Komentář k úlohám probíraným na cvičeních dle plánu 11

Petr Olšák

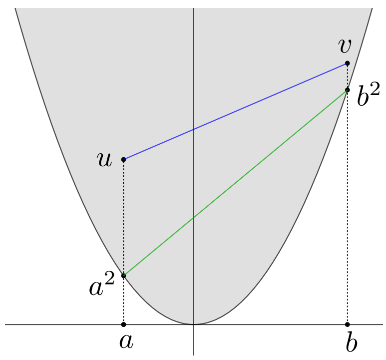
13.1 a) Množina M je konvexní, když $\forall u, v \in M$, $\forall t \in \langle 0,1\rangle$ leží $t\,u + (1-t)\,v$ také v množině M. V případě $M = \langle a, b\rangle$ máme v předpokladech definice $a \le u \le b$, $a \le v \le b$. Platí:

$$a = t\,a + (1-t)\,a \le t\,u + (1-t)\,v \le t\,b + (1-t)\,b = b$$

takže prvek $t\,u + (1-t)\,v$ je vždy mezi a a b, tedy leží v M. Množina je konvexní.

13.1 b) Množina není konvexní, Načrtněte si parabolu a na ní nějaké dva body, které spojíte úsečkou. V případě konvexní množiny musí celá úsečka v množině ležet, v tomto případě v té parabole úsečka neleží. Nyní to formalizujeme podle definice. Stačí najít dva prvky v M a jeden prvek na spojnici těch dvou, který v M neleží. Takže třeba $[1,1], [-1,1] \in M$, $t = \frac{1}{2} \in \langle 0,1\rangle$. Bod $\frac{1}{2}[1,1] + \frac{1}{2}[-1,1] = [0,1] \notin M$.

13.1 c) Množina je konvexní, jak plyne z náčrtku. Vezměte si libovolné dva body nad parabolou a shledáte, že jejich spojnice (modrá úsečka mezi nimi) leží také celá nad parabolou. Dokázat to exaktně z definice bez využití derivací dá bohužel hodně zabrat, takže je otázka, zda má smysl se tím vůbec zabývat, protože při použití derivací máme výsledek rovnou. Ale v zadání je požadavek na přímé odvození z definice, tak tedy do toho. Začneme se dvěma body ležícími přímo na parabole $[a, a^2]$ a $[b, b^2]$. Jakýkoli bod na zelené spojnici $t\,[a, a^2] + (1-t)\,[b, b^2]$ má mít svou y souřadnici větší než kvadrát jeho x-ové souřadnice. Má-li se dokázat, že je první číslo větší než druhé, stačí ověřit, že rozdíl „první číslo mínus druhé“ je větší než nula:



$$t\,a^2 + (1-t)\,b^2 - (t\,a + (1-t)\,b)^2 = ta^2 + b^2 - tb^2 - t^2a^2 - 2t(1-t)\,ab - (1-2t+t^2)\,b^2 =$$
$$= \text{uff, je to dřina} = \cdots = (t-t^2)\,(a^2 - 2ab + b^2) = (t-t^2)\,(a-b)^2 \ge 0$$

Dále je třeba dokázat, že obecná úsečka (modrá) s koncovými body v M vždy leží „nad“ případně jinou úsečkou s koncovými body přímo na parabole (zelenou) a tudíž i modrá úsečka leží celá v M. Z obrázku to je zase vidět rovnou, matematická argumentace vypadá následovně: Je-li $u \ge a^2$ a $v \ge b^2$ pak pro každé $t \in \langle 0,1\rangle$ je $t\,u + (1-t)\,v \ge t\,a^2 + (1-t)\,b^2$.

13.1 e) Vezmeme $\mathbf{u}, \mathbf{v} \in M$, tedy $\mathbf{Au} \le \mathbf{b}, \mathbf{Cu} = \mathbf{d}$, $\mathbf{Av} \le \mathbf{b}, \mathbf{Cv} = \mathbf{d}$. O vektoru $\mathbf{w} = t\,\mathbf{u} + (1-t)\,\mathbf{v}$ pro $t \in \langle 0,1\rangle$ máme dokázat, že leží v M, tedy splňuje sledované dvě podmínky:

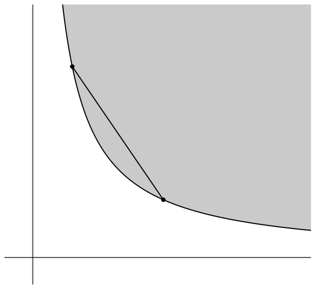
$$\mathbf{Aw} = \mathbf{A}\,(t\,\mathbf{u} + (1-t)\,\mathbf{v}) = t\,\mathbf{Au} + (1-t)\,\mathbf{Av} \le t\,\mathbf{b} + (1-t)\,\mathbf{b} = \mathbf{b},$$
$$\mathbf{Cw} = \mathbf{C}\,(t\,\mathbf{u} + (1-t)\,\mathbf{v}) = t\,\mathbf{Cu} + (1-t)\,\mathbf{Cv} = t\,\mathbf{d} + (1-t)\,\mathbf{d} = \mathbf{d}.$$

13.1 g) Množina $\mathbb{Z}$ není konvexní. Protipříklad je v řešení ve skriptech a jistě každý vymyslí další. Připomínám při té příležitosti známý fakt, že pokud chceme definici použít k ověření konvexity, pak kvůli třem obecným kvantifikátorům v ní použitých ($\forall\mathbf{u}, \forall\mathbf{v}, \forall t$) je třeba v argumentaci zůstat u neurčených těchto třech proměnných, lidově řečeno „postupovat obecně“. Když naopak chceme dle definice konvexitu vyvrátit, tak díky logické negaci těch tří kvantifikátorů ($\exists\,\mathbf{u}$,$\exists\,\mathbf{v}$,$\exists\,t$) stačí najít jeden konkrétní příklad, pro který tvrzení z definice neplatí.

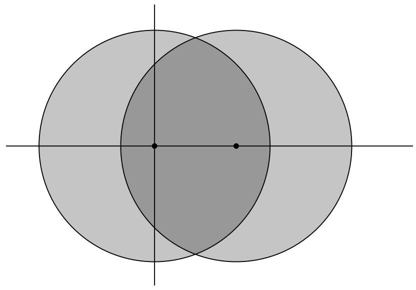
13.2 a) Je to konvexní množina. Je to totiž dokonce konvexní mnohostěn: proměnné se vyskytují v lineárních (neostrých) nerovnostech. Každou rovnost můžeme nahradit dvěma nerovnostmi, takže i tady je $\sum_{i=1}^n x_i = 1$ ekvivalentní s $\sum_{i=1}^n x_i \ge 1$ a $\sum_{i=1}^n x_i \le 1$.

13.2 b) Povrch n-rozměrné koule (sféra) není konvexní, dva protilehlé body ležící na sféře mají spojnici procházející počátkem, přitom počátek není na sféře. Dokonce žádná spojnice dvou různých bodů uvnitř sféry neleží. Konvexním obalem je jednotková koule, tj. množina $\{\mathbf{x} \in \mathbb{R}^n; \|\mathbf{x}\| \le 1\}$.

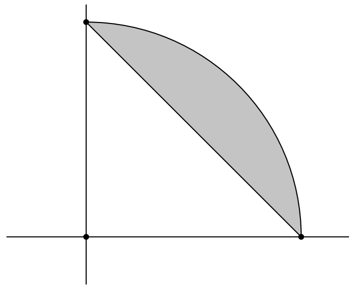
13.2 e) Zkoumaná množina je hyperbola (graf funkce $y = 1/x$) v prvním kvadrantu, protože $xy = 1$ je totéž jako $y = 1/x$. Tento graf není konvexní, např. střed spojnice bodů $[2, \frac{1}{2}]$ a $[\frac{1}{2}, 2]$ je $[\frac{5}{4}, \frac{5}{4}]$ a ten neleží na zadané hyperbole. Obrázek je asi názornější než počítání čtvrtin. Konvexním obalem je množina $\{[x, y] \in \mathbb{R}^2;\ y \geq 1/x,\ x > 0\}$.



12.2 e)



12.2 f)



12.2 g)

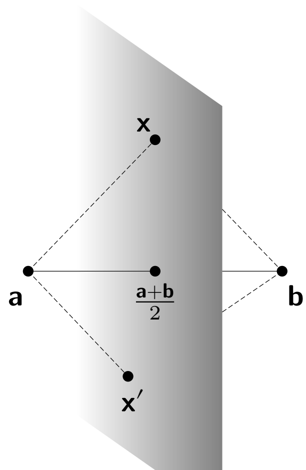
13.2 f) Množina je zadaná jako průnik dvou kruhů o poloměru $\sqrt{2}$ s různými středy. Kruhy samotné jsou konvexní množiny a obecně platí, že průnik konvexních množin je konvexní.

13.2 g) Množina je čtvrtkružnicí v prvním kvadrantu. Není konvexní, protože např. střed spojnice bodů $[1, 0]$ a $[0, 1]$ je $[\frac{1}{2}, \frac{1}{2}]$ a ten neleží v zadané množině. Pozor, konvexní obal není „čtvrtkruh“, jak by možná mohlo někoho bez většího přemýšlení napadnout, ale množina uvedená v řešení skript.

13.2 h,i,j) Řešení ve skriptech je dostatečně podrobné.

13.7 a) Protože je $\mathbf{C}$ pozitivně definitní, je zkoumaná množina elipsoid, tedy něco jako ragbyový míč, takže je konvexní. Není to ale konvexní mnohostěn, protože její hranice se neskládá z „rovných ploch“, jež by se daly vymezit pomocí konečně mnoha lineárních nerovnic. Výjimkou je případ dimenze 1, kdy množina $\{x;\ cx^2 \leq 1\} = \langle -1/\sqrt{c}, 1/\sqrt{c}\rangle$ se dá zapsat dvěma lineárními nerovnostmi $-1/\sqrt{c} \leq x \leq 1/\sqrt{c}$.

13.7 b) Přímka procházející počátkem ve směru vektoru $\mathbf{v}$ je konvexní mnohostěn. Její (ne)rovnicové vyjádření získáme tak, že najdeme všech $n-1$ lineárně nezávislých kolmic $\mathbf{k}_i$ na tuto přímku (tedy bázi řešení homogenní soustavy s jednou rovnicí $\mathbf{v}^T\mathbf{x} = 0$) a přímku vyjádříme jako průnik $n-1$ nadrovin procházejících přímkou s normálovým vektorem $\mathbf{k}_i$. První nadrovina má rovnici $\mathbf{k}_1^T\mathbf{x} = 0$ což se dá dále rozepsat na dvě nerovnosti $\mathbf{k}_1^T\mathbf{x} \leq 0$, $\mathbf{k}_1^T\mathbf{x} \geq 0$. Podobně pro ostatní nadroviny. Přímka je tedy vyjádřitelná pomocí $2(n-1)$ nerovností.



12.7 d)

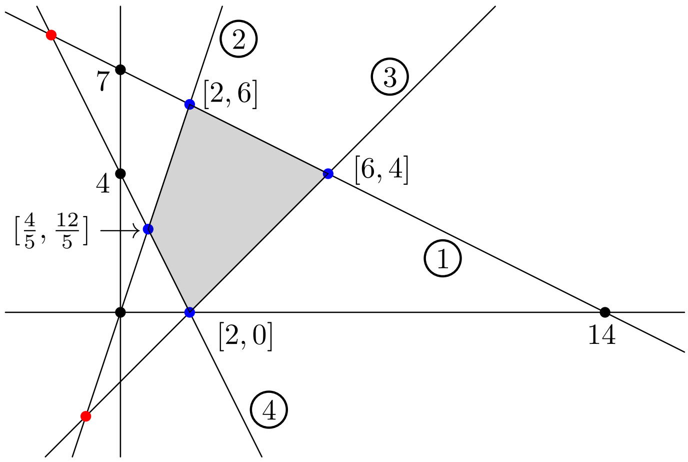
13.7 d) Zkusme se nejprve zamyslet, jak vypadá hranice zkoumané množiny, tedy množina $\{\mathbf{x} \in \mathbb{R}^n;\ \|\mathbf{x} - \mathbf{a}\| = \|\mathbf{x} - \mathbf{b}\|\}$. Je to množina bodů, které mají stejnou vzdálenost od bodu $\mathbf{a}$ jako od bodu $\mathbf{b}$, je to tedy nadrovina procházející středem spojnice bodů $\mathbf{a}$, $\mathbf{b}$ a je na tuto spojnici kolmá. Dá se vyjádřit rovnicí $(\mathbf{a} - \mathbf{b})^T\mathbf{x} = c$, protože je kolmá na $(\mathbf{a} - \mathbf{b})$. Konstantu c je nutno dopočítat z faktu, že nadrovina prochází bodem $(\mathbf{a}+\mathbf{b})/2$, takže $c = (\mathbf{a}-\mathbf{b})^T(\mathbf{a}+\mathbf{b})/2 = (\|\mathbf{a}\|^2 - \|\mathbf{b}\|^2)/2$. Zkoumanou množinu (s nerovností) lze tedy vyjádřit jako $(\mathbf{a} - \mathbf{b})^T\mathbf{x} \geq (\|\mathbf{a}\|^2 - \|\mathbf{b}\|^2)/2$ a je to tedy konvexní mnohostěn.

13.9 Vzdálenost středu koule $\mathbf{c}$ od i-té nadroviny dané rovnicí $\mathbf{a}_i^T\mathbf{x} = b_i$ je rovna $|\mathbf{a}_i^T\mathbf{c} - b_i|/\|\mathbf{a}_i\|$, jak plyne z výsledku cvičení 5.14 c). Přitom $\mathbf{a}_i^T$ je i-tý řádek matice $\mathbf{A}$ v zadání úlohy. Nechť $\mathbf{c}$ vyhovuje všem nerovnicím $\mathbf{Ac} > \mathbf{b}$. Když je r dostatečně malé, aby koule neprotla žádnou nadrovinu, pak máme vzdálenost koule od i-té nadroviny danou vzorcem $|\mathbf{a}_i^T\mathbf{c} - b_i|/\|\mathbf{a}_i\| - r$. Platí $\mathbf{a}_i^T\mathbf{c} - b_i > 0$, takže není třeba psát absolutní hodnotu a máme pro kouli s poloměrem $r > 0$ uvnitř průniku poloprostorů podmínku

$$\forall i\ (\mathbf{a}_i^T\mathbf{c} - b_i)/\|\mathbf{a}_i\| - r \geq 0, \quad \text{tj.} \quad \forall i\ \mathbf{a}_i^T\mathbf{c} - b_i \geq \|\mathbf{a}_i\|\, r.$$

Proměnné optimalizační úlohy jsou $\mathbf{c} \in \mathbb{R}^n$ a $r \in \mathbb{R}$. Maximalizujeme poloměr r za podmínek, že koule leží uvnitř průniku poloprostorů, tedy $\max\{r;\ r \geq 0;\ \mathbf{a}_i^T\mathbf{c} - b_i \geq \|\mathbf{a}_i\|\, r\}$. Maticový zápis vyžaduje nejprve sestavit vektor $\mathbf{z}$ tak, aby $\mathbf{z}_i = \|\mathbf{a}_i\|$. Pak hledáme $\max\{r;\ r \geq 0;\ \mathbf{Ac} \geq r\,\mathbf{z}\}$.

13.10 a) V úloze máme 4 nerovnosti, které vymezují 4 poloroviny podle obrázku. V kroužku u hranice poloroviny je uvedeno číslo nerovnosti, které polorovinu vymezuje a kroužek je umístěn ve směru „do“ poloroviny. Všechny průsečíky těch přímek jsou potenciálními extremálními body. Takže je třeba hledat průsečík přímek odpovídající každé rovnici s každou. Pro případ první rovnice ① s druhou ② máme:



$$\left[\begin{array}{cc|c} -1 & -2 & -14 \\ 3 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 3 & -1 & 0 \\ 0 & -7 & -42 \end{array}\right] \sim \left[\begin{array}{cc|c} 3 & 0 & 6 \\ 0 & 1 & 6 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & \mathbf{2} \\ 0 & 1 & \mathbf{6} \end{array}\right]$$

Toto provedeme s každou dvojicí rovnic. Máme tedy 6 průsečíků. Z obrázku je vidět, že z těch 6 průsečíků dva (červené) nejsou vůbec ve zkoumané množině, tj. máme 4 extremální body (modré).

13.10 b) Je-li dáno m nerovnic s n proměnnými, $m \geq n$, pak je potřeba najít kandidáty všech extremálních bodů, tj. vybrat všechny n-tice odpovídajících rovnic ze zadané množiny m rovnic. Těch je dohromady $\binom{m}{n}$. Pro každou takovou n-tici rovnic je třeba nalézt řešení jako potencionální extremální bod. Když řešení neexistuje, pak daná n-tice negeneruje potenciální extremální bod. Ze všech potenciálních bodů je pak třeba vybrat jen ty, které navíc leží v zadané množině, tj. splňují soustavu $\mathbf{A}\mathbf{x} \geq \mathbf{b}$. Omlouvám se, ale program realizující tento algoritmus jsem nesesmolil.

Doplním k tomuto příkladu ještě jedno pozorování. Člověk, má-li málo rovnic a úloha je v $\mathbb{R}^2$ (jako v případě 12.10 a), může po vytvoření náčrtku rovnou vyloučit body, které jsou „očividně“ mimo množinu (červené body) a nemusí ani počítat hodnoty takových průsečíků. Stroj tuto „očividnou“ intuici obvykle nemá naprogramovánu a musí tedy vypočítat opravdu všechny potenciální extremální body a pak z nich vyloučit ty, které nevyhovují zadané soustavě rovnic. Na druhé straně je stroj poněkud rychlejší než člověk a zvládne i větší dimenze než jen $\mathbb{R}^2$.

13.11 a) Příklad 13.3 nemá bohužel konvexní množiny očíslované, ale jen opuntíkované. Každý jeden puntík v mém textu po řadě odpovídá puntíku v textu skript. Vyžaduje to trochu soustředění nebo počítání puntíků. Držím Vám palec, ať se Vám nedělají z puntíků mžitky před očima. • nemá přímku ani extremální bod (dále jen „e. bod“), • jakákoli přímka v množině leží, tj. bez e. bodu, • e. bod je a nebo a a b, • při dim > 0 (tj. afinní podprostor není jen samotný bod) obsahuje přímku, tj. nemá e. body, • e. bod je $\mathbf{x}$, • obsahuje přímky, je bez e. bodu, • je to vícedimenzionální „první kvadrant“, e. bod je $\mathbf{0}$, • neprázdný panel obsahuje přímku (např. na hranici $\mathbf{a}^T\mathbf{x} = b_2$) a je bez e. bodů, • vrcholy hyperkrychle jsou e. body, • vrcholy hyperkvádru jsou e. body, • vrcholy simplexu jsou e. body, • e. bod je např. $[1, 0, 0, \ldots, 0]$, • např. $[1, 0, 0, \ldots, 0]$ nebo $[-1, 0, 0, \ldots, 0]$ jsou e. body.

13.11 b) Množina je nekonečný hranol (ve směru osy z) s průřezem tvaru trojúhelníka. Například přímka $\mathbf{0} + t\,[0, 0, 1]^T$ v něm celá leží, tedy množina nemá e. bod.

13.11 c) Je to panel, který obsahuje přímku (např. $y = -x$), tj. neobsahuje e. bod, jak jsme o tom už v obecnějším případě **hovořili** u cvičení 12.11 a), puntík osmý. Nálseduje citát z klasické divadelní hry. Král: „Vy jste spolu **nehovořili**?“ Pseudo-princezna: „**Hovořili**.“ Princ Jasoň: „**Nehovořili**, za současné výjimečné situace a vládních opatřeních se nám spolu obtížně **hovoří**.“ Do řešení, které budete zasílat, můžete připojit název citované hry, ale není to nutné.