Konveksne množice

Vesna Rebselj

20. oktober 2009

Naj bo $\mathcal{D} \subseteq \mathbb{R}^n$. S $\mathrm{cl}(\mathcal{D})$ označimo *zaprtje*, z $\mathrm{int}(\mathcal{D})$ pa *notranjost* množice $\mathcal{D}$ v $\mathbb{R}^n$.

Definicija 1 *Množica $\mathcal{D} \subseteq \mathbb{R}^n$ je* konveksna, *če za vsaka $x, y \in \mathcal{D}$ in vsak $\lambda \in [0,1]$ velja $\lambda x + (1-\lambda)y \in \mathcal{D}$.*

Vsakemu vektorju $\sum_{i=1}^m \lambda_i x_i$, kjer je $m \in \mathbb{N}$, $\lambda_i \geq 0$ in $\sum_{i=1}^m \lambda_i = 1$, pravimo *konveksna kombinacija* vektorjev $x_i \in \mathbb{R}^n$. Vsaka konveksna množica vsebuje vse konveksne kombinacije svojih elementov.

Presek poljubne družine konveksnih množic je spet konveksna množica (prazna množica je po definiciji konveksna; res pa je, da pri trditvah največkrat tiho predpostavimo, da so konveksne množice neprazne).

Konveksna ovojnica (ali *konveksna ogrinjača*) $\mathrm{conv}(\mathcal{D})$ množice $\mathcal{D}$ je najmanjša konveksna množica, ki vsebuje $\mathcal{D}$.

Trditev 2 *Konveksna ovojnica množice $\mathcal{D}$ je presek vseh konveksnih množic, ki vsebujejo $\mathcal{D}$, oziroma množica vseh konveksnih kombinacij točk iz $\mathcal{D}$.*

Dokaz. Oglejmo si množico

$$S := \left\{ \sum_{i=1}^m \lambda_i x_i;\ m \in \mathbb{N},\ x_i \in \mathcal{D},\ \lambda_i \geq 0,\ \sum_{i=1}^m \lambda_i = 1 \right\}.$$

Hitro preverimo, da je množica S konveksna: Če si izberemo $x, y \in S$, potem sta x in y konveksna kombinacija elementov iz $\mathcal{D}$, torej $x = \sum_{i=1}^m \lambda_i x_i$ in $y = \sum_{i=1}^n \mu_i x_i$ in $m, n \in \mathbb{N}$, $\lambda_i, \mu_i \geq 0$ in $\sum_{i=1}^m \lambda_i = \sum_{i=1}^n \mu_i = 1$. Potem pa je za $\lambda \in [0,1]$ vsota $\lambda x + (1-\lambda)y = \sum_{i=1}^m\ \lambda l a_i + \sum_{i=1}^n \lambda \mu_i$ tudi v S,saj velja $\lambda\lambda_i, (1-\lambda)\mu_i \in [0,1]$ za vse λ_i, μ_i in $\lambda\lambda_i + (1-\lambda)\mu_j \in [0,1]$ za vse λ, λ_i, $\mu_j \in [0,1]$, kar ravno ustreza definiciji konveksnosti množice S. Ker vsak x iz $\mathcal{D}$ lahko zapišemo v obliki $x = \sum_{i=1}^1 1x$, velja $\mathcal{D} \subseteq S$. Ker je vsak x iz S konveksna kombinacija elementov iz $\mathcal{D}$ in ker vsaka konveksna množica (torej tudi $conv(\mathcal{D})$) vsebuje vse konveksne kombinacije svojih elementov, je $S \subseteq conv(\mathcal{D})$. Torej je S konveksna in velja $\mathcal{D} \subseteq S \subseteq conv(\mathcal{D})$. Potem mora biti $S = conv(\mathcal{D})$. Jasno je tudi $conv(\mathcal{D})$ res presek vseh konveksnih množic, ki vsebujejo $\mathcal{D}$. $\square$

Množica je konveksna natanko tedaj, ko je enaka svoji konveksni ovojnici.

TRDITEV 3 *Za kompaktne (tj. zaprte in omejene v $\mathbb{R}^d$) množice je tudi njihova konveksna ovojnica kompaktna. Za odprte množice je njihova konveksna ovojnica odprta množica.*

Dokaz. Ker je K omejena, konveksna ovojnica K pa konveksna kombinacija elementov iz K, je tudi $conv(K)$ omejena. Za prvi del trditve nam preostane dokazati še, da je $conv(K)$ zaprta.
Vemo, da je množica K v metričnem prostoru zaprta natanko tedaj, ko ima vsako konvergento zaporedje iz K limito v K. Oglejmo si zdaj poljubno konvergentno zaporedje $\{x^n\}_{n\in\mathbb{N}}$. Naj bo x limita tega zaporedja. Radi bi pokazali, da je x iz konveksne ovojnice K. Vsak člen zaporedja $\{x^n\}_{n\in\mathbb{N}}$ lahko zapišemo v obliki $x^n = \sum_{i=1}^{d+1} \lambda_i^n x_i^n$, pri čemer je $\sum_{i=1}^{d+1} \lambda_i^n = 1$, in vsi x_i^n so iz konveksne ogrinjače K.

$$x_1 = \sum_{i=1}^{d+1} \lambda_i^1 x_i^1$$
$$x_2 = \sum_{i=1}^{d+1} \lambda_i^2 x_i^2$$
$$x_3 = \sum_{i=1}^{d+1} \lambda_i^3 x_i^3$$
$$\vdots$$

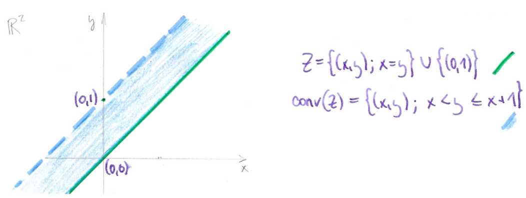
Če to zaporedje gledamo po komponentah, dobimo v j-ti komponenti vsakega vektorja iz zaporedja člen $\lambda_j^n x_j^n$. Ker so x_j^n vsi iz kompaktne množice K, $\{x_j^n\}_{n\in\mathbb{N}}$ pa konvergentno zaporedje, mora biti tudi njegova limita, x_j v K. Prav tako je za vsako komponento kovergentno tudi zaporedje koeficientov $\{\lambda_j^n\} n \in \mathbb{N}$. In ker je tudi to zaporedje na kompaktni množici $[0, 1]$, je tudi njegova limita $\lambda_j \in [0, 1]$. Zdaj pa lahko x zapišemo kot konveksno kombinacijo elementov iz K:

$$x = \lim_{n\to\infty} \sum_{j=1}^{d+1} \lambda_j^n x_j^n = \sum_{j=1}^{d+1} \lim_{n\to\infty} \lambda_j^n x_j^n = \sum_{j=1}^{d+1} \lambda_j x_j.$$

S tem je trditev za kompaktne množice dokazana.
Naj bo zdaj $U \subseteq \mathbb{R}^n$ odprta. Izberimo poljuben x iz konveksne ogrinjače U. Dovolj bo dokazati, da obstaja neka odprta krogla okoli x, ki je cela vsebovana v konveksni ogrinjači U. Če je $x = 0$, to pomeni, da je 0 v U, zato si lahko mislimo, da je $x \neq 0$. Tak x lahko zapišemo kot konveksno kombinacijo elementov iz U, torej $x = \sum_{i=1}^{n} \lambda_i u_i$; $n \in \mathbb{N}$, $u_i \in U$, $\sum_{i=1}^{n} \lambda_i = 1$, in ker je x

različen od 0, lahko brez škode za splošnost predpostavimo, da je $\lambda_i > 0$ za vsak $i \in \{1, 2, \ldots, n\}$. Ker je U odprta, za vsak $u_i \in U$ lahko najdemo kroglo, ki je še cela vsebovana v U. Torej takšne krogle obstajajo tudi za vse u_i, ki nastopajo v zapisu x-a. Vzemimo zdaj za δ polmer najmanjše izmed teh krogel oziroma $\delta = 1$, če ta ne obstaja. Naj bo $\epsilon = min\{\lambda_i;\ i \in \{1, 2, ..., n\}\}$. Vzemimo zdaj $r = \delta\epsilon$. Ker sta oba, δ in ϵ strogo pozitivna, je takšen tudi njun produkt, krogla s polmerom r in s središčem v x pa je cela vsebovana v $conv(U)$. $\square$

Za zaprte množice taka trditev ne velja. Če denimo v $\mathbb{R}^2$ vzamemo zaprto množico $Z = \{(x, y);\ x = y\} \cup \{(0, 1)\}$, je $conv(Z) = \{(x, y);\ x \leq y < x + 1\}$, ki pa ni zaprta.
Skica 1:



Trditev 4 *Če je množica $\mathcal{D}$ konveksna, sta konveksni tudi množici* $\mathrm{cl}(\mathcal{D})$ *in* $\mathrm{int}(\mathcal{D})$ *(seveda je notranjost neprazne množice lahko prazna).*

Trditev 5 *Naj bo $\mathcal{D}$ konveksna množica z $\mathrm{int}(\mathcal{D}) \neq \emptyset$ ter $x_1 \in \mathrm{int}(\mathcal{D})$ in $x_2 \in \mathrm{cl}(\mathcal{D})$. Potem polodprta daljica $[x_1, x_2)$ cela leži v $\mathrm{int}(\mathcal{D})$.*

Dokaz. Lahko si mislimo, da se vse dogaja v ravnini $\mathbb{R}^2$. Ker je x_1 iz notranjosti $\mathcal{D}$, bo dovolj, če pokažemo, da je tam tudi preostanek daljice $[x_1, x_2)$, to je daljica (x_1, x_2). Vzemimo zdaj poljuben $x \in (x_1, x_2)$. Tak x lahko zapišemo kot konveksno kombinacijo x_1 in x_2: $x = \lambda x_1 + (1 - \lambda)x_2$. Ker $\lambda \notin \{0, 1\}$, lahko najdemo λ', x_2', da bo $x = \lambda' x_1 + (1 - \lambda')x_2'$, pri čemer velja $\lambda' > \lambda$ in $\lambda' \in [0, 1]$). Ker je $\mathcal{D}$ konveksna podmnožica $\mathbb{R}^2$, mora biti $x_2' \in \mathrm{int}(\mathcal{D})$. Notranjost $\mathcal{D}$ je po prejšnji trditvi konveksna, x pa je konveksna kombinacija vektorjev iz $\mathrm{int}(\mathcal{D})$, zato je tudi $x \in \mathrm{int}(\mathcal{D})$, torej je res $x \in \mathcal{D}$.
$\square$

Posledica 6 *Naj bo $\mathcal{D}$ konveksna množica z $\mathrm{int}(\mathcal{D}) \neq \emptyset$. Potem je*

$$\mathrm{cl}(\mathrm{int}(\mathcal{D})) = \mathrm{cl}(\mathcal{D}) \qquad \textit{in} \qquad \mathrm{int}(\mathcal{D}) = \mathrm{int}(\mathrm{cl}(\mathcal{D}))\,.$$

Če denimo v $\mathbb{R}$ vzamemo $\mathcal{D} = (0, 2) \cup \{4\}$ in za x_1 izberemo 1, za x_2 pa 4, potem $\mathcal{D}$ zadošča pogojem iz trditve 5, daljica $[1, 4)$ pa ne leži v celoti niti v $\mathcal{D}$, kaj šele v $\mathrm{int}(\mathcal{D})$. Poglejmo si zdaj, kako je v tem primeru s posledico 6:

$$\mathrm{cl}(\mathrm{int}(\mathcal{D})) = [0, 2] \neq [0, 2] \cup \{4\} = \mathrm{cl}(\mathcal{D})$$

Če vzamemo še nekonveksno množico $\mathcal{D}' = [0, 1] \cap \mathbb{Q}$ in to vstavimo v drugo formulo iz posledice, dobimo:

$$\mathrm{int}(\mathcal{D}') = \emptyset \neq (0, 1) = \mathrm{int}(\mathrm{cl}(\mathcal{D}')).$$

Vidimo, da za nekonveksne množice v splošnem ne veljata niti trditev 5 niti posledica 6.

Izrek 7 **(Carathéodoryjev izrek, 1911)** *Naj bo $P \subseteq \mathbb{R}^d$ in $x \in \mathrm{conv}(P)$. Potem lahko x zapišemo kot konveksno kombinacijo $d+1$ (ne nujno različnih) točk iz P.*

Dokaz. Naj bo x iz konveksne ovojnice P. Potem ga lahko zapišemo kot kombinacijo končno mnogo točk iz P: $x = \sum_{i=1}^n \lambda_i x_i$, pri čemer so vsi x_i iz P, vsi λ_i pozitivni in $\sum_{i=1}^n \lambda_i = 1$. Vzemimo zdaj $n > d + 1$ (sicer ni kaj dokazati). Potem so točke $x_2 - x_1$, $x_3 - x_1$, ..., $x_n - x_1$ linearno odvisne, torej obstajajo skalarji α_2, ..., α_n, ki niso vsi ničelni, da je $\sum_{i=2}^n \alpha_i (x_i - x_1) = 0$. Če definiramo $\alpha_1 := -\sum_{i=2}^n \alpha_i$, potem je $\sum_{i=1}^n \alpha_i x_i = 0$, $\sum_{i=1}^n \alpha_i = 0$ in ker niso vsi α_i ničelni, je vsaj en, imenujmo ga α_j, večji od nič. Velja

$$\sum_{i=1}^n \lambda_i x_i - \mu \sum_{i=1}^n \alpha_i x_i = \sum_{i=1}^n (\lambda_i - \mu \alpha_i) x_i$$

za vsak realen μ. Enakost bo torej držala tudi za

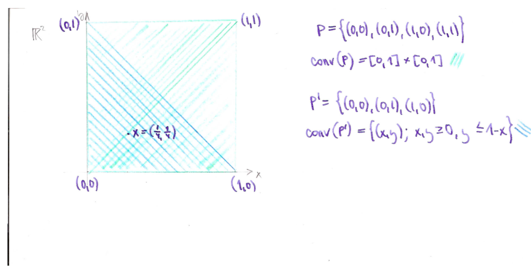
$$\mu := \min\{\frac{\lambda_i}{\alpha_i}; \alpha_i > 0, 1 \leq i \leq n\} = \frac{\lambda_j}{\alpha_j}.$$

Opazimo, da je $\mu > 0$, in da za vsak i med 1 in n velja $\lambda_i - \mu\alpha_i \geq 0$. V posebnem iz definicije μ sledi $\lambda_j - \mu\alpha_j = 0$. Torej $x = \sum_{i=1}^n (\lambda_i - \mu\alpha_i) x_i$, pri čemer so vsi $\lambda_i - \mu\alpha_i$ nenegativni, njihov seštevek je 1 in $\lambda_j - \mu\alpha_j = 0$. Drugače povedano: x je predstavljen s konveksno kombinacijo največ $n - 1$ točk iz P.
Ta postopek lahko ponavljamo, dokler x ni zapisan kot konveksna kombinacija največ $d + 1$ točk iz P. □

Vzemimo zdaj podmnožico prostora $\mathbb{R}^2$, $P = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$.

Konveksna ogrinjača te množice je pravokotnik $[0,1] \times [0,1]$. Če zdaj vzamemo točko $x = (\frac{1}{4}, \frac{1}{4})$, ki je vsebovana v konveksni ogrinjači P-ja, lahko konstruiramo množico $P' = \{(0,0),(0,1),(1,0)\}$. Konveksna ogrinjača P' je trikotnik $\{(x,y);\, x,y \geq 0,\, y \leq 1-x\}$, točka x pa njegova notranja točka. Tu torej izrek deluje, saj je moč množice P' tri.
V tem primeru ne moremo v P' imeti le dveh točk, saj je konveksna ogrinjača poljubnih dveh le daljica med njima, x pa ne leži na nobeni takšni daljici.

Skica 2:



Definicija 8 *Naj bo $\mathcal{D}$ konveksna množica. Točka $x \in \mathcal{D}$ je* ekstremna točka *množice $\mathcal{D}$, če je množica $\mathcal{D} \setminus \{x\}$ konveksna.*

Drugače povedano, točka je ekstremna natanko tedaj, ko je ni moč zapisati v obliki $\lambda x_1 + (1-\lambda)x_2$, kjer $x_1, x_2 \in \mathcal{D}$, $x_1 \neq x_2$ in $0 < \lambda < 1$. V splošnem je lahko konveksna množica brez ekstremnih točk. To se lahko zgodi tudi pri zaprtih konveksnih množicah. Velja pa, da ima vsaka kompaktna konveksna množica kakšno ekstremno točko. Nobena ekstremna točka ne leži v notranjosti množice $\mathcal{D}$.

Izrek 9 *Naj bo $\mathcal{D} \subseteq \mathbb{R}^n$ kompaktna konveksna množica in $\mathcal{E}$ množica njenih ekstremnih točk. Potem je*

$$\mathcal{D} = \mathrm{cl}(\mathrm{conv}(\mathcal{E}))\,.$$

Izrek 9 velja tudi za tako imenovane lokalno konveksne topološke vektorske prostore in ne le za $\mathbb{R}^n$. V tej splošni obliki mu pravimo **Krein-Milmanov izrek**.

V $\mathbb{R}^n$ je moč izrek 9 še izboljšati, in sicer velja

$$\mathcal{D} = \mathrm{conv}(\mathcal{E})\,.$$

Na konveksno množico $\mathcal{D} \subseteq \mathbb{R}^n$ pogosto ne gledamo kot na množico v $\mathbb{R}^n$, ampak raje kot na podmnožico njene afine ovojnice. Ta ovojnica je v bistvu spet prostor $\mathbb{R}^d$, kjer je $d \leq n$. Število d je tedaj *razsežnost* (oz. *dimenzija*) množice $\mathcal{D}$. Znotraj svoje afine ovojnice ima vsaka konveksna množica neprazno notranjost (pravimo ji *relativna notranjost* in jo označimo z $\mathrm{ri}(\mathcal{D})$). Tako lahko posplošimo mnoge trditve, pri katerih smo zahtevali neprazno notranjost množice, pri čemer $\mathrm{int}(\mathcal{D})$ zamenjamo z $\mathrm{ri}(\mathcal{D})$. Primera sta trditev 5 in posledica 6.

DEFINICIJA 10 *Naj bosta $\mathcal{D}, \mathcal{D}' \subseteq \mathbb{R}^n$. Hiperravnina* Π loči *množici $\mathcal{D}$ in $\mathcal{D}'$, če sta $\mathcal{D}$ in $\mathcal{D}'$ podmnožici različnih odprtih polprostorov, ki ju določa Π.*

Hiperravnina Π krepko loči *$\mathcal{D}$ in $\mathcal{D}'$, če loči $\mathcal{D}$ in $\mathcal{D}'$, hkrati pa obstaja tak $\varepsilon > 0$, da je vsaka točka iz $\mathcal{D} \cup \mathcal{D}'$ oddaljena vsaj ε od Π.*

Hiperravnina Π šibko loči *$\mathcal{D}$ in $\mathcal{D}'$, če sta $\mathcal{D}$ in $\mathcal{D}'$ podmnožici različnih zaprtih polprostorov, ki ju določa Π, in velja $\mathcal{D} \cup \mathcal{D}' \not\subseteq \Pi$.*

TRDITEV 11 *Za vsako zaprto konveksno množico $\mathcal{D}$ in točko $x \notin \mathcal{D}$ obstaja hiperravnina, ki krepko loči x in $\mathcal{D}$. V $\mathcal{D}$ tudi obstaja natanko ena točka, ki je najbližja x.*

Dokaz. Točka, ki je najbližja x očitno obstaja: Vemo, da je v metričnem prostoru razdalja med zaprto množico in točko, ki ni iz te množice, strogo pozitivna. Recimo torej $d(\mathcal{D}, x) = \rho > 0$. Če definiramo $C := \mathcal{D} \cap K(0, 2\rho)$, potem je funkcija $f : C \to \mathbb{R}$, definirana s predpisom $f(u) = d(u, x)$, zveznas in slika v $[\rho, 2\rho]$, kar je kompakten prostor, zato f na njem zavzame minimum. Naj bo v zdaj točka, v kateri je f zavzame minimum, torej $f(v) = \rho$. Premislimo zdaj, da je takšna točka res ena sama. Pa recimo, da sta dve. Torej imamo točki $u, v \in \mathcal{D}$ taki, da je $d(u, x) = d(v, x) = d(\mathcal{D}, x) = \rho$. Ker je $\mathcal{D}$ konveksna, velja $\frac{u+v}{2} \in \mathcal{D}$. Trikotniška neenakost nam pove, da je $d(\frac{u+v}{2}, x) < \rho$, kar pa je v protislovju s predpostavko $d(\mathcal{D}, x) = \rho$.

Hiperravnina Π, ki krepko loči x in $\mathcal{D}$ je zdaj ravnina, ki gre skozi razpolovišče daljice $[v, x]$, njena normala pa je vzporedna s to daljico. Res. Množici $\mathcal{D}$ in $\{x\}$ sta potemtakem očitno v različnih odprtih podprostorih, ki ju Π določa in vsaka točka iz $\mathcal{D} \cup \{x\}$ je od Π oddaljena vsaj za $\frac{\rho}{2}$. □

Pri konveksnih množicah, ki niso zaprte, se lahko zgodi, da najbližja točka v $\mathcal{D}$ ne obstaja, nikoli pa ne more biti več kot ena.

IZREK 12 *Naj bosta $\mathcal{D}$ in $\mathcal{D}'$ taki konveksni množici v $\mathbb{R}^n$, da je $\mathrm{int}(\mathcal{D}) \neq \emptyset$ in $\mathrm{int}(\mathcal{D}) \cap \mathcal{D}' = \emptyset$. Potem obstaja hiperravnina, ki šibko loči $\mathrm{cl}(\mathcal{D})$ in $\mathrm{cl}(\mathcal{D}')$.*

Izrek 12 velja tudi v neskončno razsežnih prostorih, v $\mathbb{R}^n$ pa ga je moč izboljšati. Opomba: v $\mathbb{R}^n$ sta enakovredni trditvi, da lahko šibko ločimo množici in da lahko šibko ločimo njuni zaprtji.

IZREK 13 *Naj bosta $\mathcal{D}$ in $\mathcal{D}'$ konveksni množici v $\mathbb{R}^n$. Množici $\mathcal{D}$ in $\mathcal{D}'$ lahko šibko ločimo natanko tedaj, ko je $\mathrm{ri}(\mathcal{D}) \cap \mathrm{ri}(\mathcal{D}') = \emptyset$.*

V izrekih 12 in 13 ni moč doseči ločitve množic $\mathcal{D}$ in $\mathcal{D}'$.

Izrek 14 *Naj bosta $\mathcal{D}$ in $\mathcal{D}'$ taki zaprti konveksni množici v $\mathbb{R}^n$, da je $\mathcal{D}$ kompaktna in $\mathcal{D} \cap \mathcal{D}' = \emptyset$. Potem obstaja hiperravnina, ki krepko loči $\mathcal{D}$ in $\mathcal{D}'$.*

Če v izreku 14 izpustimo zahtevo po kompaktnosti množice $\mathcal{D}$, izrek ni več resničen.