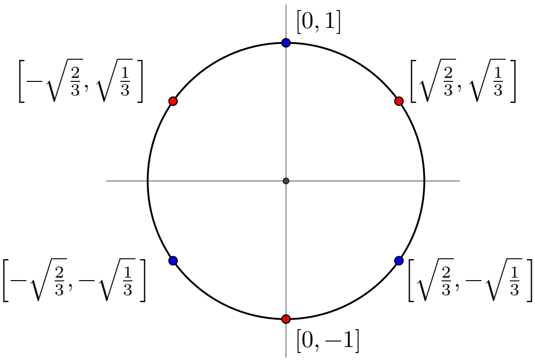
Komentář k úlohám probíraným na cvičeních dle plánu 09

Petr Olšák

11.1 d) Úlohu můžeme vyřešit třemi způsoby a porovnat efektivitu práce.

1. Převedením do polárních souřadnic máme $(x, y) = (\cos t, \sin t)$, takže hledáme extrémy funkce $\cos^2 t \sin t$, derivace rovná nule: $-2\cos t \sin^2 t + \cos^2 t \cos t = \cos t\,(\cos^2 - 2\sin^2 t) = 0$. Takže $\cos t = 0$, tj. $t = \pm\frac{\pi}{2}$ nebo $\cos^2 t - 2\sin^2 t = 0$, tj. $1 - 3\sin^2 t = 0$. Z první části řešení máme $x = 0, y = \pm 1$. Z druhé části řešení sice nevyjádříme přesně t, ale víme, že $\sin^2 t = \frac{1}{3}$, takže $y^2 = \frac{1}{3}$ a tedy $y = \pm\frac{1}{\sqrt{3}}$. Na kružnici tedy máme šest bodů $(\frac{\sqrt{2}}{\sqrt{3}}, \frac{1}{\sqrt{3}})$, $(0, 1)$, $(-\frac{\sqrt{2}}{\sqrt{3}}, \frac{1}{\sqrt{3}})$, $(-\frac{\sqrt{2}}{\sqrt{3}}, -\frac{1}{\sqrt{3}})$, $(0, -1)$, $(\frac{\sqrt{2}}{\sqrt{3}}, -\frac{1}{\sqrt{3}})$. Jak plyne z funkčních hodnot, v těchto bodech je po řadě max (kladná hodnota), min (nula) max (kladná hodnota) min (záporná hodnota) max (nula) min (záporná hodnota). Na obrázku jsou maxima vyznačena červeně a minima modře, přirozeně se na kružnici střídají.



2. Vyjádříme-li z podmínky $x^2 = 1 - y^2$ funkci f, máme $f = (1 - y^2)y$, na intervalu $\langle -1, 1\rangle$, což po zderivování a položení rovno nule dává $y^2 = \frac{1}{3}$, a to dává čtyři z šesti bodů. Dále musíme nezapomenout, že lokálními extrémy této funkce jedné proměnné jsou i krajní body intervalu, tedy $y = \pm 1$, což dává zbylé dva body.

3. Lagrangeova funkce $L(x, y, \lambda) = x^2 y + \lambda(x^2 + y^2 - 1)$. Po jejím zderivování postupně podle všech třech proměnných a položením rovno nule máme $2xy + 2x\lambda = 0$, $x^2 + 2y\lambda = 0$ a třetí rovnice je okrajová podmínka $x^2 + y^2 = 1$. Z první rovnice po vytknutí x máme $x = 0$ nebo $y = -\lambda$. První možnost dává dva body $(0, 1)$, $(0, -1)$ a druhou možnost dosadíme do druhé rovnice a s využitím třetí opět máme $y^2 = \frac{1}{3}$ což dává zbylé čtyři body s extrémy.

11.4 c) **1.** Z okrajové podmínky plyne $x^2 = \frac{1}{y}$, takže funkci $x^2 + y^2$, kterou minimalizujeme, máme převedenu do jedné proměnné $f(y) = \frac{1}{y} + y^2$. Zderivování a položení rovno nule dává $-\frac{1}{y^2} + 2y = 0$, tedy $y^3 = \frac{1}{2}$, $y = 1/\sqrt[3]{2}$ a zpětné dosazení pro x dává $x^2 = \sqrt[3]{2}$, tedy $x = \pm\sqrt[6]{2}$.

2. Ještě jednou jinak. $L(x, y, \lambda) = x^2 + y^2 + \lambda(x^2 y - 1)$. Zderivování této funkce postupně podle všech tří proměnných a položení rovno nule dává $2x + 2\lambda xy = 0$, $x(1 + \lambda y) = 0$ a třetí rovnice je okrajová podmínka. Z první rovnice je $x = 0$ nebo $\lambda y = -1$. První možnost nepřipadá v úvahu, protože z ní a z druhé rovnice plyne, že také $y = 0$, ale dvojice $(0, 0)$ nevyhovuje okrajové podmínce. Druhá možnost dává $\lambda = -\frac{1}{y}$ a když k tomu přidáme fakt z okrajové podmínky $x^2 = \frac{1}{y}$ a oba mezivýsledky dosadíme do druhé rovnice, máme $2y + \frac{-1}{y}\frac{1}{y} = 0$, což vede na $y = \frac{1}{2y^2}$ a to vede na stejný výsledek, jako v případě 1.

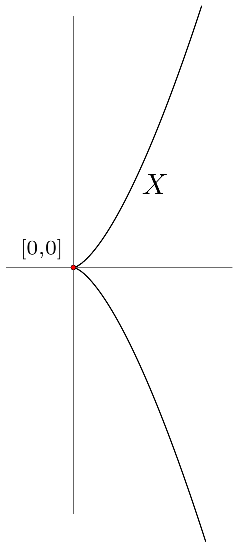
11.5 d) Pokud mě paměť neklame, tak půllitr jsem na cvičení dělal v době, kdy jsme se mohli fyzicky vídat. Budu komentovat tedy něco jiného, třeba 10.5 b).

11.5 b) Hledá se minimum $ab + 2ac + 2bc$ za podmínky $abc = 1$. Strany a, b značí základnu a c výšku krabice bez víka. Okrajovou podmínku $c = 1/(ab)$ zaneseme do účelové funkce a máme z ní funkci bez podmínky se dvěma proměnnými $f(a, b) = ab + 2/b + 2/a$. Hledáme stacionární body zderivováním v obou proměnných a položením nule: $b = 2/a^2 = 0$, $a - 2/b^2 = 0$, tedy $a^2 b = 2$ a taky $ab^2 = 2$, tj. $a^2 b - ab^2 = 0$. Po vytknutí a máme $a = 0$ což zamítneme, neboť se to neslučuje s okrajovou podmínkou. Druhá možnost vede na nepřekvapivý výsledek $ab = b^2$ a tedy $a = b$. Takže z $a^2 b = 2$ plyne $a^3 = 2$, což vede na výsledek $a = b = \sqrt[3]{2}$ a z okrajové podmínky pak $c = 1/\sqrt[3]{4}$. Protože ale nás nezajímal konkrétně objem=1, výsledek určuje pouze poměry stran, tedy $a = b$, $c = 1/(ab)$ a $a/c = \sqrt[3]{8} = 2$. Všechny strany je třeba pronásobit takovým koeficientem, aby abc bylo rovno požadovanému objemu.

2. Metoda Lagrangeových multiplikátorů dá po menší námaze stejný objevný výsledek, že totiž $a = b$. Zkuste si to spočítat sami.

3. Asi nejrychleji jsme hotovi při použití „selského rozumu“. Ta krabice nemůže být „asymetrická“ ve směru stran a a b, protože žádná z těch stran není v účelové funkci nijak preferovaná. Takže $a = b$ a z okrajové podmínky $c = 1/(ab)$. Nyní po využití okrajové podmínky je účelová funkce jen v jedné proměnné a její minimum najedeme v bodě $a = \sqrt[3]{2}$.

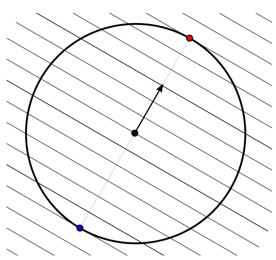
11.8 Metoda Lagrangeových multiplikátorů kolabuje obdobně jako státní rozpočet po koronavirové pandémii kvůli tomu, že množina X z okrajové podmínky nemá v bodě minima $(0,0)$ tečný prostor. Je tam hrot. Množina $X = \{(x,y);\ x^3 = y^2\}$ se skládá z grafů funkcí $y = \pm\sqrt{x^3}$. Zkusíme si v tomto příkladě zahájit metodu Lagrangeových multiplikátorů a uvidíme, co to udělá…



$L = x + \lambda(x^3 - y^2)$. Zderivováním a položením rovno nule dostáváme $1 + 3\lambda x^2 = 0$ a $-2\lambda y = 0$ a okrajová podmínka. Z druhé rovnice plyne, že musí být $y = 0$ nebo $\lambda = 0$. První případ společně s okrajovou podmínkou $x^3 = y^2$ vede na $x = 0$, ale to nevyhovuje první rovnici. Druhý případ přímo nevyhovuje první rovnici. Takže soustava sestavená metodou Lagranegeových multiplikátorů nemá řešení.

11.12 a) Bez té pohádky o entropii hledáme maximum funkce $f(\mathbf{p}) = -\sum_{i=1}^{n} p_i \log p_i$ za podmínky $\sum p_i = 1$. Lagrangeova funkce je $L(\mathbf{p},\lambda) = -\sum_{i=1}^{n} p_i \log p_i + \lambda\left(\sum_{i=1}^{n} p_i - 1\right)$ a její derivace podle každé proměnné p_i položíme rovny nule: $\frac{\partial L}{\partial p_i} = L'_i = -(\log p_i + p_i \frac{1}{p_i}) + \lambda = 0$, tedy $\lambda = \log p_i + 1$. Protože pravá strana naposledy zapsané rovnosti je prostá v p_i a máme zde n takových rovností rovných stále stejnému λ, musí se všechny p_i vzájemně rovnat. Nebo jinak: aplikujeme prostou exp na obě strany rovnosti (po přesunutí jedničky nalevo) a máme $e^{\lambda-1} = p_i$. Podstatné tedy je, že $p_1 = p_2 = \cdots = p_n = \frac{1}{n}$. Ta poslední rovnost plyne z okrajové podmínky. (Pozn: metoda „selského rozumu“ nám taky říká, že se proměnné musejí rovnat, když žádná není ve vzorcích preferovaná.) To je stacionární bod. Hodnota je v tomto bodě kladná, protože logaritmy jsou záporné a před vzorečkem pro f je jedno mínus. Jdeme-li limitně ke kraji definičního oboru (který už není součástí definičního oboru), tj. jedno z $p_i = 1$ a ostatní jsou nulové, pak $\lim f(\mathbf{p}) = 0$, protože $1 \log 1 = 0$ a $\lim_{t\to 0} t \log t = 0$ z L'Hospitalova pravidla. Takže v bodě $(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$ jsme našli maximum funkce f.

11.? Maximalizujeme $\mathbf{a}^T\mathbf{x}$ za podmínky $\mathbf{x}^T\mathbf{x} = 1$. $L = \mathbf{a}^T\mathbf{x} + \lambda(\mathbf{x}^T\mathbf{x} - 1)$. Nyní naráz zderivujeme dle všech n proměnných a příslušných n rovnic položíme rovno nule: $L' = \mathbf{a}^T + \lambda(2\mathbf{x}^T) = \mathbf{0}^T$. K tomu ještě přidáme okrajovou podmínku $\mathbf{x}^T\mathbf{x} = 1$. Systém rovnic transponuji a spočítám z něj $\mathbf{a} = -2\lambda\mathbf{x}$, takže $\mathbf{x} = \mathbf{a}/(-2\lambda)$. Tento výsledek dosadím do okrajové podmínky



$(\mathbf{a}/(-2\lambda))^T(\mathbf{a}/(-2\lambda)) = \|\mathbf{a}\|^2/(4\lambda^2) = 1$, takže $\|\mathbf{a}\|^2 = 4\lambda^2$ a máme spočítáno $\lambda = \pm\|\mathbf{a}\|/2$. Když tento poznatek uplatníme v systému rovnic $(L')^T = \mathbf{0}$, dostáváme $\mathbf{a} \pm (\|\mathbf{a}\|/2)\,(2\mathbf{x}) = \mathbf{0}$ a z toho plyne $\mathbf{x} = \mp\mathbf{a}/\|\mathbf{a}\|$, což je normalizovaný gradient funkce $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x}$. Protože funkce roste ve směru gradientu, je maximum v bodě $\mathbf{a}/\|\mathbf{a}\|$ a minimum v bodě $-\mathbf{a}/\|\mathbf{a}\|$. Geometricky: hledáme na sféře místo, které maximalizuje lineární funkci plynule rostoucí ve směru gradientu. Na obrázku pro případ 2D jsou nakresleny její vrstevnice a její gradient. Maximum je přirozeně v místě normovaného gradientu.

11.18 $L = \mathbf{x}^T\mathbf{C}\mathbf{x} + \boldsymbol{\lambda}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$. Tentokrát (poprvé v tomto cvičení) máme více než jednu proměnou λ. Je jich obecně m a jsou sdruženy do vektoru $\boldsymbol{\lambda} \in \mathbb{R}^m$, kde m je počet řádků matice $\mathbf{A}$. Okrajová podmínka má totiž m lineárních rovnic, každá je reprezentovaná jedním řádkem matice $\mathbf{A}$ a jedním prvkem z vektoru $\mathbf{b}$. Maticovým derivováním podle $\mathbf{x}$ a položením nulovému vektoru dostáváme $L' = 2\mathbf{x}^T\mathbf{C} + \boldsymbol{\lambda}^T\mathbf{A} = \mathbf{0}^T$. Tento systém rovnic zapíšeme transponovaně $-2\,\mathbf{C}\mathbf{x} = \mathbf{A}^T\boldsymbol{\lambda}$ a přihodíme okrajovou podmínku $\mathbf{A}\mathbf{x} = \mathbf{b}$. Máme systém $m + n$ rovnic s n neznámými x_i a m neznámými λ_j. První podsystém násobíme zleva maticí $\mathbf{C}^{-1}$ ($\mathbf{C}$ je regulární, protože je pozitivně definitní) a máme $-2\,\mathbf{x} = \mathbf{C}^{-1}\mathbf{A}^T\boldsymbol{\lambda}$. Dále zleva přihodíme $\mathbf{A}$ a vzpomeneme si na systém okrajových podmínek $-2\,\mathbf{b} = -2\,\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{C}^{-1}\mathbf{A}^T\boldsymbol{\lambda}$. Z tohoto zápisu lze vyjádřit $\boldsymbol{\lambda}$ jako $\boldsymbol{\lambda} = -2\,(\mathbf{A}\mathbf{C}^{-1}\mathbf{A}^T)^{-1}\mathbf{b}$ což zpětně uplatníme v prvním podsystému rovnic $-2\,\mathbf{C}\mathbf{x} = \mathbf{A}^T(-2\,(\mathbf{A}\mathbf{C}^{-1}\mathbf{A}^T)^{-1}\mathbf{b})$. Po vykrácení konstantou -2 a vynásobení maticí $\mathbf{C}^{-1}$ zleva máme konečný výsledek $\mathbf{x} = \mathbf{C}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{C}^{-1}\mathbf{A}^T)^{-1}\mathbf{b}$.