Zkouška OPT 19.1.2024 | Jméno: Příjmení:

Řešení pište do připravených mezer. Odevzdává se pouze tento čistopis.
Každý příklad musí mít nejen odpověď, ale i postup. Odpověď bez postupu nestačí.

---

1. Jste na pravém břehu řeky široké 1 km a chcete se dostat ke stanu na levém břehu, který je 2 km po proudu od bodu, který je na levém břehu nejblíže vám. Řeka teče velmi pomalu, zanedbatelnou rychlostí. Plavete rychlostí 1 km/h a chodíte rychlostí 3 km/h. Jaký je nejkratší možný čas, za který se můžete dostat ke stanu?

   (a) **(2b)** Zformulujte tuto úlohu. Formulaci zvolte tak, aby šla dobře řešit. Ilustrujte ji obrázkem.

   (b) **(2b)** Úlohu vyřešte. Odpověď bude hodnota nejkratšího času.

   Tento příklad je ve skriptech, jen s jinými čísly. Minimalizujeme funkci $t(x) = \frac{1}{1}\sqrt{1+x^2} + \frac{1}{3}(2-x)$. Zderivujeme, položíme rovno nule. Minimální čas nastane pro $x = 1/(2\sqrt{2}) \approx 0.3536$ a je $t(x) = \frac{2}{3}(1+\sqrt{2}) \approx 1.6095$ hod $\approx$ 96.57 min.

---

2. Máme funkci $f(x,y) = x^2 + 2ay(x+y) - 4x - 8y + 1$ s parametrem $a \in \mathbb{R}$.

   (a) **(1b)** Je funkce f polynom? Jestliže ano, tak jakého stupně? Je homogenní? Polynom stupně 2, nehomogenní.

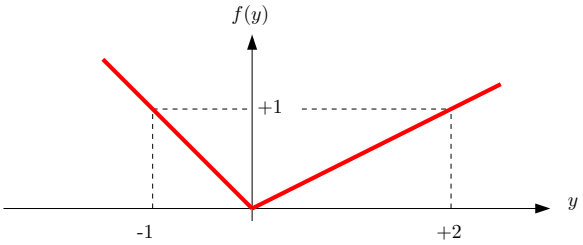
   (b) **(2b)** Pro jaké hodnoty parametru a je funkce konvexní? Proč?

   (c) **(1b)** Pro jaké hodnoty parametru a je funkce konkávní? Proč?

   Pro konvexitu musí být Hessova matice (pro pohodlí napíšeme její dvojnásobek) $2f''(x,y) = \begin{bmatrix} 1 & a \\ a & 2a \end{bmatrix}$ positivně semidefinitní. To nastane právě tehdy, když $a \geq 0$ a $2a - a^2 \geq 0$, tj. $a \in [0,2]$. Konkávní bude pro Hessián negativně semidefinitní, což nebude nikdy, protože Hessián má na diagonále kladné číslo (to negativně semidef. matice mít nesmí).

---

3. **(3b)** Jsou dány vektory $\mathbf{a}_i \in \mathbb{R}$ a čísla $b_i \in \mathbb{R}$ pro $i = 1, \ldots, m$. Dále je dána funkce $f\colon \mathbb{R} \to \mathbb{R}$ obrázkem. Chceme minimalizovat funkci $g(\mathbf{x}) = \max_{i=1}^{m} f(\mathbf{a}_i^T\mathbf{x} - b_i)$ na množině $\mathbb{R}^n$. Zformulujte jako lineární program.



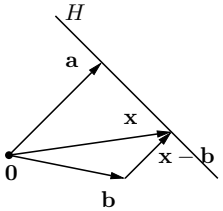
Skoro stejné jako cvičení 12.4.e ve skriptech. Fce f je maximum dvou afinních funkcí, $f(y) = \max\{-y, y/2\}$. Tedy $g(\mathbf{x})$ je maximum $2m$ funkcí $-\mathbf{a}_i^T\mathbf{x} + b_i$ a $(\mathbf{a}_i^T\mathbf{x} - b_i)/2$ pro $i = 1, \ldots, m$. Zavedeme pomocnou proměnnou z a napíšeme jako LP: minimalizuj z za podmínek $-\mathbf{a}_i^T\mathbf{x} + b_i \leq z$, $(\mathbf{a}_i^T\mathbf{x} - b_i)/2 \leq z \; \forall i = 1, \ldots, m$, s proměnnými $\mathbf{x} \in \mathbb{R}^n$ a $z \in \mathbb{R}$.

---

4. Jsou dány $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, kde $\mathbf{a}^T\mathbf{a} = 1$. Hledáme bod $\mathbf{x} \in \mathbb{R}^n$ splňující $\mathbf{a}^T\mathbf{x} = 1$, který je nejblíže bodu $\mathbf{b}$.

   (a) **(3b)** Úlohu vyřešte. Výsledkem bude vzorec pro $\mathbf{x}$.
   To má několik způsobů řešení. Lze to vidět jako nalezení vadálenosti bodu od afinního podprostoru, což se dělá ve skriptech v §5.2.1. Jiný je napsat si úlohu jako min $(\mathbf{b} - \mathbf{x})^T(\mathbf{b} - \mathbf{x})$ za podm. $\mathbf{a}^T\mathbf{x} = 1$ a použít metodu Lagrangeových multiplikátorů. Vyjde $\mathbf{x} = (1 - \mathbf{a}^T\mathbf{b})\mathbf{a} + \mathbf{b}$.

   (b) **(3b)** Nakreslete situaci pro $n = 2$. Na obrázku bude počátek (tj. bod $\mathbf{0}$), zvolené vektory $\mathbf{a}, \mathbf{b}$ (zvolte je v obecné vzájemné poloze), množina přípustných řešení úlohy (označte ji H) a vektor $\mathbf{x}$. Pokud jsou v obrázku nějaké dvojice objektů kolmé či rovnoběžné, vyznačte to standardními značkami.
   Viz obrázek. Množina H je nadrovina, zde přímka, bod $\mathbf{a}$ leží na H. Značky pro kolmost a rovnoběžnost jsem nekreslil, neb na počítači se to špatně kreslí – napíšu to slovy: vektor $\mathbf{a}$ je rovnoběžný s vektorem $\mathbf{x} - \mathbf{b}$ a oba tyto vektory jsou kolmé na nadrovinu H.



Zkouška OPT 19.1.2024 | Jméno: Příjmení:

5. Je dáno n čísel $c_1, \ldots, c_n \in \mathbb{R}$. Maximalizujeme $\sum_{i=1}^n c_i x_i$ za podmínek $-1 \leq x_i \leq 1$ pro $i = 1, \ldots, n$.

   (a) **(2b)** Vyřešte úvahou. Výsledek bude co nejjednodušší vzorec pro optimální *hodnotu* (nikoliv *argument*!).
   $\sum_{i=1}^n |c_i|$

   (b) **(3b)** Napište duální úlohu a zjednodušte ji (např. pokud jste duální úlohu získali v maticovém tvaru, nenechávejte ji v maticovém tvaru ale roznásobte matice a zjednodušte).
   Minimalizujeme $\sum_i (u_i + v_i)$ za podmínek $u_i \geq 0$, $v_i \geq 0$, $v_i - u_i = c_i$. Duální proměnné jsou u_i, v_i pro $i = 1, \ldots, n$. Může být i v mírně jiném ekvivalentním tvaru, např. jedna sada proměnných je vynásobená mínus jednou.

   (c) **(1b)** Napište podmínky komplementarity.
   - Pro každé i je $x_i = -1$ nebo $u_i = 0$ (neboli $(x_i + 1)u_i = 0$).
   - Pro každé i je $x_i = 1$ nebo $v_i = 0$ (neboli $(x_i - 1)v_i = 0$).

   (d) **(2b)** Najděte číselné hodnoty optimálních primárních a duálních proměnných pro toto zadání: $n = 3$, $(c_1, c_2, c_3) = (-2, 3, 4)$. Jestliže primární či duální úloha má více optimálních řešení, popište je všechna. Jestliže je primární či duální úloha nepřípustná či neomezená, vysvětlete proč.
   $(x_1, x_2, x_3) = (-1, 1, 1)$
   $(u_1, u_2, u_3) = (2, 0, 0)$
   $(v_1, v_2, v_3) = (0, 3, 4)$.

---

6. **(3b)** Danými body $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ chceme proložit nadrovinu $\{\, \mathbf{x} \in \mathbb{R}^d \mid \mathbf{a}^T\mathbf{x} = b \,\}$ tak, aby součet čtverců vzdáleností bodů od nadroviny byl minimální. Napište postup (tj. posloupnost kroků, nemusí to být kód v nějakém programovacím jazyce) jak najít parametry nadroviny $\mathbf{a} \in \mathbb{R}^d$, $b \in \mathbb{R}$.

   To je příklad na PCA. Nejdříve od každého bodu $\mathbf{x}_i$ odečteme těžiště $\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^n \mathbf{x}_i$. Pak naskládáme tyto body jako sloupce do matice $\mathbf{X} \in \mathbb{R}^{d \times n}$. Uděláme vlastní rozklad matice $\mathbf{X}\mathbf{X}^T \in \mathbb{R}^{d \times d}$. Vektor $\mathbf{a}$ je normálový vektor hledané nadroviny (tj. tvoří bázi jejího ortogonálního doplňku), tedy $\mathbf{a}$ je vlastní vektor matice $\mathbf{X}\mathbf{X}^T$ příslušný jejímu *nejmenšímu* vlastnímu číslu. Alternativně uděláme SVD $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ matice $\mathbf{X}$, pak vektor $\mathbf{a}$ je sloupec matice $\mathbf{U}$ příslušný nejmenšímu singulárnímu číslu (tj. poslední sloupec, řadíme-li sing. čísla sestupně).

   Málokdo vymyslel, jak získat skalár b. To už není mechanický postup, ale musí se malinko zapřemýšlet. Z nadroviny $\mathbf{a}^T\mathbf{x} = b$ již známe normálu $\mathbf{a}$ a navíc víme, že nadrovina prochází těžištěm $\bar{\mathbf{x}}$. Z toho plyne $\mathbf{a}^T\bar{\mathbf{x}} = b$.

---

7. Jsou dány body $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^2$ v rovině. Hledáme kružnici takovou, aby součet čtverců vzdáleností bodů od kružnice byl minimální.

   (a) **(3b)** Formulujte tuto optimalizační úlohu. **Z formulace musí být jasné, co jsou proměnné úlohy, co účelová funkce a co omezení.**
   Bylo předmětem domácí úlohy (takže kdo to neměl tak buď domácí úlohu nepochopil nebo opsal). Vzdálenost bodu $\mathbf{a}$ od kružnice se středem $\mathbf{c}$ a poloměrem r je $|\|\mathbf{a} - \mathbf{c}\|_2 - r|$. Takže minimalizujeme funkci $f(\mathbf{c}, r) = \sum_i (\|\mathbf{a}_i - \mathbf{c}\|_2 - r)^2$ přes proměnné $\mathbf{c} \in \mathbb{R}^2$ a $r \in \mathbb{R}$ bez omezujících podmínek.

   (b) **(2b)** Váš spolužák tvrdí, že střed optimální kružnice leží v těžišti daných bodů. Má pravdu? Odpověď dokažte.
   Nemá. Protipříklad je množina tří bodů které jsou téměř ale ne úplně kolineární – tyto body leží na právě jedné kružnici (která je tedy optimálním řešením naší úlohy) ale jejich těžiště je úplně jinde než ve středu této kružnice.

   (c) **(2b)** Kdybychom znali optimální polohu středu kružnice, lze jednoduše spočítat její optimální poloměr? Jestliže ano, tak jak? Jestliže ne, vysvětlete.
   Jde to snadno. Pro známý střed $\mathbf{c}$ chceme minimalizovat funkci $f(\mathbf{c}, r)$. Označíme pro názornost $\|\mathbf{a}_i - \mathbf{c}\| = b_i$, tedy minimalizujeme $\sum_i (b_i - r)^2$ přes $r \in \mathbb{R}$. Řešení je aritmetický průměr $\frac{1}{n}\sum_i b_i$ čísel b_i.

---

8. Připomeňme, že vzdálenost množin $X, Y \subseteq \mathbb{R}^n$ je rovna číslu $\min_{\mathbf{x} \in X,\, \mathbf{y} \in Y} \|\mathbf{x} - \mathbf{y}\|_2$ (pokud minimum existuje).

   (a) **(3b)** Najděte vzdálenosti dvou množin v $\mathbb{R}^2$: přímky $y = 2x - 3$ a paraboly $y = x^2$. Výsledkem bude hodnota vzdálenosti mezi těmito množinami (tj. jediné číslo).
   Mechanický postup: vzdálenost bodu (x, y) od dané přímky je $|2x - y - 3|/\sqrt{5}$. Bod na parabole parametrizujeme jako $(x, y) = (x, x^2)$. Tedy minimalizujeme $(2x - x^2 - 3)^2$, kde jsme pro pohodlí zanedbali konstantu

$\sqrt{5}$ a umocnili na druhou. Zderivujeme, položíme rovno nule, vyjde $x = 1$. Tedy bod na parabole nejbližší přímce je $(1, 1)$, vzdálenost tohoto bodu od přímky je $2/\sqrt{5}$.

Alternativně můžeme použít tvrzení z podúlohy (b). Pokud (x, y) je nejbližší bod na parabole k dané přímce, tak normála (gradient) k parabole v tom bodě musí být rovnoběžná (tedy násobkem) normálového vektoru přímky $(2, -1)$. Je třeba opatrnost při počítání normály (= gradientu) k parabole: musíme parabolu chápat jako množinu v rovině $\mathbb{R}^2$ (tedy jako graf funkce x^2), tedy jako (nulovou) vrstevnici funkce $f(x, y) = x^2 - y$. Ta má gradient $\nabla f(x, y) = (2x, -1)$. Tedy máme podmínku $(2x, -1) = \lambda(2, -1)$ (z toho nějak čouhají Lagranegeovy multiplikátory, že jo?), z toho $x = \lambda = 1$.

(b) **(1b)** Uvažujme následující tvrzení. Nechť $X = \{\, \mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) = 0 \,\}$, kde $f\colon \mathbb{R}^n \to \mathbb{R}$ je spojitě diferencovatelná funkce. Nechť $Y = \{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T\mathbf{x} = b \,\}$, kde $\mathbf{0} \neq \mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$. Nechť $X \cap Y = \emptyset$. Nechť $\mathbf{x}^*$ je bod množiny X, který je nejblíže nadrovině Y, a zároveň splňuje $\nabla f(\mathbf{x}^*) \neq \mathbf{0}$. Tvrdíme, že tečný prostor k množině X v bodě $\mathbf{x}^*$ je rovnoběžný s nadrovinou Y.
Jak by se tento výsledek použil na úlohu (a)? Jako odpověď napište, co v kontextu úlohy (a) jsou $n, f, \mathbf{a}, b$.

$n = 2$, $f(x, y) = y - x^2$, $\mathbf{a} = (2, -1)$, $b = 3$.

(c) **(1b)** Dokažte tvrzení z úlohy (b). Použijte metodu Lagrangeových multiplikátorů.

Toto tvrzení ihned dostaneme z podmínky prvního řádu na extrémy vázané rovnostmi (metoda Lagr multiplikátorů). Vzdálenost bodu $\mathbf{x}$ od nadroviny Y je $|\mathbf{a}^T\mathbf{x} - b|/\|\mathbf{a}\|$. Hledejme bod $\mathbf{x} \in X$ pro který je tato vzdálenost nejmenší, tedy minimalizujme $\frac{1}{2}(\mathbf{a}^T\mathbf{x} - b)^2$ (kde jsme pro pohodlí vypustili jmenovatel $\|\mathbf{a}\|$, umocnili na druhou a přidali $\frac{1}{2}$) za podmínky $f(\mathbf{x}) = 0$. Podmínky prvního řádu na extrémy vázané rovností dají $(\mathbf{a}^T\mathbf{x} - b)\mathbf{a} = \lambda\nabla f(\mathbf{x})$. Protože dle předpokladů je $\mathbf{x} \notin Y$ (neboli $\mathbf{a}^T\mathbf{x} \neq b$) a $\nabla f(\mathbf{x}) \neq \mathbf{0}$, musí být $\mathbf{a}$ násobkem gradientu $\nabla f(\mathbf{x})$, tedy rovnoběžný s tečným prostorem k množině X v bodě $\mathbf{x}$ (protože gradient je normála k tečnému prostoru, za předpokladu $\nabla f(\mathbf{x}) \neq \mathbf{0}$).

Alternativně můžeme minimalizovat $\|\mathbf{x} - \mathbf{y}\|^2$ za podmínek $f(\mathbf{x}) = 0$ a $\mathbf{a}^T\mathbf{y} = b$. Je $L(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|^2 + \alpha f(\mathbf{x}) + \beta(b - \mathbf{a}^T\mathbf{y})$. Derivace: $L_x = \mathbf{x} - \mathbf{y} + \alpha\nabla f(\mathbf{x}) = \mathbf{0}$, $L_\mathbf{y} = \mathbf{y} - \mathbf{x} - \beta\mathbf{a} = \mathbf{0}$. Sečtení těchto dvou rovnic dá $\alpha\nabla f(\mathbf{x}) = \beta\mathbf{a}$. To může platit buď pro $\alpha = \beta = 0$ nebo pro $\alpha, \beta \neq 0$. První případ by implikoval $\mathbf{x} = \mathbf{y}$, což je nemožné, protože se nadpovrch a nadrovina neprotínají. Tedy $\nabla f(\mathbf{x})$ je násobkem $\mathbf{a}$, což je dokazované tvrzení.