bo zaporedje realnih števil $\{f_i(x_i)\big|i = 1, 2, \ldots\}$ omejeno in bo imelo vsaj eno konvergentno podzaporedje. Torej lahko najdemo realno število α_1 in neskončno podmnožico I_1 množice $\{1, 2, \ldots\}$, tako da vrednosti funkcij f_i v podzaporedju ki korespondira I_1 konvergirajo k α_1 v x_1. Ker je $\{f_i(x_2)\big|i \in I_1\}$ omejena, lahko najdemo realno število α_2 in neskončno podmnožico I_2 množice I_1, ki ne vsebuje prvo (oz. najmanjše) število v I_1 tako da vrednosti funkcij f_i v podzaporedju ki korespondira I_2 konvergirajo k α_2 v x_2 in tudi k α_1 v x_1. Nadaljujmo z tistim premislekom in za vsak x_j dobimo pripadajoče I_j in λ_j. Naj bo I neskončna množica ki sestoji iz prvega števila v I_1, prvega števila v I_2, in tako naprej. Zaporedje realnih števil $f_i(x_j), i \in I$, konvergira k α_j za vsak j. Končno, zaporedje funkcij f_i, $i \in I$, konvergira po točkah na C''. $\square$

6 Diferencialnost konveksnih funkcij

Kot bomo videli v tem poglavju, imajo konveksne funkcije veliko uporabnih diferenciabilnih lastnosti. Ena od teh pravi da enostranski smerni odvodi obstajajo povsod. Bomo spoznali kako lahko opišemo enostranske smerne odvode s pomočjo koncepta subgradienta, analognem konceptu gradienta za dvostranske smerne odvode.

Definicija 6.1. Naj bo f poljubna funkcija iz $\mathbb{R}^n$ v $[-\infty, +\infty]$, in naj bo x točka kjer je f končna. *Enostranski smerni odvod* f v x glede na vektor y definiramo kot limito

$$f'(x; y) = \lim_{\lambda \downarrow 0} \frac{f(x + \lambda y) - f(x)}{\lambda},$$

seveda če limita sploh obstaja ($+\infty$ in $-\infty$ sta lahko limita).
Dvostranski smerni odvod definiramo kot limito

$$\lim_{\lambda \to 0} \frac{f(x + \lambda y) - f(x)}{\lambda}.$$

Opazimo, da

$$-f'(x; -y) = \lim_{\lambda \uparrow 0} \frac{f(x + \lambda y) - f(x)}{\lambda}.$$

Sledi, da je enostranski smerni odvod $f'(x; y)$ dvostranski če in samo če $f'(x; -y)$ obstaja in

$$f'(x; -y) = -f'(x; y).$$

Izrek 6.2. *Naj bo f konveksna funkcija, in naj bo x točka v kateri je f končna. Za vsak y, je diferenčni kvocient v 6.1 je nenaraščajoča funkcija $\lambda > 0$, torej $f'(x; y)$ obstaja in sicer*

$$f'(x; y) = \inf_{\lambda > 0} \frac{f(x + \lambda y) - f(x)}{\lambda}.$$

Še več, $f'(x; y)$ je pozitivno homogena konveksna funkcija y, $f'(x; 0) = 0$ in

$$-f'(x; -y) \leq f'(x; y), \quad \forall y.$$

Dokaz. Diferenčni kvocient za $\lambda > 0$ lahko izrazimo v obliki $\lambda^{-1} h(\lambda y)$, kjer je $h(y) = f(x + y) - f(x)$. Zaradi

$$h(\lambda y_1 + (1 - \lambda) y_2) = f(x + \lambda y_1 + (1 - \lambda) y_2) - f(x)$$

$$\leq f(x) + \lambda f(y_1) + (1-\lambda)f(y_2) - f(x) = \lambda f(y_1) + (1-\lambda)f(y_2)$$

in

$$\lambda h(y_1) + (1-\lambda)h(y_2) = \lambda f(x+y_1) - \lambda f(x) + (1-\lambda)f(x+y_2) - (1-\lambda)f(x)$$

$$= \lambda f(x+y_1) + (1-\lambda)f(x+y_2) - f(x) \leq \lambda f(y_1) + (1-\lambda)f(y_2),$$

je funkcija h konveksna. Konveksno množico epi$h$ dobimo premikanjem epi$f$ tako da točko $(x, f(x))$ damo v $(0,0)$. Velja pa $\lambda^{-1}h(\lambda y) = (h\lambda^{-1})(y)$, kjer je $h\lambda^{-1}$ po definiciji konveksna funkcija katere epigraf je λ^{-1}epi$h$. Naj bo $\lambda^{-1} < \delta^{-1}$ in $(x,\mu) \in \lambda^{-1}$epi$h$, oz. $(\lambda x, \lambda\mu) \in$ epi$h$. Opazimo, da je $\kappa := \frac{\lambda^{-1}}{\delta^{-1}} < 1$. Ker je epi$h$ konveksna in ker vsebuje koordinatno izhodišče $(0,0)$ in točko $(\lambda x, \lambda\mu)$, sledi da tudi vsebuje točko

$$(1-\kappa)(0,0) + \kappa(\lambda x, \lambda\mu) = \delta(x,\mu).$$

To je ekvivalentno pogoju $(x,\mu) \in \delta^{-1}$epi h, se pravi, smo dokazali da λ^{-1}epi$h$ narašča, takoj ko λ^{-1} narašča. Z drugimi besedami, za vsak y, diferenčni kvocient $(h\lambda^{-1})(y)$ se zmanjša takoj ko se λ zmanjša. Sledi,

$$\exists \inf_{\lambda>0}(h\lambda^{-1})(y) = f'(x;y), \quad \forall y.$$

Se pravi, smerni odvod funkcije $f'(x;\cdot)$ obstaja in je pozitivna homogena konveksna funkcija ki je generirana z h. Po definiciji je $f'(x;0) = 0$. Poleg tega, za $\mu_1 > f'(x;-y)$ in $\mu_2 > f'(x;y)$, velja

$$\frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 \geq f'(x; \frac{1}{2}(-y) + \frac{1}{2}y) = 0$$

zaradi konveksnosti. Torej, $-f'(x;-y) \leq f'(x;y)$ za vsak y. □

Na podlagi prejšnjega izreka sklepamo, da v slučaju ko je f konveksna funkcija na $\mathbb{R}$, lahko definiramo levi in desni odvod, ki popolnoma opisujejo smerne odvode f v x.

Definicija 6.3. Za konveksno funkcijo f na $\mathbb{R}$, *desni odvod* definiramo kot

$$f'_+(x) = f'(x;1) = \lim_{\lambda\downarrow 0} \frac{f(x+\lambda) - f(x)}{\lambda},$$

Analogno temu, *levi odvod* se definira kot

$$f'_-(x) = -f'(x;-1) = -\lim_{\lambda\downarrow 0} \frac{f(x-\lambda) - f(x)}{\lambda},$$

Definicija 6.4. Pravimo, da je vektor x^* *subgradient* konveksne funkcije $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ v točki x če velja

$$f(z) \geq f(x) + \langle x^*, z - x\rangle, \quad \forall z.$$

Zadnji neenakosti pravimo *subgradientska neenakost.*

Definicija 6.5. Množico vseh subgradientov funkcije f v točki x imenujemo *subdiferencial* funkcije f v točki x, in jo označamo z $\partial f(x)$. Preslikava $\partial f : x \longrightarrow \partial f(x)$ se imenuje *subdiferencial funkcije* f. Če je $\partial f(x)$ neprazna množica, rečemo da je f *subdiferenciabilna* v x.

Spomnimo se, da je podporna funkcija $\delta^*(\cdot|C)$ konveksne množice C v $\mathbb{R}^n$ definirana kot

$$\delta^*(x|C) = \sup\{\langle x, y\rangle \big| y \in C\}.$$

Izrek 6.6. *Za konveksno funkcijo f, in točko x v kateri je f končna, velja da je x^* subgradient funkcije f v x če in samo če je*

$$f'(x; y) \geq \langle x^*, y\rangle, \quad \forall y.$$

Lahko povemo še več, in sicer da je zaprtje konveksne funkcije $y \longrightarrow f'(x; y)$ podporna funkcija zaprte konveksne množice $\partial f(x)$.

Dokaz. Naj bo $z = x + \lambda y$. Subgradientska neenakost se spremeni v obliko

$$\frac{f(x + \lambda y) - f(x)}{\lambda} \geq \langle x^*, y\rangle$$

za vsak y in $\lambda > 0$. Diferenčni kvocient se zmanjša do $f'(x; y)$ ko $\lambda \downarrow 0$, ta neenakost je ekvivalentna tisti v izreku. Če uporabimo 4.44 na pozitivno homogeno funkcijo $f'(x; y)$ nam zagotavlja da izrek drži. □

Izrek 6.7. *Naj bo f prava konveksna funkcija. Za $x \notin domf$, je $\partial f(x)$ prazna. Za $x \in ri(domf)$, je $\partial f(x)$ neprazna, $f'(x; y)$ je zaprta in prava kot funkcija y in*

$$f'(x; y) = \sup\{\langle x^*, y\rangle \big| x^* \in \partial f(x)\} = \delta^*(y \big| \partial f(x)).$$

Končno, $\partial f(x)$ je neprazna omejena množica če in samo če je $x \in int(domf)$. V tem primeru velja da je $f'(x; y)$ končna za vsak y.

Dokaz. Če vstavimo $x \in \text{dom}f$ v subgradientsko neenakost, se iskaže da neenakosti ne more biti zadoščeno za noben x^* ko je $f(x) = +\infty$, saj bi tedaj za vsak z imeli: $f(z) \geq +\infty + \langle x^*, z - x\rangle$, kar ni možno ker je f prava konveksna funkcija. Za $x \in \text{ri}(\text{dom}f)$ je po 6.2, efektivna domena $f'(x; \cdot)$ afina množica, podprostor paralelen afini ogrinjači dom$f$. Ker $f'(x; \cdot)$ izgine v koordinatnem izhodišču, sledi da ne more biti enakaa $-\infty$ na tisti afini množici. Torej je po 4.22 $f'(x; \cdot)$ prava in po 4.30 zaprta. Po 6.6 je tedaj $f'(x; \cdot)$ podporna funkcija $\partial f(x)$, kar nam zagotavlja 6.7 in nepraznost $\partial f(x)$. Če je $\text{ri}(\text{dom}f) = \text{int}(\text{dom}f)$, efektivna domena $f'(x; \cdot)$ je cel prostor, tako da je podporna funkcija $\delta^*(\cdot|\partial f(x))$ končna povsod. Po drugi strani, ker je $\delta^*(\cdot|\partial f(x))$ zaprtje $f'(x; \cdot)$, potem sledi da če je $\delta^*(\cdot|\partial f(x))$ končna povsod tudi $f'(x; \cdot)$ mora biti končna povsod, kar po 4.11 pomeni, da je $x \in \operatorname{int}(\operatorname{dom} f)$. Zadnji del izreka zdaj sledi iz dejstva da je neprazna konveksna množica omejena če in samo če je njena podporna funkcija končna povsod (4.45). □

Izrek 6.8. *Naj bo f zaprta prava konveksna funkcija na $\mathbb{R}$. Podaljšajmo levi in desni odvod (sta dobro definirana povsod na dom$f$ za pravo funkcijo f po 6.2) f'_- in f'_+ izven intervala dom$f$ tako da sta oba enaka $+\infty$ za točke ki ležijo na desni strani dom$f$ in $-\infty$ za točke ki ležijo na levi strani dom$f$. Potem sta f'_- in f'_+ nepadajoče funkcije na $\mathbb{R}$, končne na notranjosti dom$f$ tako da velja*

$$f'_+(z_1) \leq f'_-(x) \leq f'_+(x) \leq f_-(z_2), \qquad z_1 < x < z_2.$$

Še več, za vsak x velja

$$\lim_{z \downarrow x} f'_+(z) = f'_+(x), \quad \lim_{z \uparrow x} f'_+(z) = f'_-(x)$$

$$\lim_{z \downarrow x} f'_-(z) = f'_+(x), \quad \lim_{z \uparrow x} f'_-(z) = f'_-(x).$$

Dokaz. Za vsak $x \in \operatorname{dom} f$, imamo po definiciji

$$f'_+(x) = \lim_{z \downarrow x} \frac{f(z) - f(x)}{z - x} = \lim_{\lambda \downarrow 0} \frac{f(x + \lambda) - f(x)}{\lambda} = f'(x; 1),$$

$$f'_-(x) = \lim_{z \downarrow x} \frac{f(z) - f(x)}{z - x} = \lim_{\lambda \downarrow 0} \frac{f(x - \lambda) - f(x)}{-\lambda} = -f'(x; -1).$$

Po 6.2, te limite obstajajo v monotono padajočem in monotono naraščajočem smislu, in $f'_-(x) < f'_+(x)$. Je očitno, zaradi monotonosti diferenčnih kvocientov, da $f'_+(x) < \infty$ če in samo če je x na levi strani desne končne točke cl(dom$f$) in $f'_-(x) > -\infty$ če in samo če je x na desni strani leve končne točke cl(dom$f$). Točke, kjer sta f'_+ in f'_- obe končni funkciji sta natančno tiste v int(dom$f$). Če sta y in z oba v dom$f$ in $y < z$, imamo

$$f'_+(y) \leq \frac{f(z) - f(y)}{z - y} = \frac{f(y) - f(z)}{y - z} \leq f'_-(z).$$

Če y in z nista oba v dom$f$, in $y < z$, potem po definiciji sledi, da $f'_+(y) \leq f'_-(z)$. Smo dokazali trojno neenakost. Neenakost pomeni da sta f'_+ in f'_- nepadajoči funkciji. Nadalje imamo

$$f'_+(x) \leq \lim_{z \downarrow x} f'_-(z) \leq \lim_{z \downarrow x} f'_+(z).$$

Zadošča pokazati, da druga limita ni večja od $f'_+(x)$ v slučaju ko dom$f$ vsebuje interval $(x, x + \varepsilon)$ za nekateri $\varepsilon > 0$, če bi hoteli pokazati da velja enakost. V nasprotnem nam definiciji f'_+ in f'_- zagotavljata enakost. v tem primeru limita $f(z)$ ko $z \downarrow x$ je $f(x)$ po 4.32 tako da za $x < y < x + \varepsilon$ imamo

$$\frac{f(y) - f(x)}{y - x} = \lim_{z \downarrow x} \frac{f(y) - f(z)}{y - z} \geq \lim_{z \downarrow x} f'_+(z)$$

Odtod pride

$$\lim_{z\downarrow x} f'_+(z) \leq \lim_{y\downarrow x} \frac{f(y)-f(x)}{y-x} = f'_+(x)$$

□

Na podlagi tega izreka lahko sklepamo da za vsak x velja

$$\partial f(x) = \{x^* \in \mathbb{R} \big| f'_-(x) \leq x^* \leq f'_+(x)\}.$$

Primer 6.9. Naj bo f je zaprta prava konveksna funkcija na $\mathbb{R}$ definirana kot:

$$f(x) = \begin{cases} |x| - 2(1-x)^{1/2} & \text{če } -3 \leq x \leq 1, \\ +\infty & \text{sicer.} \end{cases}$$

v tem primeru imamo

$$f'_+(x) = \begin{cases} +\infty & \text{če } x \geq 1, \\ 1 + (1-x)^{-1/2} & \text{če } 0 \leq x < 1, \\ -1 + (1-x)^{-1/2} & \text{če } -3 \leq x < 0, \\ -\infty & \text{če } x < -3, \end{cases}$$

$$f'_-(x) = \begin{cases} +\infty & \text{če } x \geq 1, \\ 1 + (1-x)^{-1/2} & \text{če } 0 < x < 1, \\ -1 + (1-x)^{-1/2} & \text{če } -3 < x \leq 0, \\ -\infty & \text{če } x \leq -3, \end{cases}$$

tako da

$$\partial f(x) = \begin{cases} \emptyset & \text{če } x \geq 1, \\ \{1 + (1-x)^{-1/2}\} & \text{če } 0 < x < 1, \\ [0,2] & \text{če } x = 0, \\ \{-1 + (1-x)^{-1/2}\} & \text{če } -3 < x < 0, \\ [-\infty, -\frac{1}{2}] & \text{če } x = -3, \\ \emptyset & \text{če} x < -3. \end{cases}$$

Primer 6.10. Izrek ne velja ko f ni zaprta funkcija! Npr. naj bo

$$f(x) = \begin{cases} 0 & \text{če } x > 0, \\ 1 & \text{če } x = 0, \\ +\infty & \text{če } x < 0. \end{cases}$$

v tem primeru je

$$f'_+(x) = \begin{cases} 0 & \text{če } x > 0, \\ -\infty & \text{če } x \leq 0, \end{cases}$$

tj. $f'_+(x)$ ni desno zvezna v 0.

Lema 6.11. *Naj bo f konveksna funkcija na $\mathbb{R}^n$ in C odprta konveksna množica na kateri je f končna. Naj bo $f_1, f_2, \ldots,$ zaporedje konveksnih funkcij končnih na C ki konvergirajo po točkah k funkciji f na C. Naj bo $x \in C$ in $x_1, x_2, \ldots$ zaporedje točk v C ki konvergirajo proti x. Potem za poljuben $y \in \mathbb{R}^n$ in poljubno zaporedje $y_1, y_2, \ldots$ ki konvergira proti y velja*

$$\limsup_{i\to\infty} f'_i(x_i; y_i) \leq f'(x; y).$$

Še več, za poljuben $\varepsilon > 0$, obstaja indeks i_0 takšen da je

$$\partial f_i(x_i) \subset \partial f(x) + \varepsilon B, \quad \forall i \geq i_0,$$

kjer je B evklidska enotska krogla v $\mathbb{R}^n$.

Dokaz. Za poljuben $\mu > f'(x; y)$, obstaja $\lambda > 0$ takšen da je $x + \lambda y \in C$ in

$$\frac{f(x + \lambda y) - f(x)}{\lambda} < \mu$$

Po 5.9, $f_i(x_i + \lambda y_i) \longrightarrow f(x + \lambda y)$ ko $f_i(x_i) \longrightarrow f(x)$. Torej, za dovolj velike i imamo

$$\frac{f_i(x_i + \lambda y_i) - f_i(x_i)}{\lambda} < \mu.$$

Ker je

$$f'_i(x_i; y_i) \leq \frac{f_i(x_i + \lambda y_i) - f_i(x_i)}{\lambda},$$

sledi, da je $\limsup_{i\to\infty} f'_i(x_i; y_i) \leq \mu$. To drži za poljuben $\mu > f'(x; y)$. Nam je uspelo pokazati prvi del. Konveksni funkciji $f'_i(x_i; \cdot)$ in $f'(x; \cdot)$ sta podporni funkciji nepraznih zaprtih omejenih konveksnih množic $\partial f_i(x_i)$ in $\partial f(x)$, respektivno, po 6.7 in so zaradi tega končne povsod na $\mathbb{R}^n$. Torej za poljuben $\varepsilon > 0$, po 5.10 obstaja indeks i_0 takšen da

$$f'_i(x_i; y) \leq f'(x; y) + \varepsilon, \quad \forall y \in B, \quad \forall i \geq i_0.$$

Zaradi pozitivne homogenosti velja

$$f'_i(x_i; y) \leq f'(x; y) + \varepsilon|y|, \quad \forall y \in \mathbb{R}^n, \quad \forall i \geq i_0,$$

ali, z drugimi besedami,

$$\delta^*(y|\partial f_i(x_i)) \leq \delta^*(y|\partial f(x)) + \varepsilon\delta^*(y|B)$$

$$= \delta^*(y|\partial f(x) + \varepsilon B), \quad \forall y \in \mathbb{R}^n, \quad \forall i \geq i_0.$$

Po 4.42, to pomeni

$$\partial f_i(x_i) \subset \partial f(x) + \varepsilon B, \quad \forall i \geq i_0.$$

□

Posledica 6.12. *Če je f prava konveksna funkcija na $\mathbb{R}^n$, potem je $f'(x;y)$ navzgor polzvezna funkcija $(x,y) \in int(domf) \times \mathbb{R}^n$. Še več, za poljuben $x \in int(domf)$ in $\varepsilon > 0$, obstaja $\delta > 0$ takšen da je*

$$\partial f(z) \subset \partial f(x) + \varepsilon B, \quad \forall z \in (x + \delta B),$$

kjer je B evklidska enotska kroglav $\mathbb{R}^n$.

Dokaz. Po 6.11. Naj bo $C = \mathrm{int}(\mathrm{dom} f)$ in $f_i = f$ za vsak i. □

Definicija 6.13. Naj bo f funkcija iz $\mathbb{R}^n$ v $[-\infty, +\infty]$ in x točka v kateri je f končna. Pravimo, da je f *diferenciabilna* v x če in samo če obstaja vektor x^* z lastnostjo

$$\lim_{z \to x} \frac{f(z) - f(x) - \langle x^*, z - x\rangle}{|z - x|} = 0.$$

Če takšen x^* obstaja, ga imenujemo *gradient* f v x in uporabljamo oznako $\bigtriangledown f(x)$ zanj.

Naj bo funkcija f diferenciabilna v x. Potem, po definiciji, za vsak $y \neq 0$,

$$0 = \lim_{\lambda \downarrow 0} \frac{f(x + \lambda y) - f(x) - \langle \bigtriangledown f(x), \lambda y\rangle}{\lambda |y|}$$

$$= \frac{f'(x;y) - \langle \bigtriangledown f(x), y\rangle}{|y|}$$

Se pravi, $f'(x;y)$ obstaja kot linearna funkcija y:

$$f'(x;y) = \langle \bigtriangledown f(x), y\rangle, \quad \forall y.$$

Za $j = 1, \ldots, n$,

$$\langle \bigtriangledown f(x), e_j\rangle = \lim_{\lambda \to 0} \frac{f(x + \lambda e_j) - f(x)}{\lambda} = \frac{\partial f}{\partial \varepsilon_j}(x),$$

kjer je e_j vektor iz j-te vrstice $n \times n$ identitete, in ε_j j-ta komponenta x. Odtod pa sledi

$$\bigtriangledown f(x) = (\frac{\partial f}{\partial \varepsilon_1}(x), \ldots, \frac{\partial f}{\partial \varepsilon_n}(x)),$$

tako da za vsak $y = (\eta_1, \ldots, \eta_n)$ velja

$$f'(x;y) = \frac{\partial f}{\partial \varepsilon_1}(x)\eta_1 + \cdots + \frac{\partial f}{\partial \varepsilon_n}(x)\eta_n.$$

To nam zagotavlja enoličnost vektorja x^*.Vprašamo se, ali obstaja kakšna zveza med konceptimi gradienta in subgradienta. Se iskaže da ja, in sicer zelo enostavna.

Izrek 6.14. *Naj bo f konveksna funkcija, in x točka v kateri je f končna. Če je f diferenciabilna v x, potem je subgradient f v x enak $\bigtriangledown f(x)$, tako da velja*

$$f(z) \geq f(x) + \langle \bigtriangledown f(x), z - x \rangle, \quad \forall z.$$

Obratno, če f ima natančno določen subgradient v x, potem je f diferenciabilna v x.

Dokaz. Najprej, recimo da je f diferenciabilna v x. Potem je $f'(x;\cdot)$ linearna funkcija $\langle \bigtriangledown f(x), \cdot \rangle$. Po 6.6, subgradienti v x so vektorji x^* takšni da

$$\langle \bigtriangledown f(x), y \rangle \geq \langle x^*, y \rangle \quad \forall y,$$

ta pogoj pa je zadoščen če in samo če je $x^* = \bigtriangledown f(x)$ (če bi imeli $x^* \neq \bigtriangledown f(x)$, potem vektorji y_1 in $y_2 = -y_1$ ne bi hkrati ustrezali 6). Torej, $\bigtriangledown f(x)$ je natančno določen subgradient f v x. Recimo po drugi strani da f ima natančno določen subgradient x^* v x. Vstavimo v definicijo subgradienta $z = x + y$ in definirajmo konveksno funkcijo g kot (afine funkcije so konveksne, vsota konveksnih funkcij je konveksna, se pravi: g je konveksna)

$$g(y) = f(x+y) - f(x) - \langle x^*, y \rangle.$$

Potem g ima 0 kot svoj natančno določeni subgradient v koordinatnem izhodišču. Pokažimo, da to pomeni

$$\lim_{y \to 0} \frac{g(y)}{|g|} = 0.$$

Zaprtje $g'(0;\cdot)$ je podporna funkcija množice $\partial g(0)$. Po definiciji je podporna funkcija enaka $\delta^*(x|\{0\}) = sup\{\langle x, y \rangle | y \in \{0\}\} = 0$. Se pravi, $g'(0;\cdot)$ je enaka 0, ker $g'(0;\cdot)$ ne more biti različna od svojega zaprtja razen na točkah na meji svoje efektivne domene. Torej imamo

$$0 = g'(0;u) = \lim_{\lambda \downarrow 0} \frac{g(\lambda u) - g(0)}{\lambda} \quad \forall u.$$

V tej enakosti je $g(0) = 0$ in diferenčni kvocient je nenaraščajoča funkcija λ. Konveksne funkcije h_λ, ki jih definiramo kot

$$h_\lambda(u) = \frac{g(\lambda u)}{\lambda}, \quad \lambda > 0$$

torej padajo po točkah proti konstantni funkciji 0 ko λ gre proti 0. Naj bo B evklidska enotska krogla in $\{a_1, \ldots, a_m\}$ poljubna končna kolekcija točk katerih konveksna ogrinjača vsebuje B. Vsak $u \in B$ se lahko izraža kot konveksna kombinacija

$$u = \lambda_1 a_1 + \cdots + \lambda_m a_m,$$

in potem velja

$$0 \leq h_\lambda(u) \leq \sum_{i=1}^{m} \lambda_i h_\lambda(a_i)$$

$$\leq max\{h_\lambda(a_i)\big| i = 1, \dots, m\}.$$

Ker $h_\lambda(a_i) \downarrow 0$ za vsak i ko $\lambda \downarrow 0$, lahko sklepamo da $h_\lambda(u) \downarrow 0$ enakomerno za vsak $u \in B$ ko $\lambda \downarrow 0$. Za poljuben $\varepsilon > 0$ torej obstaja $\delta > 0$, tako da

$$\frac{g(\lambda u)}{\lambda} \leq \varepsilon, \quad \forall \lambda \in (0, \delta], \quad \forall u \in B.$$

Ker vsak y ki zadošča $0 < |y| < \delta$ lahko izrazimo kot λu za $\lambda = |y|$ in $u \in B$, imamo $\frac{g(y)}{|y|} \leq \varepsilon$ kadarkoli $0 < |y| \leq \delta$. Končno, sklepamo da je limita $\lim_{y \to 0} \frac{g(y)}{|y|}$ enaka 0. □

Posledica 6.15. *Naj bo f konveksna funkcija. Če je f končna in diferenciabilna v dani točki x, potem je f prava in $x \in int(domf)$.*

Dokaz. Neenakost v prejšnjem izreku pomeni da je $f(z) > -\infty$ za vsak z, in da je torej f prava. Očitno je, iz same definicije diferenciabilnosti, da če je f diferenciabilna v x, potem mora biti končna v nekateri okolici x. □

Izrek 6.16. *Naj bo f konveksna funkcija na $\mathbb{R}^n$, in x točka v kateri je f končna. Potreben in zadosten pogoj, da je f diferenciabilna v x je da je smerni odvod $f'(x; \cdot)$ linearen. Še več, ta pogoj je izpoljnjen če n dvostranskih parcialnih odvodov $\frac{\partial f(x)}{\partial \varepsilon_j}$ obstajajo v x in če so končni.*

Dokaz. Če je smerni odvod $f'(x; \cdot)$ linearen, je $f'(x; \cdot)$ zaprta konveksna funkcija in je po 6.6 enaka podporni funkciji $\partial f(x)$. Potem $\partial f(x)$ sestoji iz le ene točke po 4.44, kar po 6.14 pomeni da je f diferenciabilna v x. Naj bo $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ vektor z 1 na j-tem mestu. Imamo

$$f'(x; e_j) = \frac{\partial f}{\partial \varepsilon_j}(x) = -f'(x; -e_j), \quad j = 1, \dots, n$$

Efektivna domena $f'(x; \cdot)$ torej vsebuje $2n$ vektorjev e_j in $-e_j$ $(j = 1, \dots, n)$, in zato vse pozitivne večkratnike teh, zaradi pozitivne homogenosti. Ker je efektivna domena konveksna, mora biti enaka $\mathbb{R}^n$. Sledi, da je $f'(x; \cdot)$ prava, ker bi v nasprotnem bila enaka $-\infty$ po 4.22. Na koncu nam 3.35 poda linearnost. □

Izrek 6.17. *Naj bo f prava konveksna funkcija na $\mathbb{R}^n$. Za poljuben $y \neq 0$, naj bo D množica točk x v $int(domf)$ kjer je $f'(x; y) = -f'(x; -y)$, oz. kjer obstaja navadni dvostranski smerni odvod*

$$\lim_{\lambda \to 0} \frac{f(x + \lambda y) - f(x)}{\lambda}$$

Potem D sestoji prav iz točk iz $int(domf)$ kjer je $f'(x; y)$ zvezna kot funkcija x. Še več, D je gosta v $int(domf)$. Komplement D v $int(domf)$ je množica mere 0.

Dokaz. Po 6.12 točke v int(dom$f$) kjer je $f'(x;y)$ zvezna kot funkcija x su iste kot tiste kjer je navzdol polzvezna kot funkcija x. Trdimo, da

$$\liminf_{z\to y} f'(z;y) = -f'(x;-y), \quad \forall x \in \text{int}(\text{dom}f).$$

Dokaz te enakosti bo pomenil zveznost v izreku. Očitno, ker je $f'(z;y) \geq -f'(z;-y)$ za vsak z v dom$f$ in ker je $f'(z;-y)$ navzgor polzvezna v z na int(dom$f$), po 6.2 sledi da "$\geq$" del velja. Po drugi strani, ker za enodimenzionalno konveksno funkcijo $g(\lambda) = f(x+\lambda y)$ velja

$$\lim_{\lambda\uparrow 0} f'(x+\lambda y;y) = \lim_{\lambda\uparrow 0} g'_+(\lambda) = g'_-(0) = -f'(x;-y)$$

po 6.8, sklepamo, da je velja tudi "$\leq$" del. Komplement D v int(dom$f$) sestoji iz točk x v int(dom$f$) kjer

$$0 < f'(x;y) + f'(x;-y) = h(x)$$

Torej je unija naraščajočega zaporedja množic

$$S_k = \{x \in \text{int}(\text{dom}f) \big| h(x) \geq \frac{1}{k}\}, \quad k = 1, 2, \ldots$$

Kot vsota dveh navzgor polzveznih funkcij x, h je navzgor polzvezna na int(dom$f$). Torej vsaka S_k je zaprta relativno int(dom$f$), in je merljiva. Za $x \in \mathbb{R}^n$, naj bo L_x premica skozi x v smeri y. Naj se L_k in S_k sekata. Z omejitvijo f na L_x dobimo enodimenzionalno konveksno funkcijo g kot zgoraj, in točke $z = x + \lambda y \in L_x \cap S_x$ tako da $g'_+(\lambda) - g'_-(\lambda) \geq \frac{1}{k}$. Neenakost v 6.8 poskrbi, da ne more obstajati več kot končno mnogo takšnih točk v poljubnem omejenem intervalu. Smo dokazali, da vsaka S_k ima potrebno lastnost preseka, kar pomeni da S_k ima mero 0. (Mero S_k lahko dobimo tako, da integriramo mero $L_k \cap S_k$ kot funkcijo $x \in S'_k$, kjer je S'_k projekcija S_k na podprostor $\mathbb{R}^n$ ki je ortogonalen y). Ker je komplement D v int(dom$f$) unija množic S_k, je tudi mere 0. Torej, ne more imeti notranjih točk in D je končna v int(dom$f$). □

Najbolj pomemben izrek o gradientu konveksne funkcije je naslednji.

Izrek 6.18. *Naj bo f prava konveksna funkcija na $\mathbb{R}^n$ in D množica točk kjer je f diferenciabilna. Potem je D gosta množica v int(dom$f$), in njen komplement v int(dom$f$) je množica mere 0. Poleg tega, $\bigtriangledown f : x \to \bigtriangledown f(x)$ je zvezna na D.*

Dokaz. Naj bodo $e_1, \ldots, e_n$ vrstice $n \times n$ identitetne matrike. Z uporabo 6.17 na $y = e_j$ sklepamo, da podmnožica D_j v int(dom$f$) kjer obstaja $\frac{\partial f}{\partial \varepsilon_j}$ ima komplement ki ima mero 0. Unija teh komplementov za $j = 1, \ldots, n$ tudi ima mero 0. Ta pa je komplement množice $D_1 \cap \cdots D_n$, a slednja množica je prav zaprav D po 6.16. Komplement D v int(dom$f$) nima notranjosti, oz. D je gosta v int(dom$f$). Vsak parcialni odvod $\frac{\partial f}{\partial \varepsilon_j}$ je zvezen na pripadajoči množici D_j, po 6.17, in so torej vsi zvezni na D. Ker je $\bigtriangledown f(x)$ vektor prvih parcialnih odvodov v točkah kjer obstaja, sledi da je funkcija $\bigtriangledown f$ zvezna na D. $\square$

Smo prišli do konca tega poglavja in do najbolj presentljivega rezultata. Namreč, vemo da v splošnem za zaporedje diferenciabilnih funkcij $f_1, f_2 \ldots$ na odprtem intervalu I ki konvergirajo po točkah proti diferenciabilni funkciji f na I ne velja nujno da zaporedje odvodov $f_1', f_2', \ldots$ konvergira proti f'. Če pa so funkcije $f_1, f_2 \ldots$ konveksne, ne samo da konvergirajo zaporedje odvodov $f_1', f_2', \ldots$ konvergira k f', ampak konvergira enakomerno na vsakem zaprtem omejenem podintervalu I. To zelo pomembno dejstvo je posledica naslednjega izreka.

Izrek 6.19. *Naj bo C odprta konveksna množica in f konveksna funkcija ki je končna in diferenciabilna na C. Naj bo $f_1, f_2 \ldots$ zaporedje konveksnih funkcij ki so končne in diferenciabilne na C tako da $\lim_{i\to\infty} f_i(x) = f(x)$ za vsak $x \in C$. Potem*

$$\lim_{i\to\infty} \bigtriangledown f_i(x) = \bigtriangledown f(x), \quad \forall x \in C.$$

Preslikave $\bigtriangledown f_i$ konvergirajo enakomerno k $\bigtriangledown f$ na vsaki zaprti omejeni podmnožici množice C.

Dokaz. Naj bo S zaprta omejena podmnožica v C. Zadošča pokazati, da parcialni odvodi f_i konvergirajo enakomerno k parcialnim odvodom f na S, oz. da za poljuben vektor y in $\varepsilon > 0, \exists i_0$:

$$|\langle \bigtriangledown f_i(x), y\rangle - \langle \bigtriangledown f(x), y\rangle| \leq \varepsilon, \quad \forall i \geq i_0, \quad \forall x \in S.$$

To neenakost lahko zapišemo kot par neenakosti

$$\langle \bigtriangledown f_i(x), y\rangle \leq \langle \bigtriangledown f(x), y\rangle + \varepsilon, \langle \bigtriangledown f_i(x), -y\rangle \leq \langle \bigtriangledown f(x), -y\rangle + \varepsilon.$$

Pokažimo, da obstaja i_1 takšen da prva neenakost velja za vsak $i \geq i_1$ in $x \in S$. Podobno, poiščimo i_2 za drugo neenakost in naj bo potem $i_0 = \max\{i_1, i_2\}$. A takšen i_1 nujno obstaja? Denimo nasprotno. Potem obstaja neskončno mnogo indeksov i za katere lahko izberemo pripadajoči vektor $x_i \in S$ takšen da

$$\langle \bigtriangledown f_i(x_i), y\rangle > \langle \bigtriangledown f(x_i), y\rangle + \varepsilon.$$

Recimo, da to velja za vsak indeks i, in da zaporedje $x_1, x_2, \ldots$ konvergira k x v S. Za $\lambda > 0$ dovolj majhen da $x + \lambda y \in C$, imamo za dovolj velike i $x_i + \lambda y \in C$ in

$$\langle \bigtriangledown f_i(x_i), y\rangle \leq \frac{f_i(x + \lambda y) - f_i(x_i)}{\lambda}.$$

Po 5.9, funkcije f_i konvergirajo enakomerno k f na vsaki zaprti omejeni podmnožici množice C. Ker je f zvezna na C, to pomeni $f_i(x_i) \to f(x)$ in $f_i(x_i + \lambda y) \to f(x + \lambda y)$. Ker je po 6.18 $\triangledown f$ zvezna, $\triangledown f(x_i) \to \triangledown f(x)$. Torej

$$\langle \triangledown f(x), y\rangle + \varepsilon = \lim_{i\to\infty} \langle \triangledown f(x_i), y\rangle + \varepsilon \leq \limsup_{i\to\infty} \langle \triangledown f_i(x_i), y\rangle$$

$$\leq \lim_{i\to\infty} \frac{f_i(x_i + \lambda y) - f_i(x_i)}{\lambda} = \frac{f(x + \lambda y) - f(x)}{\lambda}$$

To naj bi veljalo za vsak dovolj majhen $\lambda > 0$. Ampak,

$$\langle \triangledown f(x), y\rangle = f'(x; y) = \lim_{\lambda\downarrow 0} \frac{f(x + \lambda y) - f(x)}{\lambda},$$

kar nas privede do protislovja. $\square$

7 Zaključek

V zaključni nalogi sem predvsem obravnal konveksne množice in konveksne funkcije. Sem začel z najbolj osnovnimi definicijami in lastnostimi, kot so afine množice, afine ogrinjače, konveksne množice, konveksne ogrinjače, stožci, in podobno, da bi, sprehajajoč se skozi topološko stran snovi in koncepte kot so relativna notranjost, relativna meja, efektivna domena in zaprtje prišel do bolj kompleksnih dejstev o zveznosti in diferenciabilnosti konveksnih funkcij. Naloga je bila zelo bogata konceptimi, trditvami in idejami; glede na to, naj bi vsak bralec imel bolj globoko razumevanje zanimivega matematičnega fenomena kot so konveksne funkcije. Ni bilo dovolj prostora, za obravnavanje številnih aplikacij konveksnosti v uporabni matematiki, vendar si bralec lahko prebere več o tem v [4, 5, 10, 12].

Kot smo že na začetku omenili, smo delali v prostoru $\mathbb{R}^n$. Poleg tega pa obstaja bogata teorija kompleksne konveksnosti. Radoveden bralec si lahko prebere več o tem v [1].

8 Literatura

[1] M. Andersson in M. Passare, R. Sigurdsson. V: Complex Convexity, *Complex Convexity and Analytic Functionals*, Birkhäuser Verlag AG, 2004, 15–72. *(Citirano na strani 59.)*

[2] E. Asplund in R.T. Rockfellar, Gradients of convex functions. *The American Mathematical Monthly vol. 139* 1969 (443–467) *(Ni citirano.)*

Asplund

[3] M.. Berger, Convexity. *The American Mathematical Monthly* 97 (1990) 650–678. *(Ni citirano.)*

[4] J.M. Bismut, Conjugate Convex Functions in Optimal Stochastic Control. *Journal of Mathematical Analysis and Applications* 44 (1973) 384–404. *(Citirano na strani 59.)*

[5] J.M. Borwein in A.S. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples.* Canadian Mathematical Society-Springer Books, Vol. 3, 2000. *(Citirano na strani 59.)*

[6] S. Boyd in L. Vandenberghe, *Convex Optimization.* Cambridge University Press, Vol. 1, 2004. *(Ni citirano.)*

[7] W. Fenchel, On Conjugate Convex Functions. *Canadian Journal of Mathematics, Vol. 1* 1949 (73–77) *(Ni citirano.)*

Fenchel

[8] M. Giaquinta in G. Modica, *Mathematical Analysis: Foundations and Advanced Techniques for Functions of Several Variables.* Birkhäuser Verlag AG, Volume 5, 2012. *(Ni citirano.)*

[9] J.W. Green, Recent Applications of Convex Functions. *The American Mathematical Monthly, Vol. 61* No. 7 (1954) 449–454. *(Ni citirano.)*

[10] P.M. Gruber, Aspects of Convexity and its Applications. *Expo. Math. vol. 2* 1984 (47–83) *(Ni citirano.)*

Gruber

[11] G.H. Hardy, J.E. Littlewood in G. Pólya, Some simple inequalities satisfied by convex functions. *Messenger Math. vol. 58* 1929 (145–152) *(Ni citirano.)*

Hardy

[12] C. Niculescu in L.E. Persson, *Convex Functions and their Applications: A Contemporary Approach.* Canadian Mathematical Society-Springer Books, 2004. *(Citirano na strani 59.)*

[13] R.P. Phelps, *Convex Functions, Monotone Operators and Differentiability, 2nd Ed.*, Lecture Notes in Mathematics, vol. 1364. Springer-Verlag, Berlin 1993 *(Ni citirano.)*

[14] A.W. Roberts in D.E. Varberg, *Convex Functions.* Academic Press, New York, 1973. *(Ni citirano.)*

[15] R.T. Rockfellar, Conjugate Convex Functions in Optimal Control and the Calculus of Variations. *Journal of Mathematical Analysis and Applications* 32 (1970) 174–222. *(Citirano na strani 36.)*