

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

1. (3b) Máme elipsu s délkami poloos a, b . Do této elipsy chceme vepsat obdélník s maximálním obsahem. Předpokládejte přitom, že strany obdélníku jsou rovnoběžné s osami elipsy. Jaký obsah bude mít tento obdélník?

Rovnice elipsy: $(x/a)^2 + (y/b)^2 = 1$. Uvažujeme jen čtvrtinu elipsy $x, y \geq 0$. Pak $y = b\sqrt{1 - (x/a)^2}$. Obsah čtvrtiny obdélníka je $xy = xb\sqrt{1 - (x/a)^2}$. Najdeme minimum této fce jedné proměnné. Vyjde $x = a/\sqrt{2}$, $y = b/\sqrt{2}$. Tedy obsah čtvrtiny obdélníka je $ab/2$, obsah celého obdélníka $2ab$.

2. (2b) Hledáme obdélník s nejmenším obvodem, který má jednu stranu rovnoběžnou s osou x , druhou stranu rovnoběžnou s osou y , a obsahuje dané body $(x_1, y_1), \dots, (x_n, y_n) \in \mathbb{R}^2$. Najděte vzorce pro délky a, b stran tohoto obdélníka (musí to být opravdu vzorce, nikoliv jen slovní popis jak spočítat a, b).

Rozpadá se na dvě nezávislé úlohy, pro každou souřadnici jednu: $a = \max_i x_i - \min_i x_i$, $b = \max_i y_i - \min_i y_i$

3. (2b) Napište vzorec pro gradient funkce $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{y} + \mathbf{b}^T \mathbf{x} + \|\mathbf{x} - \mathbf{y}\|_2$, kde $\mathbf{A} \in \mathbb{R}^{n \times n}$ a $\mathbf{b}, \mathbf{y} \in \mathbb{R}^n$ jsou dány.

Nejdřív spočteme totální derivaci. K tomu musíme členy, které jsou lineární funkce vektoru \mathbf{x} , dát do tvaru $\mathbf{a}^T \mathbf{x}$ kde \mathbf{a} je nějaký vektor: derivace tohoto členu bude pak \mathbf{a}^T . To se vlastně týká jen členu $\mathbf{x}^T \mathbf{A} \mathbf{y}$, který tedy ekvivalentně napišeme jako $(\mathbf{A} \mathbf{y})^T \mathbf{x}$. Pro derivaci posledního členu použijeme pravidlo pro derivaci eukleidovské normy a řetízkové pravidlo. Totální derivace tedy je $f'(\mathbf{x}) = (\mathbf{A} \mathbf{y})^T + \mathbf{b}^T + (\mathbf{x} - \mathbf{y})^T / \|\mathbf{x} - \mathbf{y}\|_2$. Gradient bude její transpozice, tedy $\nabla f(\mathbf{x}) = \mathbf{A} \mathbf{y} + \mathbf{b} + (\mathbf{x} - \mathbf{y}) / \|\mathbf{x} - \mathbf{y}\|_2$.

Všimněte si: \mathbf{y} není proměnná (jak by se mohlo zdát), ale konstanta, podle něj tedy nederivujeme.

4. (3b) Máme funkci $f(\mathbf{x}) = \mathbf{x}^T (\mathbf{A} \mathbf{x} + \mathbf{b})$, kde matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ (ne nutně symetrická) a vektor $\mathbf{b} \in \mathbb{R}^n$ jsou dány. Napište postup (posloupnost kroků), jak na základě vlastností \mathbf{A} a \mathbf{b} najít na množině \mathbb{R}^n všechny extrémy této funkce a rozhodnout, o jaký extrém se jedná (minima/maxima, lokální/globální).

Roznásobíme závorku, což dá $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x}$. To je kvadratická funkce, vše co potřebujete je tedy v §6.4 skript. Nebo úlohu můžeme vidět jako vyšetřování volných extrémů diferencovatelné funkce, o čemž je §10.1. Možný postup:

1 Hodnota kvadratické formy $\mathbf{x}^T \mathbf{A} \mathbf{x}$ závisí jen na symetrické části matice \mathbf{A} , pro pohodlí tedy budeme nadále používat jen tuto symetrickou část $\mathbf{S} = (\mathbf{A} + \mathbf{A}^T)/2$. Tedy vyšetřujeme funkci $f(\mathbf{x}) = \mathbf{x}^T \mathbf{S} \mathbf{x} + \mathbf{b}^T \mathbf{x}$.

2 Najdeme stacionární bod/body: položíme derivaci rovnou nule, $\mathbf{0}^T = \frac{d}{d\mathbf{x}}(\mathbf{x}^T \mathbf{S} \mathbf{x} + \mathbf{b}^T \mathbf{x}) = 2\mathbf{x}^T \mathbf{S} + \mathbf{b}^T$, po transpozici a úpravě $\mathbf{S} \mathbf{x} = -\mathbf{b}/2$ (tuto soustavu lze odvodit i doplněním na čtverec dle §6.4.1, což je o něco složitější). Řešení této soustavy jsou stacionární body funkce. Jestliže funkce nemá ani jeden stacionární bod (tj. $\mathbf{b} \notin \text{rng } \mathbf{S}$), pak funkce nemá extrém. Jestliže má jeden nebo více (tj. $\mathbf{b} \in \text{rng } \mathbf{S}$), extrémy může mít nebo nemusí – to zjistíme z definitnosti matice \mathbf{S} dále.

3 Jestliže má fce alespoň jeden stacionární bod, určíme definitnost matice \mathbf{S} (libovolnou metodou, tedy se nám ale hodí, že matice je symetrická). Jestliže \mathbf{S} je pozitivně [negativně] semidefinitní, tak všechny stacionární body jsou globální minima [maxima] funkce. Je-li indefinitní, funkce extrémy nemá (je neomezená shora i zdola).

5. Je dána funkce $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ předpisem $f(x, y) = x^2 + 4xy + y^2$.

(a) (1b) Najděte symetrickou matici \mathbf{A} tak aby $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$, kde $\mathbf{x} = (x, y) \in \mathbb{R}^2$. $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

(b) (2b) Rozhodněte, zda má f v počátku maximum, minimum nebo nic z toho. Odpověď odůvodněte.
Matice \mathbf{A} je indefinitní (vl. čísla jsou $-1, 3$), počátek je tedy její sedlový bod.

(c) (3b) Najděte ortogonální matici \mathbf{U} a konstanty a, b tak, aby pro každé $x, y \in \mathbb{R}$ platilo $f(x, y) = au^2 + bv^2$, kde $\begin{bmatrix} u \\ v \end{bmatrix} = \mathbf{U} \begin{bmatrix} x \\ y \end{bmatrix}$.

Toto je o interpretaci spektrálního rozkladu symetrické matice (vzorec (6.18) ve skriptech a text pod ním): $a = -1$ a $b = 3$ jsou vl. čísla matice \mathbf{A} , sloupce \mathbf{U} jsou normalizované vl. vektory $1/\sqrt{2}(-1, 1)$ a $1/\sqrt{2}(1, 1)$.

6. (3b) Máme dvě kuželosečky $\{ \mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x}^T \mathbf{A}\mathbf{x} = 1 \}$ a $\{ \mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x}^T \mathbf{B}\mathbf{x} = 1 \}$, kde $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{2 \times 2}$ jsou dané symetrické regulární matice. Hledáme průsečík těchto kuželoseček Newtonovou metodou. Napište iteraci metody (tj. vzorec, jak se spočítá příští odhad průsečíku z aktuálního odhadu).

Řešíme nelineární soustavu $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ kde zobrazení $\mathbf{g}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ a jeho derivace jsou

$$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} \mathbf{x}^T \mathbf{A}\mathbf{x} - 1 \\ \mathbf{x}^T \mathbf{B}\mathbf{x} - 1 \end{bmatrix} \in \mathbb{R}^2, \quad \mathbf{g}'(\mathbf{x}) = 2 \begin{bmatrix} \mathbf{x}^T \mathbf{A} \\ \mathbf{x}^T \mathbf{B} \end{bmatrix} \in \mathbb{R}^{2 \times 2}.$$

Iterace pak bude známý vzorec $\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{g}'(\mathbf{x}_k)^{-1} \mathbf{g}(\mathbf{x}_k)$.

7. (3b) Jsou dány obdélníková úzká matice \mathbf{C} , obdélníková široká matice \mathbf{A} a vektor \mathbf{b} . Minimalizujeme výraz $\|\mathbf{Cx}\|_2$ za podmínky $\mathbf{Ax} = \mathbf{b}$ přes proměnné $\mathbf{x} \in \mathbb{R}^n$. Najděte řešení úlohy; toto řešení musí být jediný vzorec (používající maticové operace) pro optimální \mathbf{x} . Použijte metodu Lagrangeových multiplikátorů. Předpokládejte, že inverze všech čtvercových matic, které budete potřebovat, existují.

Je důležité místo výrazu $\|\mathbf{Cx}\|_2$ minimalizovat jeho čtverec $\|\mathbf{Cx}\|_2^2 = \mathbf{x}^T \mathbf{C}^T \mathbf{C} \mathbf{x}$ (což je pouhá kvadratická forma), protože odmocnina by nám to velmi zkomplikovala (stejný trik jsme použili opakováně v metodě nejmenších čtverců). Ještě lépe je (ale to už není kritické) minimalizovat výraz $\frac{1}{2} \|\mathbf{Cx}\|_2^2$, aby nám derivace vyšla bez dvojkdy (to se taky ve skriptech použilo).

Lagr. funkce $L(\mathbf{x}, \boldsymbol{\lambda}) = \frac{1}{2} \|\mathbf{Cx}\|_2^2 + \boldsymbol{\lambda}^T (\mathbf{b} - \mathbf{Ax})$, podmínky rvního řádu jsou

$$\begin{aligned} L_{\mathbf{x}}(\mathbf{x}, \boldsymbol{\lambda})^T &= \mathbf{C}^T \mathbf{C} \mathbf{x} - \mathbf{A}^T \boldsymbol{\lambda} \mathbf{A} = 0 \\ \mathbf{Ax} &= \mathbf{b} \end{aligned}$$

Tuto soustavu řešíme pro proměnné $(\mathbf{x}, \boldsymbol{\lambda})$. Z první rovnice máme $\mathbf{x} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{A}^T \boldsymbol{\lambda}$ (kde inverze čtvercové matice $\mathbf{C}^T \mathbf{C}$ dle předpokladu existuje), dosazením do druhé máme $\mathbf{A}(\mathbf{C}^T \mathbf{C})^{-1} \mathbf{A}^T \boldsymbol{\lambda} = \mathbf{b}$. Z toho dostaneme $\boldsymbol{\lambda}$ a dosadíme do rovnice výše, výsledek $\mathbf{x} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{A}^T [\mathbf{A}(\mathbf{C}^T \mathbf{C})^{-1} \mathbf{A}^T]^{-1} \mathbf{b}$.

8. Máme množinu $\{ (x, y) \in \mathbb{R}^2 \mid \min\{x^2 + y^2, (x-1)^2 + y^2\} \leq 1 \}$.

- (a) (3b) Načrtněte množinu (její hranici plnou čarou, její vnitřek vyšrafujte).

Nápověda: nepopleťte si minimum s maximem.

Zde je klíčové si uvědomit, že tvrzení $\min\{a, b\} \leq c$ (pro $a, b, c \in \mathbb{R}$) je ekvivalentní tvrzení $(a \leq c) \vee (b \leq c)$. (Kdo to nevidí, ujasní si to např. tak, že si různé trojice čísel a, b, c kreslí na číselné osu. Pozor: neplést s tvrzením $\max\{a, b\} \leq c$, které je ekvivalentní tvrzení $(a \leq c) \wedge (b \leq c)$.) Tedy libovolný bod $(x, y) \in \mathbb{R}^2$ patří do množiny, jestliže splňuje $x^2 + y^2 \leq 1$ nebo $(x-1)^2 + y^2 \leq 1$ (v nevylučovacím smyslu, tedy klidně i obojí). Tedy množina je sjednocení dvou kruhů s poloměrem 1 a středy $(0, 0)$ a $(1, 0)$.

- (b) (3b) Je množina konvexní? Svou odpověď dokažte algebraicky z definice konvexní množiny (neformální argument pomocí obrázku nestačí).

Z obrázku ihned vidíme, že sjednocení uvedených kruhů konvexní není. K algebraickému důkazu stačí najít dva body (x, y) a (x', y') a skalár $0 < \alpha < 1$ tak, aby oba body byly v množině, ale bod $(1-\alpha)(x, y) + \alpha(x', y')$ v ní nebyl. S pomocí obrázku můžeme zvolit např. $(x, y) = (0, 1)$ a $(x', y') = (1, 1)$ a $\alpha = \frac{1}{2}$, tedy bod $(1-\alpha)(x, y) + \alpha(x', y') = (\frac{1}{2}, 1)$ je v polovině jejich spojnice. Tento bod v množině není, což je patrné z obrázku. Musíme to ale taky dokázat algebraicky: pro bod $(x, y) = (\frac{1}{2}, 1)$ podmínka $\min\{x^2 + y^2, (x-1)^2 + y^2\} = \min\{\frac{1}{4} + 1, \frac{1}{4} + 1\} = \frac{5}{4} \leq 1$ není splněna.

9. (4b) Minimalizujeme funkci $f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_\infty + \|\mathbf{x}\|_1$ na množině \mathbb{R}^n . Transformujte tuto úlohu na lineární program.

Skoro stejně jako Cvičení 12.4h ve skriptech. Je třeba zkombinovat dva obraty, které jsme použili při převodu minimalizace max-normy a 1-normy na lineární program (skripta §12.4.2). Výsledek (zapsán ve vektorové formě): $\min\{y + \mathbf{1}^T \mathbf{z} \mid -y\mathbf{1} \leq \mathbf{Ax} - \mathbf{b} \leq y\mathbf{1}, \mathbf{x} \leq \mathbf{z} \leq \mathbf{x}, \mathbf{x} \in \mathbb{R}^n, y \in \mathbb{R}, \mathbf{z} \in \mathbb{R}^n\}$

10. Máme úlohu $\min\{\mathbf{b}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{a}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}\}$, kde $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ jsou dané vektory \mathbf{s} nezápornými složkami.

- (a) **(2b)** Pro jaká \mathbf{a}, \mathbf{b} má úloha optimální řešení? Odpověď' podrobně odůvodněte.

Úloha LP má (aspoň jedno) optimální řešení, právě když je přípustná (tj. množina příustných řešení je neprázdná) a omezená (tj. účelová funkce je na množině příp. řešení zdola omezená).

To první platí právě když $\mathbf{a} \neq \mathbf{0}$ (tj. aspoň jedno číslo a_i je nenulové), jinak by podmínka $\mathbf{a}^T \mathbf{x} = 1$ nešla splnit.

Je-li $\mathbf{a} \neq \mathbf{0}$, pak je účelová funkce na množině přípustných řešení vždy zdola omezená, protože pro každé $\mathbf{b} \geq \mathbf{0}$ a $\mathbf{x} \geq \mathbf{0}$ je vždy $\mathbf{b}^T \mathbf{x} \geq 0$.

- (b) **(3b)** Má-li úloha optimální řešení, najdete ho úvahou a popište, jak se spočítá. Svoji úvahu vysvětlete.

Neřeklo se, že musíme najít všechna optimální řešení, stačí tedy nalézt jen jedno libovolné. Úlohu vyřešíme tak, že ji jednoduchou substitucí proměnných převedeme na úlohu, kterou dobře známe.

Nejdříve si všimneme, že když pro nějaké i je $a_i = 0$, tak proměnná x_i v podmínce $\mathbf{a}^T \mathbf{x} = 1$ nevystupuje a proto v optimu bude $x_i = 0$, protože zvětšením x_i by se úč. fce zhoršila (jestliže $b_i \geq 0$) nebo aspoň nezlepšila (pokud $b_i = 0$). Tedy můžeme dále uvažovat jen ty indexy i pro které $a_i > 0$. Dále tedy předpokládáme, že jsme z vektorů \mathbf{a}, \mathbf{b} již odstranily složky, pro které je $a_i = 0$ (a tedy máme $a_i > 0$ pro všechna i).

Ted' můžeme udělat substituci $x_i = y_i/a_i$, tj. $y_i = a_i x_i$. Úloha se tím změní na ekvivalentní úlohu $\min\{\mathbf{c}^T \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^n, \mathbf{1}^T \mathbf{y} = 1, \mathbf{y} \geq \mathbf{0}\}$, kde tedy $c_i = b_i/a_i$. Tuto úlohu dobře známe (její opt. řešení bude mít složky $y_{i^*} = 1$ pro libovolné $i^* \in \arg\min_i c_i$ a $y_i = 0$ pro všechna $i \neq i^*$).

Poznamenejme (i když se to v zadání nechtělo), že opt. hodnota úlohy jde napsat vzorcem jako $\min_{i \mid a_i \neq 0} b_i/a_i$.

- (c) **(3b)** K úloze napište duální úlohu.

To je snadné dle návodu ve skriptech (zač. kapitoly 15). Na levé straně napíšeme primár a na pravé duál:

$$\begin{array}{ll} \min & \mathbf{b}^T \mathbf{x} \\ \text{z.p.} & \mathbf{a}^T \mathbf{x} = 1 \\ & \mathbf{x} \geq \mathbf{0} \end{array} \quad \begin{array}{ll} \max & z \\ \text{z.p.} & z \in \mathbb{R} \\ & \mathbf{a} z \leq \mathbf{b} \end{array}$$

Pozn. (nebylo v zadání): všimněte si, že optimální hodnotu duálu najdeme úvahou snadněji než opt. hodnotu primáru: hledáme největší číslo z splňující $a_i z \leq b_i$ pro všechna $i = 1, \dots, n$, kde dle předpokladu $a_i, b_i \geq 0$.