Konveksnost i optimizacija

Slišković, Marijana

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Marijana Slišković

KONVEKSNOST I OPTIMIZACIJA

Diplomski rad

Voditelj rada:
prof. dr. sc. Marko Vrdoljak

Zagreb, veljača, 2017.

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

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

Povjerenstvo je rad ocijenilo ocjenom _______________.

Potpisi članova povjerenstva:

1. ____________________________
2. ____________________________
3. ____________________________

Ovaj rad posvećujem svojoj obitelji.

Sadržaj

Uvod

U diplomskom radu proučit ću konveksnost i optimizaciju. Teorija konveksnosti je relativno novo područje matematike u kojem se isprepliću svojstva iz geometrije, analize, linearne algebre, topologije i kombinatorike. Bazirana je na jednostavnom geometrijskom konceptu koji joj i daje ime. Teorija konveksnosti je fundamentalna u mnogim područjima optimizacije, a od važnosti je i u statistici, matematičkoj ekonomiji, funkcionalnoj analizi, teoriji aproksimacija ,itd. Ona je osnovni matematički alat za linearno programiranje. U ovom radu se ograničavamo na konveksnost u konačno-dimenzionalnom vektorskom prostoru R^n.

Poglavlje 1

Konveksni skupovi

Definicija i primjeri

Definicija 1.0.1. *Za $x, y \in R^n$ skup*

$$[x, y] := \{(1 - \lambda)x + \lambda y : \lambda \in [0, 1]\}$$

nazivamo segment (ili spojnica) s krajevima x i y.

Kažemo da je skup $K \subseteq \mathbb{R}^n$ konveksan ako vrijedi

$$(\forall x, y \in K)\, [x, y] \subseteq K$$

Primjer 1.0.2. *Prazan skup $\emptyset$ (po definiciji),jednočlan skup $\{a\}$ i R^n su konveksni skupovi.*

Primjer 1.0.3. *Kugla sa centrom u x_0 i polumjerom r:*

$$\bar{B}\,(x_0, r) = x_0 + r\bar{B} \ \text{ gdje je } \bar{B}_i = B\,(\bar{0}, 1) \text{ jedinična kugla.}$$

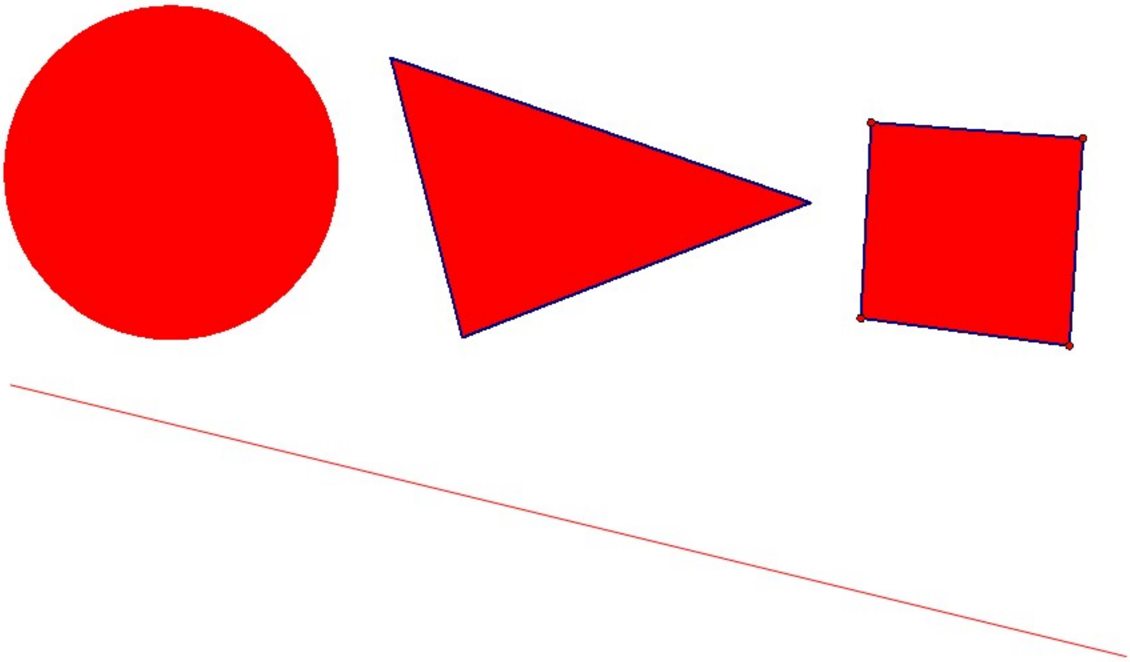
Zaista, za $x_1, x_2 \in B$, $\lambda \in [0, 1]$ imamo

$$||0 - (1 - \lambda)x_1 - \lambda x_2|| \leq (1 - \lambda)\,||x_1|| + \lambda\,||x_2|| \leq (1 - \lambda) * 1 + \lambda * 1 = 1$$

1.1 Geometrijsko svojstvo konveksnih skupova

Skup K je konveksan ako za svake svoje dvije točke sadrži i segment koji spaja te dvije točke. Očito je da je jednočlan skup konveksan. Prazan skup također smatramo konveksnim.

Definicija 1.1.1. *Konveksna kombinacija vektora $x_1, \ldots, x_m \in R^n$ je svaki vektor x oblika*



Slika 1.1: Primjeri konveknih skupova u ravnini

$$x = \sum_{i=1}^{m} \lambda_i m_i$$

gdje su $\lambda_i \geq 0$*,* $i = 1, \ldots, m$ *realni brojevi takvi da je* $\sum_{i=1}^{m} \lambda_i = 1$.
Skup svih konveksnih kombinacija točaka iz skupa $K \subseteq \mathbb{R}^n$ *naziva se konveksna ljuska skupa* K *i označava sa* $convK$ *.*

Propozicija 1.1.2. *Skup* $K \subseteq \mathbb{R}^n$ *je konveksan i ako i samo ako se podudara sa svojom konveksnom ljuskom.*

Dokaz. Ako je $K = convK$, budući da je $convK$ konveksan, K je konveksan skup. Obratno, pretpostavimo da je$K$ konveksan skup i pokažimo da je $K = convK$. Kako je $K \subseteq convK$ preostaje pokazati da je $convK \subseteq K$. Dokaz provodimo pomoću matematičke indukcije po m, gdje m označava broj točaka koje ulaze u konveksnu kombinaciju. Za $m = 1$ i $m = 2$ tvrdnja proizlazi iz definicije konveksnosti.

Pretpostavimo da je tvrdnja dokazana za prirodni broj m i pretpostavimo da je x konveksna kombinacija od $m + 1$ vektora iz S. Tada postoje vektori $x_1, \ldots, x_{m+1} \in K$ takvi da je

$$x = \sum_{i=1}^{m+1} \lambda_i x_i$$

Pri čemu je

$$\sum_{i=1}^{m+1} \lambda_i = 1$$

Zapišimo vektor $x = \mu y + \lambda_{m+1} x_{m+1}$
gdje su

$$\mu = \sum_{i=1}^{m} \lambda_i \text{ i } y = \sum_{i=1}^{m} \frac{\lambda_i}{\mu} x_i.$$

Budući je vektor y konveksna kombinacija od m vektora iz K, po pretpostavci indukcije on leži u skupu K. No tada i vektor x kao konveksna kombinacija vektora y i x_{m+1}
također pripada skupu K. □

Teorem 1.1.3. *Konveksna ljuska skupa $K \subseteq \mathbb{R}^n$ je najmanji konveksan skup koji sadrži K tj. presjek svih konveksnih skupova koji sadrže K.*

Dokaz. Neka je W presjek svih konveksnih skupova iz R^n koji sadrže skup K. Treba pokazati da je $convK = W$. Iz $K \subseteq W$ slijedi $convK \subseteq convW$, a zbog konveksnosti skupa W i propozicije 1.1.2 je $convW = W$. Dakle, $convK \subseteq W$. Kako je $K \subseteq convK$ konveksan skup te kako je$W$ najmanji konveksan skup koji sadrži K, slijedi $W \subseteq convK$. □

Primjer 1.1.4. *Neka su zadani vektor $a \in R^n$ i realan broj β. Skup*

$$H_{\alpha,\beta} = \{x \in R^n : \langle a, x_i \rangle = \beta\}$$

nazivamo hiperravnina, a skupovi

$$H^-_{\alpha,\beta} = \{x \in R^n : \langle a, x_i \rangle \leq \beta\}$$
$$H^+_{\alpha,\beta} = \{x \in R^n : \langle a, x_i \rangle \geq \beta\}$$

nazivaju se zatvoreni poluprostori odredeni hiperravninom $H_{\alpha,\beta}$. Vektor $a \in R^n$ zovemo vektor normale na hiperravnine $H_{\alpha,\beta}$. Hiperravnine i poluprostor su konveksni skupovi.

Primjer 1.1.5. *Neka su $A \in R^{n \times m}$, $b \in R^n$. Skup*

$$P = \{x \in R^n : Ax \leq b\}$$

zovemo poliedar u R^n.
Uočimo da je po definiciji poliedar presjek konačno mnogo zatvorenih poluprostora. Zato je on takoder konveksan i zatvoren, a zovemo ga još i konveksni poliedar.

1.2 Konveksna funkcija

Definicija 1.2.1. *Za funkciju $f : K \to R$ definiranu na konveksnom skupu $K \subseteq R^n$ kažemo da je konveksna ako za sve $x, y \in K$ i za svaki $\lambda \in [0, 1]$ vrijedi:*

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

Fukcija f je konkavna ako je funkcija - f konveksna.

Teorem 1.2.2. *Ako je f konveksna funkcija definirana na konveksnom skupu K, onda je skup K^* točaka u kojima f prima minimum konveksan i svaki lokalni minimum je ujedno i globalni minimum.*

Dokaz. Dokažimo prvo da je K^* konveksan skup. Ako je K^* prazan ili jednočlan skup, po definiciji on je konveksan. Neka su $x^*, y^* \in K^*$. To znači da vrijedi

$$f(x) \geq f(x^*) = f(y^*), \in K$$

odakle slijedi $\forall \lambda \in [0, 1]$ dobivamo:

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

odakle slijedi

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

i zato je:

$$f(x) \geq f(\lambda x^* + (1 - \lambda) y^*)\ \forall x \in K$$

to jest

$$\lambda x^* + (1 - \lambda) y^* \in K^*.$$

Sada ćemo pokazati da je svaki lokalni minimum $x^* \in K^*$ ujedno i globalni minimum. Pretpostavimo suprotno, tj. da postoji točka $y^* \in K^*$ za koju je $f(y^*) < f(x^*)$. Neka je $\epsilon > 0$ bilo koji realan broj, a $B(x^*, \epsilon)$ otvorena kugla radijusa ϵ s centrom u točki x^*. Pokazat ćemo da je

$$\lambda y^* + (1 - \lambda) x^* \in B(x^*, \epsilon), \forall \lambda \in (0, \epsilon / \|y^* - x^*\|).$$

Zbog konveksnosti funkcije f i $f(y^*) < f(x^*)$ za svaki realan broj $\lambda \in (0, \epsilon / \|y^* - x^*\|)$ što je kontradikcija činjenici da je x^* lokalni minimum.

□

Poglavlje 2

Konveksna ljuska i Carathéodoryjev teorem

2.1 Afina geometrija

Skup $K \subseteq \mathbb{R}^n$ je afin ako za sve $x, y \in K$ vrijedi:

$$\{(1-\lambda)\, x + \lambda y : \lambda \in R\} \subseteq K.$$

Najjednostavniji primjeri afinog skupa su pravac i hiperravnina.

Propozicija 2.1.1. *Neka je $M \subseteq R^n$. Sljedeće tvrdnje su ekvivalentne:*

1. *K je afin skup,*
2. *Za svaki $x_0 \in K$, skup $K - x_0$ je vektorski potprostor od R^n. Osim toga, za svaki $y_0 \in K$ je $K - y_0 = K - x_0$,*
3. *Postoji $r \in \mathbb{N}, r \leq n$, matrica $A \in R^{n\times m}$ ranga r i vektor $b \in R^n$ takvi da je sustav $Ax = b$ konzistentan i i $K = x \in R^n : Ax = b$.*

Definicija 2.1.2. *Pod dimenzijom afinog skupa K podrazumijevamo dimenziju vektorskog potprostora*

$$L := K - x_0$$

gdje je $x_0 \in K$ proizvljna točka. Dakle, $dimK := dimL$.

Definicija 2.1.3. *Afina kombinacija vektora $x_1, \ldots, x_m \in R^n$ je svaki vektor x oblika*

$$x = \sum_{i=1}^{m} \lambda_i m_i$$

gdje su λ_i realni brojevi takvi da je $\sum_{i=1}^{m} \lambda_i = 1$
Skup svih afinih kombinacija točaka iz skupa $K \subseteq \mathbb{R}^n$ naziva se afina ljuska skupa K označava s $\mathrm{aff}K$.

Definicija 2.1.4. *Skup $K \subseteq R^n$ je konus s vrhom u nuli ako za svaki $x \in K$ i za svaki $\lambda \geq 0$ vrijedi $\lambda x \in K$. Ako je $x_0 \in R^n$ i $K \in R^n$ konus s vrhom u nuli, onda skup*

$$K_{x_0} := x_0 + K = \{x_0 + x : x \in K\}$$

zovemo konus s vrhom u x_0. Ako je konus ujedno i konveksan skup, onda se on naziva konveksni konus.

Primjer 2.1.5. *Prazan skup takoder smatramo konusom.*

Primjer 2.1.6. *Za zadane x_0, $q \in R^n$, $q \neq 0$ skup*

$$\{\lambda q : \lambda \geq 0\}$$

je konveksan konus s vrhom u nuli, a skup

$$x_0 + \{\lambda q : \lambda \geq 0\} = \{x_0 + \lambda q : \lambda \geq 0\}$$

je konveksan konus s vrhom u x_0. Zovemo ga još i zraka ili polupravac u R^n, te kažemo da je točka x_0 ishodište zrake, a q vektor smjera te zrake.

Propozicija 2.1.7. *Skup $K \subseteq \mathbb{R}^n$ je afin ako i samo akoje $K = \mathrm{aff}K$.*

Teorem 2.1.8. *Afina ljuska skupa $K \subseteq \mathbb{R}^n$ je najmanji afin skup koji sadrži K, tj. presjek svih afinih skupova koji sadrže K.*

Definicija 2.1.9. *Za točke $x_1, \ldots, x_m \in R^n$ kažemo da su afino nezavisne ili u općem položaju ako su jednakosti:*

$$\sum_{i=1}^{m} \lambda_i x_i = 0 \quad i \quad \sum_{i=1}^{m} \lambda_i = 0$$

moguće jedino za $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$.
U suprotnom kažemo da su afino zavisne.

Definicija 2.1.10. *Za dimenziju skupa $K \subseteq \mathbb{R}^n$ uzima se dimenzija njegove afine ljuske, tj.*

$$dim K = dim\,(\mathrm{aff}K).$$

2.2 Topologija konveksnih skupova

Konveksnost je čisto algebarsko svojstvo, za njegovo potpunije razumijevanje potrebna je i topologija (inducirana euklidskom metrikom). U daljnjem tekstu koristit ćemo sljedeće oznake:

- Int za nutrinu (interior) skupa K,
- Cl za zatvarača skupa K,
- Bd za rub (granicu) skupa K.

Podsjetimo se da je$BdK = ClK \setminus IntK$.
Sa $B(x_0, \varepsilon)$ označavat ćemo otvorenu kuglu u R^n s centrom u $x_0 \in R^n$ i radijusom ϵ. U relativnoj topologiji na aff$K$ otvoreni skupovi su oblika $O \bigcap \text{aff}K$, gdje je O
otvoren u R^n.
Specijalno, skup

$$B(x_0, r) \bigcap \text{aff}K = \{x \in \text{aff}K : \|x - x_0\| < r\}$$

je otvorena kugla (s centrom u x_0 i radijusa $r > 0$) u relativnoj topologiji na aff$K$.
Interior skupa K u relativnoj topologiji aff$K$ označavat ćemo sa $RelIntK$, zatvarač s $RelClK$, a rub s $RelBdK$. Sve točke gomilanja skupa K nalaze se u aff$K$ i zato je $RelClK = ClK$ $RelBdK = ClK \setminus RelIntK$.

Lema 2.2.1. *Neka je $K = conv\{x_0, \ldots, x_m\}$ m-dimenzionalni simpleks iz R^n generiran s afino nezavisnim točkama $x_0, \ldots, x_m \in R^n$. Tada je$K$ zatvoren skup*

$$RelIntK = \{\textstyle\sum_{i=0}^m \lambda_i x_i : \sum_{i=0}^m \lambda_i = 1,\ \lambda_i > 0\ \forall i\},$$

$$RelBdK = \{\textstyle\sum_{i=0}^m \lambda_i x_i : \sum_{i=0}^m \lambda_i = 1,\ \lambda_i = 0 \text{ za barem jedan } i\}.$$

Dokaz. Uočimo da je

$$K = \{x_0 + \textstyle\sum_{i=0}^m \lambda_i (x_i - x_0) : \sum_{i=1}^m \lambda_i \le 1, \lambda_i \in [0, 1]\}$$

Sada definirajmo neprekidnu funkciju $f : R^m \to \text{aff}K$ formulom

$$f(\lambda_1, \ldots, \lambda_m) = x_0 + \textstyle\sum_{i=0}^m \lambda_i (x_i - x_0)$$

Tada je

$$f(\Lambda) = K$$

gdje je

$$\Lambda = \{(\lambda_1, \ldots, \lambda_m) \in R^m : \textstyle\sum_{i=1}^m \lambda_i \le 1, \lambda_i \in [0, 1]\}$$

Uočimo da je Λ zatvoren skup u R^m i

$$Int\Lambda = \{(\lambda_1, \ldots, \lambda_m) \in R^m : \textstyle\sum_{i=1}^m \lambda_i \leq 1, \lambda_i \in (0,1)\}$$

Koristeći afinu nezavisnost točka $x_i - x_0$, $i = 1, \ldots, m$ lako je pokazati da je f homeomorfizam(f i f^{-1} su neprekidne funkcije) Zato f " komutira" s interiorom i zatvaračem, tj.

$$RelIntf(A) = f(IntA),$$

$$Clf(A) = f(ClA)$$

za svaki $A \subseteq \mathbb{R}^n$. Specijalno, za $A = \Lambda$ dobivamo

$$RelIntK = RelIntf(\Lambda) = f(Int\Lambda)$$

□

Propozicija 2.2.2. *Ako je $K \subseteq \mathbb{R}^n$ neprazan i konveksan skup, onda je $RelIntK \neq \emptyset$ i $dim\,(RelIntK) = dimK$.*

Dokaz. Neka je $m = dimK$. Tada K sadrži $m + 1$ afino nezavisnih točaka $x_0, x_1, \ldots, x_m$ takvih da je

$$\mathrm{aff}K = \mathrm{aff}\,\{x_0, x_1, \ldots, x_m\}.$$

Skup K je konveksan pa sadrži politop

$$S := conv\,(\{x_0, x_1, \ldots, x_m\}).$$

Znamo da je aff S = aff K i $S \subseteq K$ pa dobivamo $RelIntS \subseteq RelIntK$. Prema lemi 2.2.1 je $RelIntS \neq \emptyset$ pa je zato $RelIntK \neq \emptyset$.
Prije dokaza jednakosti

$$dim\,(RelIntK) = dimK$$

prisjetimo se da se dimenzija $dimA$ nekog skupa $A \subseteq \mathbb{R}^n$ definira kao dimenzija njegove afine ljuske aff A ; dimenzija $dimA$ je za jedan manja od maksimalnog broja afino nezavisnih vektora iz skupa A. Dimenzija ima svojstvo " monotonosti", tj.

$$A \subseteq B \Rightarrow dimA \leq dimB.$$

Kako je $RelIntK \subseteq K$ monotonost dimenzije daje $dim\,(RelIntK) \leq dimK = m$. □

Lema 2.2.3. *(Princip linijskog segmenta)*
Neka je $K \subseteq \mathbb{R}^n$ konveksan skup, $x \in RelIntK$ i $x' \in ClK$. Tada je poluotvoreni segment

$$[x, x'] := \{\alpha x + (1 - \alpha)\,x' : 0 < \alpha \leq 1\}$$

sadržan u $RelIntK$.

Dokaz. Kako je $x \in RelIntK$ postoji $\epsilon > 0$ takav da je $B(x, \epsilon) \bigcap \text{aff } K \subseteq K$. Neka je

$$y = \alpha x + (1 - \alpha) x', \ 0 < \alpha < 1.$$

Kako je $x' \in ClK$ postoji niz $\left(x'_k\right)$ točaka iz K koji konvergira prema x'_k. Neka je

$$x_k = \tfrac{1}{\alpha} \left(y - (1 - \alpha) x'_k\right), \ k \in \mathbb{N},$$

tada je

$$\lim_{k \to \infty} x_k = x$$

pa možemo pretpostaviti da je k dovoljno velik tako da bude

$$B\left(x_k, \tfrac{\epsilon}{2}\right) \subset B(x, \epsilon)$$

tada dobivamo:

$$B\left(y, \tfrac{\alpha\epsilon}{2}\right) = y + B\left(0, \tfrac{\alpha\epsilon}{2}\right) = (1 - \alpha) x_k + \alpha x'_k + B\left(0, \tfrac{\alpha\epsilon}{2}\right) = \left(1 - \alpha x'_k\right) + \alpha B\left(x_k, \tfrac{\epsilon}{2}\right),$$
$$B\left(y, \tfrac{\alpha\epsilon}{2}\right) \bigcap \text{aff } K \subseteq K$$

□

Korolar 2.2.4. *Neka je $K \subseteq \mathbb{R}^n$ konveksan skup. Zraka s početkom u točki $x \in RelIntK$ može sjeći relativnu granicu $RelBdK = ClK \setminus RelIntK$ u najviše jednoj točki.*

Propozicija 2.2.5. *Ako je $K \subseteq \mathbb{R}^n$ konveksan skup,onda je*

$$ClK = \{x \in \text{aff } K : \exists y \in K, [y, x) \subseteq K\}$$

i

$$RelIntK = \{x \in \text{aff } K : \forall y \in \text{aff } K \setminus \{x\}, \exists z \in (x, y), [x, z] \subseteq K\}.$$

Dokaz. Neka je B desna strana prve jednakosti. Ako je $x \in \text{aff } K$ takav da je $[y, x) \subseteq K$, $y \neq x$, onda je $\vartheta \bigcap K \neq \emptyset$ za svaku okolinuϑ točke x, pa je zato $x \in ClK$. Time smo pokazali da je $B \subseteq ClK$.

Za dokaz obratne inkluzije uzmimo bilo koji $x \in ClK$. Prema propoziciji 2.2.2. postoji $y \in RelIntK$, a prema lemi 2.2. 3. je $[y, x) \subseteq RelIntK \subseteq K$. Time smo dokazali prvu jednakost. Za dokaz druge jednakosti, neka je C desna strana te jednakosti. Očigledno je $RelIntK \subseteq C$.

Za dokaz obratne inkluzije uzmimo bilo koji $x \in C$. Odaberimo bilo koju točku $y \in RelIntK$, $y \neq x$. Po definiciji skupa C, za točku $2x - y \in \text{aff } K$ postoji $z \in (x, 2x - y)$ takva da je $[x, z] \subseteq K$. Uočimo da je $x \in (y, z)$. Zaista, kako je $z \in (x, 2x - y)$ postoji $\lambda > 0$ takav da je $z = x + \lambda(x - y)$, odakle se dobiva

$$x = \frac{1}{1+\lambda}z + \frac{\lambda}{1+\lambda}y$$

Konačno, kako je $z \in K \subseteq ClK$ i $y \in RelIntK$ prema lemi 2.2.3. je $x \in RelIntK$. □

Korolar 2.2.6. *Ako je $K \subseteq \mathbb{R}^n$ onda su konveksni njegova relativna nutrina $RelIntK$ i zatvorač ClK.*

Dokaz. Konveksnost skupa $RelIntK$ slijedi direktno iz leme 2.2.3. □

Korolar 2.2.7. *Ako je $K \subseteq \mathbb{R}^n$ konveksan skup, onda je*

$$ClK = Cl(RelIntK),$$

$$RelIntK = RelInt(ClK).$$

Dokaz. Očito je $Cl(RelIntK) \subseteq ClK$. Neka je $x \in ClK$. Prema propoziciji 2.2.5 postoji $y \in RelIntK$, a prema lemi 2.2.3. $[y, x) \subseteq RelIntK$. Kako je $RelIntK$ konveksan skup, prema propoziciji 2.2.5. je $x \in Cl(RelIntK)$. Očigledno vrijedi $RelIntK \subseteq RelInt(ClK)$.
Neka je $x \in RelInt(ClK)$. Odaberimo $y \in RelIntK$, $y \neq x$. Neka je $\epsilon > 0$ dovoljno malen, tako da bude

$$x + \epsilon(x - y) \in RelInt(ClK).$$

Ako primjenimo propoziciju 2.2.5. na konveksan skup ClK, dobivamo točku $z \in ClK$ takvu da je $z \in (x + \epsilon(x - y), x)$. Tada je $x \in (y, z)$ pa je po lemi 2.2.3. $x \in RelIntK$ □

Teorem 2.2.8. *(Carathéodoryjev teorem za konveksne skupove) Neka je $K \subseteq \mathbb{R}^n$. Svaka točka $x \in convK$ može se prikazati kao konveksna kombinacija od najviše $dimK + 1$ afino nezavisnih točaka iz skupa K.*

Dokaz. Neka je $x \in convK$ oblika

$$x = \sum_{i=1}^{m} \lambda_i x_i,$$

pri čemu je $x_i \in K$, $\lambda_i \geq 0$, $\sum_{i=1}^{m} \lambda_i = 1$.
Ako su vektori $x_1, \ldots, x_m$ afino nezavisni, onda je $m \leq dimK + 1$ i dokaz je gotov.
Zato nadalje pretpostavimo da su $x_1, \ldots, x_m$ afino zavisni vektori. Tada postoje skalari $\mu_1, \ldots, \mu_m$ koji nisu svi jednaki nuli, takvi da je

$$\sum_{i=1}^{m} \mu_i x_i = 0,$$
$$\sum_{i=1}^{m} \mu_i = 0.$$

Bez smanjenja općenitosti, pretpostavimo da je $\mu_1 > 0$. Neka je

$$\tau = \min\left\{\frac{\lambda_i}{\mu_i} > 0\right\} = \frac{\lambda_{i_0}}{\mu_{i_0}} > 0.$$

Definirajmo nenegativne brojeve

$$\lambda_i := \lambda_i - \tau\mu_i,\ i = 1, \ldots, m.$$

Uočimo da je $\lambda_{i_0} = 0$ i zato je:

$$x = \sum_{i=1}^{m} \lambda_i x_i = \sum_{i=1}^{m} \lambda_i x_i - \tau \sum_{i=1}^{m} \mu_i x_i = \sum_{i=1}^{m} \lambda_i^{-} x_i,\ i \neq i_0$$

pri čemu je:

$$\sum_{i=1}^{m} \lambda_i^{-} x_i = \sum_{i=1}^{m} \lambda_i^{-} = \sum_{i=1}^{m} (\lambda_i - \tau\mu_i) = 1.$$

Dobili smo novi prikaz od x kao konveksne kombinacije od $m - 1$ točaka iz K. □

Korolar 2.2.9. *Neka je $K \subseteq \mathbb{R}^n$.Vrijedi:*

1. *Ako je K otvoren, onda je i $convK$ otvoren skup.*
2. *Ako je K kompaktan, onda je i $convK$ kompaktan skup.*

Dokaz. 1. Neka je $x \in convK$ oblika

$$x = \sum_{i=1}^{m} \lambda_i x_i,$$

pri čemu je $x_i \in K$, $\lambda_i \geq 0$ i $x = \sum_{i=1}^{m} \lambda_i = 1$. Zbog otvorenosti skupa K postoji $\epsilon > 0$ takav da je

$$x_i + B(0, \epsilon) \subseteq K,\ \ \forall i = 1, \ldots, m.$$

Tada je

$$B(x, \epsilon) = x + B(0, \epsilon) = \sum_{i=1}^{m} \lambda_i (x_i + B(0, \epsilon)) \subseteq convK.$$

2. Standardni n- dimenzionalni simpleks

$$\Delta = \{(\lambda_1, \ldots, \lambda_{n+1} \in R^{n+1} : \sum_{i=1}^{n+1} \lambda_i = 1, \lambda_i \geq 0, i = 1, \ldots, m\}$$

je kompaktan skup.
Kao Kartezijev produkt kompaktnih skupova, kompaktan je i skup

$$K^{n+1} \times \Delta$$

Funkcija $f : K^{n+1} \times \Delta \to R^n$ definirana formulom

$$f(x_1, \ldots, x_{n+1}, \lambda_1, \ldots, \lambda_{n+1}) = \sum_{i=1}^{n+1} \lambda_i x_i$$

je neprekidna, pa je njezina slika $f(K^{n+1} \times \Delta)$ kompaktan skup.
Očito je $f(K^{n+1} \times \Delta) \subseteq convK$. Obratna inkluzija slijedi iz Carathéodoryjev-og teorema.

□

2.3 Diferencijabilnost

Kao prvo ustanovimo da konveksna funkcija f u unutrašnjoj točki x_0 domene, u svakom smjeru v, ima (jednostranu) derivaciju:

$$f'(x_0, v) = \lim_{t\to 0+} \frac{f(x_0+tv)-f(x_0)}{t}$$

Naime, postoji ϵ takav da je $x_0 + tv \in K$ za sve $|t| \leq \epsilon$.
Funkcija $g : [0, \epsilon] \to R$

$$g(t) = \frac{f(x_0+tv)-f(x_0)}{t}$$

je ne padajuća jer za $-\epsilon < t_1 < t_2 \leq \epsilon$ nejednakost $g(t_1) \leq g(t_2)$ glasi

$$f(x_0 + t_1 v) \leq \left(1 - \frac{t_1}{t_2}\right) f(x_0) + \frac{t_1}{t_2} f(x_0 + tv)$$

a ona vrijedi zbog konveksnosti funkcije f.

Iz toga slijedi da postoji $\lim_{t\to 0+} g(t)$, koji je konačan, budući je $g(t_0) \leq g(t)$ za sve $t \in [0, \epsilon]$ i fiksiran $t_0 \in [-\epsilon, 0]$. Dakle,

$$f'(x_0, v) = \lim_{x\to 0+} g(t) = inf_{0<t\leq\epsilon} g(t).$$

Uočimo da je za svaki $x \in K$ vektor $v = x - x_0$ dopustiv smjer, pri čemu je

$$f'(x_0, x - x_0) = \lim_{x\to 0+} g(t) = inf_{0<t\leq 1} g(t) \leq g(1)$$

Teorem 2.3.1. *Neka je f diferencijabilna na K . Nejednakosti:*

1. $f(x_2) \geq f(x_1) + \langle \nabla f(x_1), x_2 - x_1 \rangle$
2. $\langle \Delta f(x_2) - \nabla f(x_1), x_2 - x_1 \rangle \geq 0$

vrijede za sve $x_1, x_2 \in K$, ako i samo ako je f konveksna na$K$.

Teorem 2.3.2. *Neka je f neprekidna na K i dva puta neprekidno diferencijabilna na $intK \neq \emptyset$. Tada je f konveksna na K ako i samo ako za sve $x \in intK$, $v \in R^n$ vrijedi*

$$\left\langle \nabla^2 f(x) v, v \right\rangle \geq 0.$$

2.4 Subdiferencijal

Prema nejednakosti

$$f(x_2) \geq f(x_1) + \langle \nabla f(x_1), x_2 - x_1 \rangle$$

za diferencijabilnu konveksnu funkciju,za fiksiran $x_0 \in K$ i svaki $x \in K$ vrijedi:

$$f(x) - f(x_0) \geq \langle \Delta f(x_0), x - x_0 \rangle.$$

Definicija 2.4.1. *Kažemo da je $s \in \mathbb{R}^n$ subgradijent konveksne funkcije $f : K \to \mathbb{R}$ ($K \subseteq \mathbb{R}^n$ konveksan skup) u točki $x_0 \in K$ ako za svaki $x \in K$ vrijedi*

$$f(x) \geq f(x_0) + s^\tau (x - x_0)\,.$$

Skup svih subgradijenata funkcije f u x_0 zovemo subdiferencijalom i označavamo s $\partial f(x_0)$.

Teorem 2.4.2. *Subdiferencijal je zatvoren i konveksan skup.*

Dokaz. Neka je

$$\partial f(x_0) \neq 0$$

Prvi dio tvrdnje slijedi iz neprekidnosti skalarnog produkta u nejednakosti iz definicije. Dalje, uzmimo $y_0, y_1 \in \partial f(x_0)$ i $\lambda \in 0, 1$. Kako za sve $x \in D(f)$ i za $i \in [0, 1]$ imamo:

$$f(x) - f(x_0) \geq \langle y_i, x - x_0 \rangle$$

te nakon množenja sa $1 - \lambda$ (za $i = 0$), a sa λ (za $i = 1$) te zbrajanja dobijemo:

$$f(x) - f(x_0) \geq \langle (1 - \lambda) y_0 + \lambda y_1, x - x_0 \rangle.$$

odakle je

$$(1 - \lambda) y_0 + \lambda y_1 \in \partial f(x_0).$$

□

Teorem 2.4.3. *Neka je f konveksna na $x_0 \in K$, $K \subseteq \mathbb{R}^n$ .Tada je*

$$y_0 \in \partial f(x_0)$$

ako i samo ako za svaki dopustivi smjer v vrijedi

$$\langle y_0, v \rangle \leq f'(x_0, v).$$

Dokaz. Neka je y_0 subgradijent i v dopustiv pravac

Tada za sve $t \in [0, t_0]$ je

$$x_0 + tv \in K,$$

i

$$f(x_0 + tv) - f(x_0) \geq \langle y_0, tv \rangle.$$

Odakle je dijeljenjem s $t > 0$ i prelaskom na limes kada t pada prema nuli. Zaključujemo :

$$f'(x_0, v) \geq \langle y_0, tv \rangle$$

□

Obratno,neka za svaki dopustiv smjer v vrijedi

$$\langle y_0, v \rangle \leq f'(x_0, v)$$

što povlači da je $f'(x_0, v)$ konačan. Neka je $x \in K$ i $v := x - x_0$ (dopustiv smjer) jer je funkcija g u prethodnoj točki neopadajuća. Dakle

$$f(x) - f(x_0) < (y_0, x - x_0)$$

Poglavlje 3

Matematičko programiranje

Problem matematičkog programiranja sastoji se u odredivanju ekstremnih vrijednosti (minimuma ili maksimuma) realne funkcije $f : D \rightarrow R$ definirane na nepraznom skupu D iz euklidskog prostora R^n. Zbog jednakosti:

$$\max_{x \in D} f(x) = -\min_{x \in D}(-f(x))$$

dovoljno je ograničiti se na problem minimizacije, koji se kratko matematički zapisuje kao:

$$\min f(x)$$

uz ograničenja

$$x \in D.$$

Problem matematičkog programiranja sastoji se u pronalaženju svih točaka globalnog minimuma, zajedno s odgovarajućom minimalnom vrijednosti funkcije f. Nažalost, problem je taj što većina numeričkih metoda za minimizaciju funkcije cilja daje samo lokalno optimalno rješenje. U primjenama je vrlo važan tzv. slučaj konveksnog programiranja, u kojem se minimizira konveksna funkcija na konveksnom skupu.

3.1 Linearna optimizacija

Teorem 3.1.1. *(Osnovni teorem linearnog programiranj) Za zadane $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ i $c \in \mathbb{R}^n$ promotrimo kanonski oblik zadaće linearnog programiranja:*

$$\max c^T x$$
$$Ax = b,\ x \geq 0.$$

Označimo li a_i i-ti stupac matrice A, onda sustav $Ax = b$, $x \geq 0$ možemo zapisati u vektorskom obliku kao

$$b = x_1 a_1 + \cdots + x_n a_n,\ x_i \geq 0.$$

Vidimo da se problem nalaženja mogućih rješenja LP problema svodi na nalaženje svih mogućih prikaza vektora b kao linearne kombinacije stupaca matrice A s nenegativnim koeficijentima. Svako rješenje sustava

$$Ax = b,\ x \geq 0$$

u prvom je redu rješenje matrične jednadžbe

$$Ax = b.$$

Definicija 3.1.2. *Neka je $x = (x_1, \ldots, x_n)^T$ rješenje sustava i*

$$Ax = b,\ x \geq 0$$

tj.

$$b = x_1 a_1 + \cdots + x_n a_n.$$

Ako su linearno nezavisni vektori a_i za koje je $x_1 \neq 0$, onda kažemo da je x bazično rješenje sustava jednadžbi

$$Ax = b$$

Za bazično rješenje koje zadovoljava i ograničenja nenegativnosti $x \geq 0$ kažemo da je bazično dopustivo rješenje skupa uvijeta $Ax = b$, $x \geq 0$.

Kako izgledaju bazično dopustiva rješenja? Pretpostavimo da je skup dopustivih rješenja neprazan, tj. da je sustav

$$Ax = b, x \geq 0.$$

konzistentan. Tada matrica A i proširena matrica $[A, b]$ imaju isti rang r. Ako je $r < m$, onda je $m - r$ jednadžbi u sustavu $Ax = b$ suvišno i one se mogu izostaviti. Bez smanjenja općenitosti, možemo pretpostaviti da je $r = m$. Ako je $r = m = n$, onda sustav $Ax = b$ ima jedinstveno rješenje $x = A^{-1}b$ pa u tom slučaju preostaje samo da se vidi da li je $x \geq 0$ ili nije. Dakle, od interesa je samo slučaj $r = m < n$. Odaberimo m linearno nezavisnih stupaca $a_1, \ldots, a_m$ matrice A i od njh napravimo matricu $B = [a_{i1}, \ldots, a_{im}]$. U teoriji linearnog programiranja matrica B se zove matrica baze ili kraće samo baza. Bez smanjenja općenitosti, pretpostavimo da su prvih $a_{i1}, \ldots, a_{im}$ stupaca matrice A linearno nezavisni. Tada, uz oznake

$$\begin{aligned} B &:= [a_1, \ldots, a_m], \\ N &:= [a_{m+1}, \ldots, a_n], \\ x_B &:= (x_1, \ldots, x_m)^T, \\ x_N &:= (x_{m+1}, \ldots, a_n), \end{aligned}$$

sustav jednadžbi

$$Ax = b,\, x \geq 0$$

možemo pisati kao

$$Bx_B + Nx_N = b.$$

Varijable $x_1, \ldots, x_m$ zovemo bazične varijable, dok su preostale varijable $x_{m+1}, \ldots, x_n$ nebazične varijable. Kako je B regularna matrica, množenjem slijeva gornje jednakosti inverzom B^{-1} i nakon toga sređivanjem, dobivamo:

$$x_B = B^{-1}(b - Nx_N).$$

Veza između bazičnih (vektor x_B) i nebazičnih (vektor x_N) varijabli govori nam da su bazične varijable na jedinstven način određene s vrijednostima nebazičnih varijabli. Zato, na bazične varijable možemo gledati kao na zavisne, a na nebazične kao na nezavisne varijable.

3.2 Dualnost

U ovom poglavlju bavit ćemo se dualnošću u linearnom programiranju i odnosima primarnog i dualnog problema.

Neka su $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ i $A \in \mathrm{M}_{m,n}(\mathbb{R})$. Zadaću linearne optimizacije

$$\begin{aligned} c^\tau x &\to \max \\ Ax &\leq b \\ x &\geq 0 \end{aligned} \tag{3.1}$$

nazivat ćemo primarnom zadaćom. Pridružujemo joj dualnu zadaću

$$\begin{aligned} b^\tau y &\to \min \\ A^\tau y &\geq c \\ y &\geq 0\,. \end{aligned} \tag{3.2}$$

Označimo dopustive skupove primarne i dualne zadaće s $\mathcal{P}$, odnosno $\mathcal{D}$:

$$\mathcal{P} = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$$
$$\mathcal{D} = \{y \in \mathbb{R}^m : A^\tau y \leq c, y \geq 0\}.$$

Teorem 3.2.1 (slaba dualnost)**.** *Ako su $x \in \mathcal{P}$ i $y \in \mathcal{D}$ onda je*

$$c^\tau x \leq b^\tau y.$$

Dokaz. Uzmimo $x \in \mathcal{P}$ i $y \in \mathcal{D}$. Vrijedi

$$c^\tau x \leq y^\tau A x \leq y^\tau b.$$

□

Dakle, prema prethodnom teoremu, vrijednost primarne funkcije cilja nije veća od vrijednosti dualne funkcije cilja.

Korolar 3.2.2. *Ako za neke $x \in \mathcal{P}$ i $y \in \mathcal{D}$ vrijedi $c^\tau x = b^\tau y$, onda je x rješenje primarne zadaće, a y dualne zadaće.*

Zadaći (3.2) također možemo pridružiti dualnu zadaću, ali je prvo moramo zapisati u formi zadaće (3.1), što činimo uobičajenim transformacijama:

$$\begin{array}{rcl} b^\tau y & \to & \min \\ A^\tau y & \geq & c \\ y & \geq & 0. \end{array} \qquad \Longleftrightarrow \qquad \begin{array}{rcl} -b^\tau y & \to & \max \\ -A^\tau y & \leq & -c \\ y & \geq & 0. \end{array}$$

pa dualna zadaća dualne zadaće glasi

$$\begin{array}{rcl} (-c)^\tau x & \to & \min \\ (-A^\tau)^\tau x & \geq & -b \\ x & \geq & 0. \end{array} \qquad \Longleftrightarrow \qquad \begin{array}{rcl} c^\tau x & \to & \max \\ Ax & \leq & b \\ x & \geq & 0, \end{array}$$

što predstavlja primarnu zadaću (3.1). Stoga govorimo jednostavno da su zadaće (3.1) i (3.2) u dualnosti (dual jedne je ona druga).

Primjer 3.2.3. *Napišite dualnu zadaću sljedećeg problema linearnog programiranja:*

$$\begin{gathered} x_1 - x_2 + x_3 \to \min \\ 3x_1 - x_2 + x_3 \geq 0 \\ 2x_1 - x_3 \geq 12 \\ x_1, x_2, x_3 \geq 0 \end{gathered}$$

Rješenje: Dualna zadaća glasi:

$$\begin{gathered} 12y_2 \to \max \\ 3y_1 + 2y_2 \leq 1 \\ -y_1 \leq -1 \\ y_1 - y_2 \leq 1 \\ y_1, y_2, y_3 \geq 0 \end{gathered}$$

Teorem 3.2.4 (jaka dualnost)**.** *Neka su $\mathcal{P}$ i $\mathcal{D}$ neprazni. Tada (3.1) i (3.2) imaju rješenja i vrijedi*

$$\max_{x\in\mathcal{P}} c^\tau x = \min_{y\in\mathcal{D}} b^\tau y\,.$$

Dokaz teorema nećemo provoditi, a može se naći u [2].

3.3 Nelinearna optimizacija

Problemi nelinearnog programiranja su znatno teži za rješavanje od problema linearnog programiranja, jer se rješenje nelinearnog problema može naći u unutrašnjoj točki ili na granici dopustivog skupa. Problemi nelinearne optimizacije često javljaju u prirodnim znanostima.

Dopustivi skup $K \subseteq \mathbb{R}^n$, problema minimizacije zadajemo na sljedeći način

$$K := \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \ldots, m\}\,,$$

gdje su $g_i : \mathbb{R}^n \to \mathbb{R}$ zadane funkcije. Elemente dopustivog skupa zovemo dopustivim točkama.

Imajući u vidu definiciju dopustivog skupa K problem minimizacije finkcije f uz uvjete $g_i(x) \leq 0$, $i = 1, 2, \ldots, m$ možemo zapisati u skraćenom obliku:

$$\min\nolimits_{x\in K} f(x).$$

U praksi, problem optimizacije smatramo rješenim ukoliko smo našli točku globalnog minimuma (maksimuma), odnosno sve točke globalnog minimuma (maksimuma), ukoliko ih ima više, kao i vrijednost funkcije u tim točkama. Međutim, u praksi je često vrlo teško naći globalne minimume, te se često zadovoljavamo nalaženjem točaka lokalnog minimuma (maksimuma).

Primjer 3.3.1. *Pretpostavimo da na raspolaganju imamo karton povšine A od koga želimo napraviti kutiju maksimalnog volumena. Pretpostavimo da se stranice kutije posebno lijepe nekom trakom, pa se karton dodatno ne troši za spojeve. Napravite matematički model koji*

odgovara ovom problemu!
Rj: Obilježimo dužinu, širinu i visinu kutije sa x, y i z. Volumen kutije dobije se pomoću formule $V = xyz$. Površinu definiramo sa $P = 2xy + 2yz + 2xz$.

Imamo model:

$$\min V = xyz$$
$$2xy + 2yz + 2xz \leq A$$
$$x, y, z > 0.$$

Primjer 3.3.2. *Ako neka tvrtka naplaćuje svoj proizvod p i ako dnevno troši a na reklamiranje tog proizvoda, očekivan broj prodanih proizvoda iznosi* $10000 - 100p$ *.Ako se za proizvodnju jednog proizvoda dnevno potroši 10eura, koliku cjenu kompanija treba zadati kako bi maksimizirala svoj profit?*
Rj: Model:

$$\max f(p) = (10000 - 100p)(p - 10) - a$$

uz ograničenja

$$p \geq 10$$

3.4 Metode optimizacije

Uvom poglavlju promatrat ćemo neke numeričke metode za probleme optimizacije. Spomenimo najprije metode kod kojih nije potrebno poznavanje derivacije minimizirajuće funkcije. To su: metoda polovljenja ili bisekcija, metoda zlatnog reza, metoda parabole, Brentova metoda te Newtonovu metodu koja pretpostavlja dovoljnu glatkoću minimizirajuće funkcije.

U višedimenzionalnoj optimizaciji koristimo metodu najbržeg spusta, gradijentnu metodu i Broydenovu metodu. Mi ćemo se zadržati na metodama jednodimenzionalne minimizacije.

Neka je zadana je funkcija $f\colon [a, b] \to \mathbb{R}$ koja općenito nije derivabilna i koja u nepoznatoj točki x^* postiže strogi lokalni minimum. Problem pronalaženja točke x^* nećemo moći riješiti primjenom metoda koje se zasnivaju na poznavanju derivacije funkcije f, zato ćemo promotriti dvije jednostavne metode za određivanje minimuma takve funkcije: metodu zlatnog reza i metodu parabole.

Definicija 3.4.1. *Za funkciju $f\colon [a, b] \to \mathbb{R}$ kažemo da je unimodalna na intervalu $[a, b]$ ako f postiže minimum u nekoj točki $x^* \in [a, b]$ i ako za svake dvije točke $x_1 < x_2$ vrijede svojstva:*

1. *ako je* $x_1 < x_2 \leq x^*$ *tada je* $f(x_1) > f(x_2)$ *tj. funkcija je padajuća na* $[a, x^*]$
2. *ako je* $x^* \leq x_1 < x_2$ *tada je* $f(x_1) < f(x_2)$ *tj. funkcija je rastuća na* $[x^*, b]$.

Metoda polovljenja ili bisekcije

Metodu nazivamo metodom polovljenja jer u svakom koraku interval dijelimo na dvije polovine. Ako je δ manji, time je dijeljenje intervala bliže raspolavljanju. Najbolja točnost koja se može postići je točnost $\delta > 0$.

Zadana je konveksna funkcija $f : [a, b] \rightarrow \mathbb{R}$. Cilj je odrediti točku

$$x^* = \text{argmin}_{x \in [a,b]} f(x) \text{ s točnošću } \delta > 0\,.$$

Označimo: $x_a = a$, $x_b = b$ te sa $x_m = \frac{1}{2}(a + b)$

1. Izračunajmo: $F_a = F(x_a)$, $F_b = F(x_b)$ te $F_m = F(x_m)$
2. Definirajmo: $x_l = \frac{1}{2}(x_a + x_m)$ i $x_r = \frac{1}{2}(x_m + x_b)$ i izračunajmo $F_l = F(x_l)$ i $F_r = F(x_r)$

   Neka je $F_{\min} = \min\{F_a, F_b, F_m, F_l, F_r\}$

   Ako je $F_{\min} = F_a$ ili $F_{\min} = F_l$ tada definiramo: $x_b := x_m$, $x_m = x_l$, $F_b = F_m$, $F_m = F_l$

   Inače, ako je $F_{\min} = F_m$ tada definiramo: $x_a = x_l$, $x_b = x_r$, $F_a = F_l$, $F_b = F_r$

   Inače (ako je $F_{\min} = F_r$ ili $F_{\min} = F_b$) onda definiramo $x_a = x_m$, $x_m = x_r$, $F_a = F_m$, $F_m = F_r$

Korak 2. ponavljamo sve dok $|x_b - x_a|$ nije manji od δ.

Metoda parabole

Promotrimo još jednu jednostavnu metodu za odredivanje točke x^* u kojoj se postiže minimum zadane funkcije $f : [a, b] \rightarrow \mathbb{R}$.

Za početak odedimo realni broj $c \in \langle a, b \rangle$ tako da je $f(a) > f(c)$ i $f(c) < f(b)$. Za točke $A = (a, f(a))$, $C = (c, f(c))$ i $B = (b, f(b))$ sa zadanim svojstvom postoji jedinstveni polinom drugog stupnja
$P_2(x) = \alpha x^2 + \beta x + \gamma$ čiji graf prolazi danim točkama.
Za koeficijente α, β i γ polinoma P_2 dobivamo:

$$\alpha = \frac{f(a)}{(a-c)(a-b)} + \frac{f(c)}{(c-a)(c-b)} + \frac{f(b)}{(b-a)(b-c)} \tag{3.3}$$

$$\beta = -\left(\frac{bf(a)+cf(a)}{(a-c)(a-b)} + \frac{af(c)+bf(c)}{(c-a)(c-b)} + \frac{af(b)+cf(b)}{(b-a)(b-c)}\right) \tag{3.4}$$

$$\gamma = \frac{cbf(a)}{(a-c)(a-b)} + \frac{abf(c)}{(c-a)(c-b)} + \frac{acf(b)}{(b-a)(b-c)} \tag{3.5}$$

Metoda zlatnog reza

Korekcija prethodno pokazane metode polovljenja sastoji se u tome da točke oko polovišta intervala biramo u smislu omjera "zlatnog reza".

Neka je $f\colon [a,b] \to \mathbb{R}$ zadana funkcija te neka je $[a_{k-1}, b_{k-1}] \subset [a,b]$ interval takav da je $f\colon [a_{k-1}, b_{k-1}] \to \mathbb{R}$ unimodalna na $[a_{k-1}, b_{k-1}]$. Za početak promatrajmo nepoznatu točku $x^* \in [a_{k-1}, b_{k-1}]$ kojoj se postiže lokalni minimum funkcije f na segmentu $[a_{k-1}, b_{k-1}]$. Vrijednost funkcije izračunat ćemo u dvije točke $y_k < z_k$ gdje su $y_k, z_k \in [a_{k-1}, b_{k-1}]$ za koje vrijedi:

1. y_k i z_k jednako su udaljeni od krajeva segmenta $[a_{k-1}, b_{k-1}]$ tj. vrijedi
   $y_k - a_{k-1} = b_{k-1} - z_k$

2. y_k je bliže lijevom rubu segmenta $[a_{k-1}, b_{k-1}]$, a z_k je bliže desnom rubu segmenta $[a_{k-1}, b_{k-1}]$ tj. za neki $c \in \left(0, \frac{1}{2}\right)$ vrijedi
   $y_k = a_{k-1} + c\,(b_{k-1} - a_{k-1})$,
   $z_k = b_{k-1} - c\,(b_{k-1} - a_{k-1})$
   $= a_{k-1} + (1-c)\,(b_{k-1} - a_{k-1})$
   $= a_{k-1} + b_{k-1} - y_k$

3. Mogu pri tome nastupiti dva slučaja:

   a) Ako je $f(y_k) \leq f(z_k)$ zbog unimodalnosti funkcije f je $x^* \in [a_{k-1}, z_k]$ stavljamo
      $a_k{:=}a_{k-1}$
      $b_k{:=}z_k$
      $x_k{:=}y_k$.
      Tada je x_k nova aproksimacija točke x^*, koja je bolja od prethodne x_{k-1}.

   b) Ako je $f(y_k) > f(z_k)$ zbog unimodalnosti funkcije f je $x^* \in [y_k, b_{k-1}]$, pa stavljamo

a_k:=y_k
b_k:=b_{k-1}
x_k:=z_k.
Došli smo do segmenta $[a_k, b_k]$ za koji očito vrijedi $[a_k, b_k] \subset [a_{k-1}, b_{k-1}]$.

Kako vrijednost funkcije ne bismo u svakom intervalu morali računati u dvije točke proširit ćemo 3) i to na sljedećii način.
U slučaju a) kada je $f(y_k) \leq f(z_k)$ stavljamo još
z_{k+1}:=y_k.
Ovdje smo pretpostavili još jedan uvjet, a to je da točka y_k leži bliže točki z_k nego točka a_{k-1}, što je zadovoljeno dobrim izborom broja c.
U slučaju b) kada je $f(y_k) > f(z_k)$ stavljamo još
y_{k+1}:=z_k.
Ovdje smo pretpostavili da točka z_k leži bliže točki y_k nego točka b_{k-1}. Sređivanjem relacija u prvom i drugom slučaju, te uspređujući ih dobivamo:
$c = \frac{3-\sqrt{5}}{2}$,
a taj c je u smislu omjera „*zlatnog reza*„.

Brentova metoda se zasniva na kombinaciji metode zlatnog reza i metode parabole. Kompromis između dviju navedenih metoda sastoji se utome da se obje koriste ovisno o tome koja se metoda u danom koraku učini pogodnijom. Prilikom optimizacijskog procesa pratimo 6 točaka: a, b, x, u ,v i w. Minimum se nalazi unutar intervala (a,b). x je posljednja točka gdje je vrijednost funkcije bila najmanja, a w je sljedeća točka iza x koja odgovara drugoj po veličini funkcijskoj vrijednosti. v je prethodna vrijednost od w, u je točka u kojoj je vrijednost funkcije izračunata zadnji put; obično se u algoritmu prati i točka koja je u sredini intervala $\langle a, b\rangle$, ali se u njoj vrijednost funkcije ne mora računati.

Newtonova metoda

Newtonova metoda jedna od najčešće korištenih metoda za određivanje lokalnog minimuma ili maksimuma dva puta derivabilne funkcije.

Pretpostavimo da je $f \in C^2(a, b)$ i da u točki $\beta \in \langle a, b\rangle$postiže jedinstveni lokalni minimum.

Neka x_0 aproksimira minimum finkcije f. Ponovimo za $k = 0, 1, 2, \ldots$

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} \tag{3.6}$$

sve dok $|f'(x_{k+1}|$ ne postane jako mali.

Ovaj algoritam je dobiven razvojem $f(x)$ u Taylerov red po x_k:

$$f(x_k + h) = f(x_k) + hf'(x_k) + \frac{h^2}{2} f''(x_k) + O(h^3)\,,$$

odnosno za f':

$$f'(x_k + h) = f'(x_k) + hf''(x_k) + O(h^2)$$

Pretpostavimo da je h takav da je $f'(x_k + h) = 0$. Ako zanemarimo član $O(h^2)$ onda iz prethodne jednakosti slijedi

$$h = -\frac{f'(x_k)}{f''(x_k)}$$

što smo koristili u (3.6).

Neka je $S = [a, b]$ zatvoreni interval realnih brojeva i $\phi : S \to S$. Točku $x \in S$ takvu da je $\phi(x) = x$ zovemo čvrsta točka od ϕ. Preslikavanje $\phi : S \to S$ nazivamo kontrakcija ako postoji pozitivni broj $q < 1$ takav da za svake dvije točke x^1 i x^2 iz S vrijedi

$$\left|\phi(x^1) - \phi(x^2)\right| \leq q \left|x^1 - x^2\right|. \tag{3.7}$$

Neka je $\phi : S \to S$ kontrakcija na $S = [a, b] \subset \mathbb{R}$ i neka je $x^0 \in S$, $x^{k+1} = \phi(x^k)$ za $k \in \mathbb{N}_0$. Tada postoji jedinstvena čvrsta točka x_* preslikavanja ϕ te niz $\left\{x^k\right\}$ konvergira u x_* i

$$\left|x^{k+1} - x_*\right| \leq (q)^{k+1} \left|x^0 - x_*\right|, k = 0, 1, \ldots \tag{3.8}$$

Dokaz. Kako je ϕ kontrakcija to slijedi

$$|x_{k+1} - x_k| \leq q\,|x_k - x_{k-1}| \leq \cdots \leq q^k\,|x_1 - x_0|\,, \tag{3.9}$$

odnosno za $m > n$ imamo

$$|x_m - x_n| \leq |x_m - x_{m-1}| + \cdots + |x_{n+1} - x_n| \leq \left(q^{m-1} + q^{m-2} + \cdots + q^n\right)|x_1, -x_0| = \frac{q^n}{1-q}\,|x_1 - x_0|\,. \tag{3.10}$$

Stoga je (x_n) Cauchyjev niz, odnosno konvergira prema $x^* \in S$. Zbog neprekidnosti funkcije ϕ je $x_{n+1} = \phi(x_n) \to \phi(x^*)$ pa je $\phi(x^*) = x^*$, zbog jedinstvenosti limesa niza (x_n). Dakle,

$$\left|x^{k+1} - x_*\right| \leq (q)^{k+1} \left|x^0 - x_*\right|, k = 0, 1, \ldots \tag{3.11}$$

sada lako slijedi.Na kraju trebamo jos dokazat da je x^* jedinstven. Pretpostavimo suprotno:neka sux^* i y^* dvije čvrste točke.
Tada je:

$$|x^* - y^*| = |\phi(x^*) - \phi(y^*)| \leq q\,|x^* - y^*| < |x^* - y^*| \tag{3.12}$$

□

Pretpostavit ćemo da je promatrana funkcija f:$R^n \to R^n$ dvostruko neprekidno diferencijabilna i da za neke $M > m > 0$ vrijedi:

$$m\,||y||^2 \leq (f''(x)y, y) \leq M\,||y||^2\,, x, y \in \mathbb{R}^n \tag{3.13}$$

Posebno, f je strogo konveksna.

Teorem 3.4.2. *Neka funkcija $f \in C^2(R^n)$ zadovoljava uvjet*

$$m\,||y||^2 \leq (f'(x)y, y) \leq M\,||y||^2\,, x, y \in \mathbb{R}^n \tag{3.14}$$

za neke $M > m > 0$
i neka se parametri duljine koraka α_k biraju iz uvjeta

$$f(x) - f(x_k) \leq \alpha f'(x_k)p_k, p_k = -\frac{f'(x_k)}{f''(x_k)} \tag{3.15}$$

Tada neovisno o izboru početne aproksimacije x_0, iterativni proces

$$x_{k+1} = x_k + \alpha_k p_k, k = 0, 1, ..., \tag{3.16}$$

konvergira prema jedinstvenoj točki minimuma $x*$.

$$||x_{k+1} - x^*|| \leq \alpha_k\,||x_k - x^*|| \tag{3.17}$$

gdje
$\alpha_k \to 0$ *za* $k \to \infty$.

Bibliografija

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

[2] M. Florenzano, C. L. Van, *Finite dimensional convexity and optimization*, Springer, 2001.

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

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

Sažetak

U ovom diplomskom radu u prvom poglavlju smo definirali konveksni i skup te promatrali osnovna geometrijska svojstva konveknih skupova te konveksne funkcije. U drugom poglavlju definirali smo konveksnu ljusku te dokazali Carathéodoryjev teorem. U zadnjem poglavlju poglavlju diplomskog rada proučavali smo linearnu i nelinearnu optimizaciju. Opisali smo osnovne metode jednodimenzionalne minimizacije. Na kraju poglavlja smo obradili dualnost u linearnoj optimizaciji.

Summary

In this thesis in the first chapter, we define convex sets and we introduce basic geometric properties of convex sets and convex functions. In the second chapter we define the convex hull and prove Carathéodory theorem. In the last chapter we study linear and nonlinear optimization. We describe one-dimensional methods of minimization. On the end of the chapter, we study some duality results in linear optimization.

Životopis

Rođena sam 26.08.1984.g u Novoj Gradišci kao prvo dijete Dubravke i Braminira Piškor. U prvi razred krećem 1991.g u Osnovnu školu Rešetari jer smo u Rešetarima boravili kao prognanici. Nakon Bljeska moja obitelj se vraća kući i tada u peti razred krećem u Osnovnu školu Okučani. Nakon završene osnovne škole, 1998. upisujem Opću gimnaziju Nova Gradiška koju uspjesno završavam. Zatim upisujem integrirani preddiplomski i diplomski sveučilišni studij matematike i fizike na matematičkom odjelu Prirodoslovno-matematičkog fakulteta u Zagrebu.