

Optimalizace

Elektronická skripta předmětu A4B33OPT.

Text může být v průběhu semestru měněn, toto je verze z **21. října 2016**.

Upozornění na chyby posílejte na werner@cmp.felk.cvut.cz.

Tomáš Werner



Katedra kybernetiky
Fakulta elektrotechnická
České vysoké učení technické

Obsah

| | | |
|----------|---|-----------|
| 1 | Značení a základní pojmy | 6 |
| 1.1 | Matematické značení | 6 |
| 1.1.1 | Množiny | 6 |
| 1.1.2 | Zobrazení | 7 |
| 1.1.3 | Funkce a zobrazení více reálných proměnných | 7 |
| 1.1.4 | Extrémy funkce na množině | 8 |
| 1.2 | Obecná úloha spojité optimalizace | 9 |
| 1.3 | Cvičení | 11 |
| 2 | Maticová algebra | 13 |
| 2.1 | Operace s maticemi | 13 |
| 2.2 | Transpozice a symetrie | 14 |
| 2.3 | Stopa | 14 |
| 2.4 | Inverze | 15 |
| 2.5 | Determinant | 15 |
| 2.6 | Matice s jedním sloupcem nebo jedním řádkem | 15 |
| 2.7 | Matice sestavené z bloků | 16 |
| 2.8 | Hrubé chyby při práci s maticemi | 17 |
| 2.9 | Cvičení | 18 |
| 3 | Linearita | 21 |
| 3.1 | Lineární podprostory | 21 |
| 3.2 | Lineární zobrazení | 22 |
| 3.3 | Prostor obrazů a hodnost matice | 23 |
| 3.4 | Nulový prostor matice | 25 |
| 3.5 | Afinní podprostor a zobrazení | 26 |
| 3.6 | Cvičení | 29 |
| 4 | Ortogonalita | 32 |
| 4.1 | Standardní skalární součin | 32 |
| 4.2 | Ortogonální podprostory | 32 |
| 4.2.1 | Čtyři podprostory generované maticí | 33 |
| 4.3 | Ortonormální množina vektorů | 34 |
| 4.4 | Ortogonální projekce | 35 |
| 4.5 | Gramm-Schmidtova ortonormalizace | 37 |
| 4.5.1 | QR rozklad | 38 |
| 4.6 | Skalární součin a norma matic | 38 |

| | | | | | |
|----------|--|-----------|-----------|--|------------|
| 4.7 | Cvičení | 39 | 9.4.3 | Podmínky druhého řádu | 90 |
| 5 | Nehomogenní lineární soustavy | 42 | 9.5 | Cvičení | 91 |
| 5.1 | Přibližné řešení ve smyslu nejmenších čtverců | 42 | 10 | Iterační algoritmy na volné lokální extrémy | 96 |
| 5.1.1 | Řešení pomocí QR rozkladu | 44 | 10.1 | Sestupné metody | 96 |
| 5.1.2 | Lineární regrese | 45 | 10.2 | Gradientní metoda | 97 |
| 5.2 | Řešení s nejmenší normou | 45 | 10.2.1 | (*) Závislost na lineární transformaci souřadnic | 97 |
| 5.3 | Cvičení | 47 | 10.3 | Newtonova metoda | 97 |
| 6 | Spektrální rozklad a kvadratické funkce | 50 | 10.3.1 | Použití na soustavy nelineárních rovnic | 98 |
| 6.1 | Vlastní čísla a vektory | 50 | 10.3.2 | Použití na minimalizaci funkce | 99 |
| 6.1.1 | Spektrální rozklad | 51 | 10.4 | Nelineární metoda nejmenších čtverců | 100 |
| 6.2 | Kvadratická forma | 52 | 10.4.1 | Gauss-Newtonova metoda | 100 |
| 6.3 | Kvadratická funkce | 54 | 10.4.2 | Rozdíl proti Newtonově metodě | 102 |
| 6.3.1 | Kvadrika | 56 | 10.4.3 | Levenberg-Marquardtova metoda | 102 |
| 6.4 | Spektrální rozklad jako optimalizační úloha | 56 | 10.4.4 | Statistické odůvodnění kritéria nejmenších čtverců | 102 |
| 6.4.1 | Optimální proložení bodů podprostorem | 57 | 10.5 | Cvičení | 103 |
| 6.4.2 | Nejbližší matice nižší hodnosti | 58 | 11 | Lineární programování | 105 |
| 6.4.3 | Přeurčené homogenní lineární soustavy | 59 | 11.1 | Speciální tvary úloh LP | 106 |
| 6.5 | Cvičení | 60 | 11.1.1 | Po částech afinní funkce | 107 |
| 7 | Singulární rozklad | 64 | 11.2 | Některé aplikace LP | 108 |
| 7.1 | Rozšířené formy SVD | 65 | 11.2.1 | Optimální výrobní program | 108 |
| 7.2 | Nejbližší matice nižší hodnosti | 66 | 11.2.2 | Směšovací (dietní) problém | 109 |
| 7.3 | Pseudoinverze obecné matice | 67 | 11.2.3 | Dopravní problém | 109 |
| 7.3.1 | Pseudoinverze z regularizace | 67 | 11.3 | Použití na nehomogenní lineární soustavy | 110 |
| 7.4 | (*) Nejbližší ortogonální matice | 69 | 11.3.1 | Vektorové normy | 110 |
| 7.5 | Cvičení | 69 | 11.3.2 | Přibližné řešení přeurčených soustav | 111 |
| 8 | Nelineární funkce a zobrazení | 71 | 11.3.3 | Lineární regrese | 111 |
| 8.1 | Spojitosť | 72 | 11.4 | Cvičení | 113 |
| 8.2 | Parciální derivace | 72 | 12 | Konvexní množiny a mnohostěny | 117 |
| 8.3 | Totální derivace | 73 | 12.1 | Konvexní množiny | 117 |
| 8.3.1 | Derivace složeného zobrazení | 74 | 12.2 | Čtyři kombinace a čtyři obaly | 118 |
| 8.3.2 | Derivace maticových výrazů | 76 | 12.3 | Konvexní mnohostěny | 119 |
| 8.4 | Směrová derivace | 77 | 12.3.1 | Stěny konvexního mnohostěnu | 120 |
| 8.5 | Gradient | 78 | 12.3.2 | Jak byste vypsali všechny vrcholy konvexního mnohostěnu? | 121 |
| 8.6 | Parciální derivace druhého řádu | 79 | 12.3.3 | Dvě reprezentace konvexního mnohostěnu | 121 |
| 8.7 | Taylorův polynom | 79 | 12.4 | Cvičení | 122 |
| 8.8 | Cvičení | 81 | 13 | Simplexová metoda | 124 |
| 9 | Analytické podmínky na lokální extrémy | 83 | 13.1 | Stavební kameny algoritmu | 125 |
| 9.1 | Vlastnosti bodu vzhledem k podmnožině \mathbb{R}^n | 83 | 13.1.1 | Přechod k sousední standardní bázi | 125 |
| 9.2 | Lokální extrémy | 84 | 13.1.2 | Kdy je sousední báze řešení přípustné? | 126 |
| 9.3 | Volné lokální extrémy | 84 | 13.1.3 | Co když je celý sloupec nekladný? | 127 |
| 9.4 | Lokální extrémy vázané rovnostmi | 86 | 13.1.4 | Ekvivalentní úpravy účelového řádku | 127 |
| 9.4.1 | Tečný a ortogonální prostor k povrchu | 87 | 13.1.5 | Co udělá přechod k sousední bázi s účelovou funkcí? | 127 |
| 9.4.2 | Podmínky prvního řádu | 88 | 13.2 | Základní algoritmus | 128 |
| | | | 13.3 | Inicializace algoritmu | 130 |

| | |
|--|------------|
| 13.3.1 Dvoufázová simplexová metoda | 131 |
| 13.4 Cvičení | 133 |
| 14 Dualita v lineárním programování | 135 |
| 14.1 Konstrukce duální úlohy | 135 |
| 14.2 Věty o dualitě | 136 |
| 14.3 Příklady na konstrukci a interpretaci duálních úloh | 139 |
| 14.4 Cvičení | 142 |
| 15 Konvexní funkce | 144 |
| 15.1 Konvexita diferencovatelných funkcí | 146 |
| 15.2 Vztah konvexní funkce a konvexní množiny | 147 |
| 15.3 Operace zachovávající konvexitu funkcí | 148 |
| 15.4 Cvičení | 151 |
| 16 Konvexní optimalizační úlohy | 153 |
| 16.1 Ekvivalentní transformace úlohy | 154 |
| 16.2 Třídy konvexních optimalizačních úloh | 155 |
| 16.2.1 Lineární programování (LP) | 155 |
| 16.2.2 Kvadratické programování (QP) | 156 |
| 16.2.3 Kvadratické programování s kvadratickými omezeními (QCQP) | 157 |
| 16.2.4 Programování na kuželu druhého řádu (SOCP) | 157 |
| 16.2.5 Semidefinitní programování (SDP) | 158 |
| 16.3 Cvičení | 158 |
| 17 Lagrangeova dualita | 161 |
| 17.1 Minimální nerovnost | 161 |
| 17.2 Lagrangeova duální úloha | 162 |
| 17.3 Silná dualita | 163 |
| 17.4 Příklady | 164 |
| 18 Vícekriteriální optimalizace | 166 |
| 18.1 Uspořádání na množině | 166 |
| 18.2 Úlohy vícekriteriální optimalizace | 167 |

Kapitola 1

Značení a základní pojmy

1.1 Matematické značení

Pokud potkáte ve skriptech slovo vysázené **tučně**, jde o nově definovaný pojem, který máte pochopit a zapamatovat si. Slova vysázená *kurzívou* znamenají buď zdůraznění, nebo nový avšak všeobecně známý pojem. Odstavce, věty, důkazy, příklady a cvičení označené hvězdičkou (\star) jsou rozšiřující (a tedy obtížnější) a není nezbytné umět je ke zkoušce.

Zopakujeme nejdříve matematické značení, které se používá v celých skriptech a které by čtenář měl bezpečně ovládat.

1.1.1 Množiny

Názvy množin budeme psát velkými skloněnými písmeny, např. A nebo X . Budeme používat standardní množinové značení:

| | |
|-------------------------------|--|
| $\{a_1, \dots, a_n\}$ | množina s prvky a_1, \dots, a_n |
| $a \in A$ | prvek a patří do množiny A (neboli a je prvkem A) |
| $A \subseteq B$ | množina A je podmnožinou množiny B , tj. každý prvek z A patří do B |
| $A = B$ | množina A je rovna množině B , platí zároveň $A \subseteq B$ a $B \subseteq A$ |
| $\{a \in A \mid \varphi(a)\}$ | množina prvků a z množiny A , které splňují logický výrok $\varphi(a)$ |
| $A \cup B$ | sjednocení množin, množina $\{a \mid a \in A \text{ nebo } a \in B\}$ |
| $A \cap B$ | průnik množin, množina $\{a \mid a \in A, a \in B\}$ |
| $A \setminus B$ | rozdíl množin, množina $\{a \mid a \in A, a \notin B\}$ |
| (a_1, \dots, a_n) | uspořádaná n -tice prvků a_1, \dots, a_n |
| $A \times B$ | kartézský součin množin, množina $\{(a, b) \mid a \in A, b \in B\}$ |
| A^n | kartézský součin n stejných množin, $A^n = A \times \dots \times A$ (n -krát) |
| \emptyset | prázdná množina |

Číselné množiny budeme značit takto:

| | |
|-------------------|---|
| \mathbb{N} | množina přirozených čísel |
| \mathbb{Z} | množina celých čísel |
| \mathbb{Q} | množina racionálních čísel |
| \mathbb{R} | množina reálných čísel |
| \mathbb{R}_+ | množina nezáporných reálných čísel $\{x \in \mathbb{R} \mid x \geq 0\}$ |
| \mathbb{R}_{++} | množina kladných reálných čísel $\{x \in \mathbb{R} \mid x > 0\}$ |
| $[x_1, x_2]$ | uzavřený reálný interval $\{x \in \mathbb{R} \mid x_1 \leq x \leq x_2\}$ |
| (x_1, x_2) | otevřený reálný interval $\{x \in \mathbb{R} \mid x_1 < x < x_2\}$ |
| $[x_1, x_2)$ | polouzavřený reálný interval $\{x \in \mathbb{R} \mid x_1 \leq x < x_2\}$ |
| \mathbb{C} | množina komplexních čísel |

1.1.2 Zobrazení

Zobrazení z množiny A do množiny B značíme

$$f: A \rightarrow B \quad (1.1)$$

nebo (méně často) $A \xrightarrow{f} B$. Zobrazení si můžeme představit¹ jako ‘černou skříňku’, která každému prvku $a \in A$ přiřadí právě jeden prvek $b = f(a) \in B$. I když ‘zobrazení’ (*mapping*, *map*) znamená přesně totéž jako ‘funkce’ (*function*), slovo ‘funkce’ se obvykle používá pouze pro zobrazení do číselných množin (tedy $B = \mathbb{R}, \mathbb{Z}, \mathbb{C}$ apod.). Zobrazení se nazývá:

- *injektivní* (neboli prosté) pokud každý vzor má jiný obraz, tj. $(f(a) = f(a')) \Rightarrow (a = a')$;
- *surjektivní* (neboli A na B) pokud každý obraz má aspoň jeden vzor, tj. $f(A) = B$, tj. pro každé $b \in B$ existuje $a \in A$ tak, že $b = f(a)$;
- *bijektivní* (neboli vzájemně jednoznačné) pokud je zároveň injektivní a surjektivní.

Obraz množiny $A' \subseteq A$ v zobrazení $f: A \rightarrow B$ značíme

$$f(A') = \{f(a) \mid a \in A'\}. \quad (1.2)$$

Např. je-li $A' = \{1, 3, 4, -1\} \subseteq \mathbb{Z}$ a zobrazení $f: \mathbb{Z} \rightarrow \mathbb{Z}$ je definované předpisem $f(a) = a^2$, je $f(A') = \{1, 9, 16\}$.

Je-li $A' = \{a \in A \mid \varphi(a)\}$ a $f: A \rightarrow B$, používáme zkratku

$$f(A') = \{f(a) \mid a \in A, \varphi(a)\} = \{b \in B \mid b = f(a), a \in A, \varphi(a)\} \quad (1.3)$$

nebo pouze $\{f(a) \mid \varphi(a)\}$, pokud je A jasné z kontextu. Např. $\{x^2 \mid x \in \mathbb{R}, -1 < x < 1\}$ je polouzavřený interval $[0, 1)$.

Mějme dvě zobrazení $f: A \rightarrow B$ a $g: B \rightarrow C$, neboli $A \xrightarrow{f} B \xrightarrow{g} C$. *Složení* zobrazení f a g je zobrazení $g \circ f: A \rightarrow C$ definované jako $(g \circ f)(a) = g(f(a))$ pro každé $a \in A$.

1.1.3 Funkce a zobrazení více reálných proměnných

Uspořádané n -tici $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ reálných čísel říkáme (n -rozměrný) **vektor**. Zápis

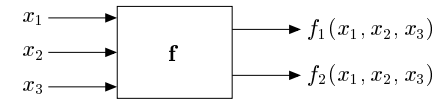
$$\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m \quad (1.4)$$

¹ Přesná definice je následující: zobrazení $f: A \rightarrow B$ je podmnožina kartézského součinu $A \times B$ (tedy *relace*) taková, že $(a, b) \in f$ a $(a, b') \in f$ implikuje $b = b'$.

označuje zobrazení, které vektoru $\mathbf{x} \in \mathbb{R}^n$ přiřadí vektor

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, \dots, x_n) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x})) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)) \in \mathbb{R}^m,$$

kde $f_1, \dots, f_m: \mathbb{R}^n \rightarrow \mathbb{R}$ jsou *složky* zobrazení. Píšeme také $\mathbf{f} = (f_1, \dots, f_m)$. Obrázek ilustruje zobrazení $\mathbf{f}: \mathbb{R}^3 \rightarrow \mathbb{R}^2$:



Pro $m = 1$ jsou hodnotami zobrazení skaláry a budeme psát jeho jméno kurzívou, f . Pro $m > 1$ jsou hodnotami zobrazení vektory a proto jméno budeme psát tučně, \mathbf{f} . I když slova ‘funkce’ a ‘zobrazení’ znamenají jedno a to samé, je často zvykem pro $m = 1$ mluvit o *funkci* a pro $m > 1$ o *zobrazení*.

1.1.4 Extrémy funkce na množině

Mějme funkci $f: X' \rightarrow \mathbb{R}$ a množinu $X \subseteq X'$. Necht' $x \in X$ je takové, že $f(x) \leq f(x')$ pro všechna $x' \in X$. Pak x nazveme *minimum funkce f* na množině X , nebo také říkáme, že funkce f *nabývá minima* na množině X v prvku x . Číslo $f(x)$ nazýváme *minimální hodnotou* funkce f na množině X a píšeme

$$f(x) = \min_{x' \in X} f(x'). \quad (1.5)$$

Pokud navíc je $f(x) < f(x')$ pro všechna $x' \in X \setminus \{x\}$, mluvíme o *ostrém minimu*. Množinu všech prvků, ve kterých f na X nabývá minima, značíme

$$\operatorname{argmin}_{x \in X} f(x) = \{x \in X \mid f(x) \leq f(x') \forall x' \in X\} \subseteq X. \quad (1.6)$$

Podobně definujeme maximum funkce na množině. Minima a maxima funkce se souhrnně nazývají její *extrémy* nebo *optima*. Pokud odkaz na množinu X chybí, myslí se $X = X'$.

Je užitečné se k minimumu funkce na množině postavit poněkud abstraktněji. Necht' $Y \subseteq \mathbb{R}$. Prvek $y \in Y$ nazveme *nejmenší prvek* (nebo také *minimální prvek*) množiny Y , jestliže $y \leq y'$ pro všechna $y' \in Y$. Nejmenší prvek značíme $\min Y$. Ne každá množina $Y \subseteq \mathbb{R}$ má nejmenší prvek (např. interval $(0, 1]$ ho nemá). Na druhou stranu, Y má nejvýše jeden minimální prvek.

Označme nyní

$$f(X) = \{f(x) \mid x \in X\} \subseteq \mathbb{R}$$

obraz množiny X funkcí f . Pokud množina $f(X)$ má nejmenší prvek, definujeme

$$\min_{x \in X} f(x) = \min\{f(x) \mid x \in X\} = \min f(X).$$

Funkce nemusí mít na množině minimum, což plyne toho, že ne každá množina $Y \subseteq \mathbb{R}$ má minimální prvek. Všimněte si, že v tom případě je množina (1.6) prázdná.

Příklad 1.1.

- Necht' $X' = X = [1, \infty)$ a $f(x) = 1/x$. Máme $f(X) = (0, 1]$. Ale množina $(0, 1]$ nemá minimální prvek, proto funkce f na množině X nemá minimum.
- $\min_{x \in \mathbb{R}} |x - 1| = \min\{|x - 1| \mid x \in \mathbb{R}\} = \min \mathbb{R}_+ = 0$, $\operatorname{argmin}_{x \in \mathbb{R}} |x - 1| = \{1\}$

- Necht' $f(x) = \max\{|x|, 1\}$. Pak $\operatorname{argmin}_{x \in \mathbb{R}} f(x) = [-1, 1]$.
- Necht' $(a_1, a_2, \dots, a_5) = (1, 2, 3, 2, 3)$. Pak² $\max_{i=1}^5 a_i = 3$, $\operatorname{argmax}_{i=1}^5 a_i = \{3, 5\}$. \square

1.2 Obecná úloha spojité optimalizace

Optimalizační úlohy se formulují jako hledání minima dané reálné funkce $f: X \rightarrow \mathbb{R}$ na dané množině X . Tato formulace je velmi obecná, neboť množina X může být zcela libovolná. Existují tři široké kategorie úloh:

- Pokud je množina X konečná (i když třeba velmi velká), mluvíme o *kombinatorické optimalizaci*. Její prvky mohou být např. cesty v grafu, konfigurace Rubikovy kostky, nebo textové řetězce konečné délky. Příkladem je nalezení nejkratší cesty v grafu nebo problém obchodního cestujícího.
- Pokud množina X obsahuje reálná čísla či reálné vektory (tedy $X \subseteq \mathbb{R}^n$), mluvíme o *spojité optimalizaci*. Příkladem je úloha lineárního programování.
- Pokud množina X obsahuje reálné funkce, mluvíme o *variačním počtu*. Příkladem je nalézt rovinnou křivku, která při dané délce obepíná co největší obsah.

Tento kurs se zabývá spojitou optimalizací. V *obecné úloze spojité optimalizace* je $X \subseteq \mathbb{R}^n$ množina řešení $(x_1, \dots, x_n) \in \mathbb{R}^n$ soustavy m nerovnic a l rovnic

$$g_i(x_1, \dots, x_n) \leq 0, \quad i = 1, \dots, m \quad (1.7a)$$

$$h_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, l \quad (1.7b)$$

pro dané funkce $g_1, \dots, g_m, h_1, \dots, h_l: \mathbb{R}^n \rightarrow \mathbb{R}$. Ve vektorovém značení píšeme

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0} \},$$

kde $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{h}: \mathbb{R}^n \rightarrow \mathbb{R}^l$ a $\mathbf{0}$ značí nulové vektory příslušné dimenze. Ve shodě s označením (1.3) se hledání minima funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině X může zapsat jako

$$\min \{ f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0} \}, \quad (1.8)$$

To je zvykem zapisovat také takto:

$$\begin{aligned} & \min && f(x_1, \dots, x_n) \\ & \text{za podmíněk} && g_i(x_1, \dots, x_n) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, l. \end{aligned} \quad (1.9)$$

Příklad 1.2. Pasterec vlastní 100 metrů pletiva a chce z něj udělat ohradu pro ovce o co největším obsahu. Ohrada bude mít tvar obdélníka, z něhož tři strany budou tvořeny plotem a zbylá strana řekou (ovce neumějí plavat).

² Místo $\max_{i=1}^5 a_i$ se obvykle píše $\max_{i=1, \dots, 5} a_i$. Používáme zde první způsob v analogii se značením $\sum_{i=1}^5 a_i$.

Označme strany obdélníka x, y . Řešíme úlohu

$$\begin{aligned} & \max && xy \\ & \text{za podmíněk} && 2x + y = 100 \\ & && x, y \geq 0 \end{aligned}$$

neboli

$$\max \{ xy \mid x \geq 0, y \geq 0, 2x + y = 100 \}.$$

Zde máme $n = 2$, $m = 2$, $l = 1$.

Tuto úlohu dokážeme snadno vyřešit. Z omezení $2x + y = 100$ máme $y = 100 - 2x$, tedy místo původní úlohy můžeme řešit ekvivalentní úlohu bez omezení

$$\max_{x \in \mathbb{R}} x(100 - 2x).$$

Maximum kvadratické funkce $x(100 - 2x)$ snadno najdeme prostředky analýzy funkcí jedné proměnné. Nabývá se v bodě $x = 25$, tedy $y = 100 - 2x = 50$. Tyto hodnoty jsou kladné, tedy podmínky $x, y \geq 0$ jsou automaticky splněny a nemuseli jsme je explicitně uvažovat. \square

Příklad 1.3. Hledejme dvojici nejbližších bodů v rovině, z nichž jeden je na kružnici se středem v počátku a jednotkovým poloměrem a druhý je ve čtverci se středem v bodě $(2, 2)$ a jednotkovou stranou. Úlohu lze samozřejmě řešit snadno úvahou. Napišme ji ale ve tvaru (1.9).

Bod (x_1, x_2) na kružnici splňuje $x_1^2 + x_2^2 = 1$. Bod (x_3, x_4) ve čtverci splňuje $-\frac{1}{2} \leq x_3 - 2 \leq \frac{1}{2}$, $-\frac{1}{2} \leq x_4 - 2 \leq \frac{1}{2}$. Máme $n = 4$, $m = 4$, $l = 1$, a

$$X = \{ (x_1, x_2, x_3, x_4) \mid x_1^2 + x_2^2 - 1 = 0, \frac{3}{2} - x_3 \leq 0, x_3 - \frac{5}{2} \leq 0, \frac{3}{2} - x_4 \leq 0, x_4 - \frac{5}{2} \leq 0 \}.$$

Řešíme úlohu

$$\begin{aligned} & \min && \sqrt{(x_1 - x_3)^2 + (x_2 - x_4)^2} \\ & \text{za podmíněk} && x_1^2 + x_2^2 - 1 = 0 \\ & && \frac{3}{2} - x_3 \leq 0 \\ & && x_3 - \frac{5}{2} \leq 0 \\ & && \frac{3}{2} - x_4 \leq 0 \\ & && x_4 - \frac{5}{2} \leq 0 \end{aligned} \quad \square$$

V matematické analýze se řešením úlohy (1.9) říká *extrémny funkce f vázané podmínkami (1.7)*. Pokud omezení chybí, mluví se o *volných extrémech* funkce f . V matematické optimalizaci se živilo poněkud odlišné názvosloví:

- Funkce f se nazývá *účelová* (také pokutová, cenová, kriteriální) funkce.
- Prvky množiny X se nazývají *přípustná řešení*, což je vlastně protimluv, protože nemusí být řešeními úlohy (1.9). Prvkům množiny $\operatorname{argmin}_{\mathbf{x} \in X} f(\mathbf{x})$ se pak říká *optimální řešení*.
- Rovnice a nerovnice (1.7) se nazývají *omezující podmínky*, krátce *omezení*.
- Omezení (1.7a) příp. (1.7b) se nazývají *omezení typu nerovnosti* příp. *typu rovnosti*. Pokud omezení chybí ($m = l = 0$), jedná se o optimalizaci *bez omezení*.
- Pokud je omezení typu nerovnosti $g_i(\mathbf{x}) \leq 0$ plněno s rovností, tedy $g_i(\mathbf{x}) = 0$, říkáme, že je v bodě \mathbf{x} *aktivní*.
- Pokud $X \neq \emptyset$, úloha se nazývá *přípustná*, v opačném případě ($X = \emptyset$) je *nepřípustná*.

- Pokud je množina $f(X)$ zdola omezená, úloha se nazývá *omezená*. V opačném případě (tedy pokud se účelová funkce při splněných omezeních může zlepšovat nade všechny meze) se úloha nazývá *neomezená*.

Všimněte si, že přípustná a omezená úloha ještě nemusí mít optimální řešení: uvažte třeba úlohu $\min\{-x \mid x \geq 0, y \geq 0, xy = 1\}$.

1.3 Cvičení

1.1. Vyřešte následující úlohy, přičemž slovní úlohy nejdříve formulujte ve tvaru (1.9). Stačí vám k tomu zdravý rozum a derivace funkcí jedné proměnné.

- $\min\{x^2 + y^2 \mid x > 0, y > 0, xy \geq 1\}$
- $\min\{(x-2)^2 + (y-1)^2 \mid x^2 \leq 1, y^2 \leq 1\}$
- Máte vyrobit papírovou krabici o objemu 72 litrů, jejíž délka je dvojnásobek její šířky. Jaké budou její rozměry, má-li se na ní spotřebovat co nejméně papíru? Tloušťka stěn je zanedbatelná.
- Jaké má rozměry válec s jednotkovým objemem a nejmenším povrchem?
- Najděte rozměry püllitru, na jehož výrobu je třeba co nejméně skla. Tloušťka stěn je zanedbatelná.
- Najděte obsah největšího obdélníka vepsaného do kružnice s poloměrem 1.
- Obdélník v rovině má jeden roh v počátku a druhý na křivce $y = x^2 + x^{-2}$, přičemž jeho strany jsou rovnoběžné se souřadnicovými osami. Pro jaké x bude jeho obsah minimální? Může být jeho obsah libovolně veliký?
- Najděte bod v rovině na parabole s rovnicí $y = x^2$ nejbliže bodu $(3, 0)$.
- Hektarová oblast obdélníkového tvaru se má obehnat ze tří stran živým plotem, který stojí 1000 korun na metr, a ze zbývajících stran obyčejným plotem, který stojí 500 korun na metr. Jaké budou rozměry oblasti při nejmenší ceně plotu?
- x, y jsou čísla v intervalu $[1, 5]$ takové, že jejich součet je 6. Najděte tato čísla tak, aby xy^2 bylo (a) co nejmenší a (b) co největší.
- Hledá se n -tice čísel $x_1, \dots, x_n \in \{-1, +1\}$ tak, že jejich součin je kladný a jejich součet minimální. Jako výsledek napište (co nejjednodušší) vzorec, který udává hodnotu tohoto minimálního součtu pro obecné n .
- Potkaní biatlon*. Potkan stojí na břehu kruhového jezírka o poloměru 1 a potřebuje se dostat na protilehlý bod břehu. Potkan plave rychlostí v_1 a běží rychlostí v_2 . Chce se do cíle dostat co nejrychleji, přičemž může běžet, plavat, nebo zvolit kombinaci obojího. Jakou dráhu zvolí? Strategie potkana může být různá pro různé hodnoty v_1 a v_2 , vyřešte pro všechny kombinace těchto hodnot.

1.2. Načrtněte následující množiny (proměnné x, y patří do \mathbb{R}):

- $[-1, 0] \times \{1\}$
- $\mathbb{Z} \times \mathbb{Z}$
- $\mathbb{R} \times \mathbb{Z}$
- $(\mathbb{R} \times \mathbb{Z}) \cup (\mathbb{Z} \times \mathbb{R})$
- $\{(x, y) \mid x^2 + y^2 = 1\} \times \mathbb{R}$
- $\{(x, y) \mid x > 0, y > 0, xy = 1\}$

- $\{(x, y) \mid \min\{x, y\} = 1\}$
- $\{1/x \mid x \geq 1\}$
- $\{1/x \mid |x| \geq 1\}$
- $\{e^{-x^2} \mid x \in \mathbb{R}\}$
- $\{x + y \mid x^2 + y^2 < 1\}$
- $\{x - y \mid x^2 + y^2 = 1\}$
- (*) $\{x + y \mid |x| \geq 1, |y| \geq 1\}$

1.3. Necht' X je libovolná množina a $f: X \rightarrow \mathbb{R}$. Necht' $g: \mathbb{R} \rightarrow \mathbb{R}$ je rostoucí funkce. Dokažte, že $\operatorname{argmin}_{x \in X} f(x) = \operatorname{argmin}_{x \in X} g(f(x))$.

Nápověda a řešení

- 1.2.h) $(0, 1]$
- 1.2.i) $[-1, 0) \cup (0, 1] = [-1, 1] \setminus \{0\}$
- 1.2.j) $(0, 1]$
- 1.2.k) $(-\sqrt{2}, \sqrt{2})$
- 1.2.l) $[-\sqrt{2}, \sqrt{2}]$

Kapitola 2

Maticová algebra

Reálná **matice** rozměru $m \times n$ je zobrazení $\{1, \dots, m\} \times \{1, \dots, n\} \rightarrow \mathbb{R}$. Toto zobrazení zapisujeme tabulkou

$$\mathbf{A} = [a_{ij}] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix},$$

kde a_{ij} jsou **prvky** matice. Množinu všech reálných matic rozměru $m \times n$ značíme $\mathbb{R}^{m \times n}$.

Budeme používat tyto názvy:

- Pro $m = n$ se matice nazývá **čtvercová** a pro $m \neq n$ **obdélníková**, přičemž pro $m < n$ je **široká** a pro $m > n$ je **úzká**.
- **Diagonální prvky** matice jsou prvky a_{11}, \dots, a_{pp} , kde $p = \min\{m, n\}$. Matice je **diagonální**, když všechny nediagonální prvky jsou nulové, tedy $a_{ij} = 0$ pro všechna $i \neq j$. Všimněte si, že diagonální matice nemusí být čtvercová. Čtvercovou ($m = n$) diagonální maticí značíme $\mathbf{A} = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$.
- **Nulová matice** má všechny prvky nulové. Značíme ji $\mathbf{0}_{m \times n}$ (pokud jsou rozměry jasné z kontextu, pak pouze $\mathbf{0}$).
- **Jednotková matice** je čtvercová diagonální, jejíž diagonální prvky jsou jedničky. Značíme ji \mathbf{I}_n (pokud jsou rozměry jasné z kontextu, pak pouze \mathbf{I}).
- **Horní [dolní] trojúhelníková matice** má $a_{ij} = 0$ pro všechna $i > j$ [$i < j$]. Všimněte si, že horní/dolní trojúhelníková matice nemusí být čtvercová.

2.1 Operace s maticemi

Na maticích jsou definovány následující operace:

- Součin skaláru¹ $\alpha \in \mathbb{R}$ a matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ je matice $\alpha \mathbf{A} = \mathbf{A} \alpha = [\alpha a_{ij}] \in \mathbb{R}^{m \times n}$.
- Součet matic $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ je matice $\mathbf{A} + \mathbf{B} = [a_{ij} + b_{ij}] \in \mathbb{R}^{m \times n}$.
- **Maticový součin** matic $\mathbf{A} \in \mathbb{R}^{m \times p}$ a $\mathbf{B} \in \mathbb{R}^{p \times n}$ je matice $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{m \times n}$ s prvky

$$c_{ij} = \sum_{k=1}^p a_{ik} b_{kj}. \quad (2.1)$$

¹ Slovo *skalár* v algebře reálných matic označuje reálné číslo. Přesněji, pohlížíme-li na množinu všech matic rozměru $m \times n$ jako na lineární prostor, jedná se o skaláry tohoto lineárního prostoru.

Všimněte si, že násobit lze jen matice, které mají vnitřní rozměr (p) stejný.

Vlastnosti maticového součinu:

- $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ a $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
- $\mathbf{A}\mathbf{I}_n = \mathbf{A} = \mathbf{I}_m \mathbf{A}$
- $(\alpha \mathbf{A})\mathbf{B} = \mathbf{A}(\alpha \mathbf{B}) = \alpha(\mathbf{AB})$

Obecně neplatí $\mathbf{AB} = \mathbf{BA}$ (maticový součin není komutativní!).

Poznamenejme, že výraz $\alpha \mathbf{A}$ pro $\alpha \in \mathbb{R}$ nelze považovat za maticový součin 'matice' α rozměru 1×1 a matice \mathbf{A} , protože vnitřní rozměr matic by byl obecně různý. Tedy násobením matice skalárem je jiná operace, než maticový součin.

Pro čtvercovou matici \mathbf{A} značí \mathbf{A}^k maticový součin k matic \mathbf{A} , tedy k -tou mocninou matice.

2.2 Transpozice a symetrie

Transpozici matice $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$ značíme $\mathbf{A}^T = [a_{ji}] \in \mathbb{R}^{n \times m}$. Vlastnosti transpozice:

- $(\alpha \mathbf{A})^T = \alpha \mathbf{A}^T$
- $(\mathbf{A}^T)^T = \mathbf{A}$
- $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
- $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$

Čtvercová matice se nazývá

- **symetrická**, když $\mathbf{A}^T = \mathbf{A}$, tj. $a_{ij} = a_{ji}$,
- **antisymetrická**, když $\mathbf{A}^T = -\mathbf{A}$, tj. $a_{ij} = -a_{ji}$ (z čehož plyne $a_{ii} = 0$).

2.3 Stopa

Stopa (angl. *trace*) čtvercové matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je součet jejích diagonálních prvků, značí se

$$\text{tr } \mathbf{A} = a_{11} + \cdots + a_{nn}. \quad (2.2)$$

Vlastnosti:

- $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr } \mathbf{A} + \text{tr } \mathbf{B}$
- $\text{tr}(\alpha \mathbf{A}) = \alpha \text{tr } \mathbf{A}$
- $\text{tr}(\mathbf{A}^T) = \text{tr } \mathbf{A}$
- $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ (kde $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{B} \in \mathbb{R}^{n \times m}$)

Z poslední rovnosti plyne např. $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{CAB})$, protože $\text{tr}(\mathbf{DC}) = \text{tr}(\mathbf{CD})$ kde $\mathbf{D} = \mathbf{AB}$. Podobně např. $\text{tr}(\mathbf{ABCD}) = \text{tr}(\mathbf{CDAB})$. Ale neplatí např. $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{CBA})$.

2.4 Inverze

Když platí

$$\mathbf{AB} = \mathbf{I}, \quad (2.3)$$

matice \mathbf{B} se nazývá **pravá inverze** matice \mathbf{A} a matice \mathbf{A} se nazývá **levá inverze** matice \mathbf{B} . Pravá či levá inverze nemusí existovat nebo nemusí být jediná. Pravá inverze matice existuje právě tehdy, má-li matice lineárně nezávislé řádky (viz Věta 3.3). Levá inverze matice existuje právě tehdy, má-li matice lineárně nezávislé sloupce (viz Věta 3.5).

Jestliže matice \mathbf{A} je čtvercová a má pravou inverzi, má zároveň i levou inverzi a obě inverze se rovnají a jsou jediné. Opravdu: je-li $\mathbf{AX} = \mathbf{I}$ a $\mathbf{YA} = \mathbf{I}$, pak $\mathbf{YAX} = \mathbf{Y} = \mathbf{X}$. Protože toto platí pro každou levou inverzi a každou pravou inverzi, jsou zároveň jediné. Pak mluvíme pouze o **inverzi** matice a značíme ji \mathbf{A}^{-1} . Matici, která má inverzi, nazýváme **regulární**, v opačném případě **singulární**. Vlastnosti inverze:

- $\mathbf{AA}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A}$
- $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
- $(\alpha\mathbf{A})^{-1} = \alpha^{-1}\mathbf{A}^{-1}$
- $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$, což krátce značíme \mathbf{A}^{-T} .

2.5 Determinant

Determinant je funkce $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ (tedy přiřazuje čtvercové matici skalár) definovaná jako

$$\det \mathbf{A} = \sum_{\sigma} \operatorname{sgn} \sigma \prod_{i=1}^n a_{i\sigma(i)}, \quad (2.4)$$

kde sčítáme přes všechny permutace n prvků $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$, přičemž $\operatorname{sgn} \sigma$ označuje znaménko permutace. Některé vlastnosti determinantu:

- Determinant je multilineární funkce sloupců matice, tj. je lineární funkcí libovolného sloupce, jsou-li všechny ostatní sloupce konstantní.
- Determinant je alternující funkce sloupců matice, tj. prohození dvou sousedních sloupců změni znaménko determinantu.
- $\det \mathbf{I} = 1$
- $\det \mathbf{A} = 0$ právě tehdy, když \mathbf{A} je singulární
- $\det \mathbf{A}^T = \det \mathbf{A}$
- $\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})$
- $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1}$ (plyne z předchozího pro $\mathbf{B} = \mathbf{A}^{-1}$)

2.6 Matice s jedním sloupcem nebo jedním řádkem

Matice s jedním sloupcem (tedy prvek $\mathbb{R}^{n \times 1}$) se také nazývá **sloupcový vektor**². Matice s jedním řádkem (tedy prvek $\mathbb{R}^{1 \times m}$) se také nazývá **řádkový vektor**.

² V lineární algebře má slovo *vektor* obecnější význam než v maticové algebře: znamená prvek lineárního prostoru (který se někdy také nazývá *vektorový prostor*).

Lineární prostor $\mathbb{R}^{n \times 1}$ všech matic s jediným sloupcem je ‘skoro stejný’ jako lineární prostor \mathbb{R}^n všech uspořádaných n -tic (x_1, \dots, x_n) . Proto je zvykem tyto prostory ztotožnit a bez upozornění přecházet mezi oběma významy. Prvkům

$$\mathbf{x} = \underbrace{(x_1, \dots, x_n)}_{\text{uspořádaná } n\text{-tice}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

matice $n \times 1$

tohoto prostoru budeme říkat krátce **vektory**. Jinak řečeno, slovem *vektor* (bez přívlastku) budeme rozumět *sloupcový vektor* nebo také uspořádanou n -tici čísel³.

Všimněme si případů, kdy se v maticovém součinu vyskytují vektory:

- Pro matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ a vektor $\mathbf{x} \in \mathbb{R}^n$, výraz \mathbf{Ax} je maticový součin matice $m \times n$ a matice $n \times 1$, což je (sloupcový) vektor délky m .
- Pro matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ a vektor $\mathbf{x} \in \mathbb{R}^m$, výraz $\mathbf{x}^T \mathbf{A}$ je maticový součin matice $1 \times m$ a matice $m \times n$, což je řádkový vektor délky n .
- Pro $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ je $\mathbf{x}^T \mathbf{y} = x_1 y_1 + \dots + x_n y_n$ je maticový součin řádkového vektoru \mathbf{x}^T a sloupcového vektoru \mathbf{y} , jehož výsledkem je skalár. Je to standardní skalární součin vektorů \mathbf{x} a \mathbf{y} (více si o něm řekneme §4.1).
- Pro $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ je \mathbf{xy}^T matice rozměru $m \times n$, které se někdy říká **vnější součin** vektorů \mathbf{x} a \mathbf{y} nebo **dyáda**.

Symbol $\mathbf{1}_n = (1, \dots, 1) \in \mathbb{R}^n$ značí (sloupcový) vektor s jedničkovými složkami. Pokud n plyne z kontextu, píšeme jen $\mathbf{1}$. Příklad: pro $\mathbf{x} \in \mathbb{R}^n$ je $\mathbf{1}^T \mathbf{x} = x_1 + \dots + x_n$.

Symbol $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^n$ (jednička na i -tém místě) značí i -tý (sloupcový) vektor standardní báze, kde dimenze n vektoru \mathbf{e}_i je určena kontextem. Standardní báze tvoří sloupce jednotkové matice, $[\mathbf{e}_1 \ \dots \ \mathbf{e}_n] = \mathbf{I}_n$.

2.7 Matice sestavené z bloků

Matici je možno sestavit z několika jejích **podmatic** (zvaných též **bloky**), např.

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \quad [\mathbf{A} \ \mathbf{B}], \quad \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}. \quad (2.5)$$

Rozměry jednotlivých bloků musí být slučitelné, např. v prvním příkladu musí mít matice \mathbf{A}, \mathbf{B} stejný počet řádků a matice \mathbf{A}, \mathbf{C} stejný počet sloupců. V posledním příkladu jsou rozměry jednotkové matice \mathbf{I} a nulové matice $\mathbf{0}$ určeny rozměry matic \mathbf{A} a \mathbf{D} .

Při násobení matic sestavených z bloků je užitečné pravidlo, že lze formálně užít obvyklý postup pro násobení matic, pouze místo prvků matice si představíme bloky.

Příklad 2.1. Jsou-li a, b, c, d, x, y skaláry, máme

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}.$$

³ Totéž bychom samozřejmě mohli udělat s řádky (a někdo to tak i dělá, např. počítačové grafici).

Jsou-li $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{X}, \mathbf{Y}$ matice vhodných rozměrů, máme tedy (ověřte dle vzorce (2.1)!) □

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{AX} + \mathbf{BY} \\ \mathbf{CX} + \mathbf{DY} \end{bmatrix}.$$

Často je užitečné vnímat matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ jako matici sestavenou z bloků

$$\mathbf{A} = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n],$$

kde sloupcové vektory $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ jsou sloupce matice \mathbf{A} . Matici lze také vnímat jako sestavenou z bloků

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix},$$

kde řádkové vektory $\mathbf{a}_1^T, \dots, \mathbf{a}_m^T$ jsou řádky matice \mathbf{A} , přičemž $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$.

Vyjádříme-li matici $\mathbf{A} \in \mathbb{R}^{m \times p}$ pomocí řádků a matici $\mathbf{B} \in \mathbb{R}^{p \times n}$ pomocí sloupců, je

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n] = [\mathbf{Ab}_1 \ \cdots \ \mathbf{Ab}_n] = \begin{bmatrix} \mathbf{a}_1^T \mathbf{B} \\ \vdots \\ \mathbf{a}_m^T \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T \mathbf{b}_1 & \cdots & \mathbf{a}_1^T \mathbf{b}_n \\ \vdots & \ddots & \vdots \\ \mathbf{a}_m^T \mathbf{b}_1 & \cdots & \mathbf{a}_m^T \mathbf{b}_n \end{bmatrix}. \quad (2.6)$$

Vyjádříme-li matici $\mathbf{A} \in \mathbb{R}^{m \times p}$ pomocí sloupců a matici $\mathbf{B} \in \mathbb{R}^{p \times n}$ pomocí řádků, je

$$\mathbf{AB} = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_p] \begin{bmatrix} \mathbf{b}_1^T \\ \vdots \\ \mathbf{b}_p^T \end{bmatrix} = \mathbf{a}_1 \mathbf{b}_1^T + \cdots + \mathbf{a}_p \mathbf{b}_p^T. \quad (2.7)$$

2.8 Hrubé chyby při práci s maticemi

Při manipulaci s maticovými výrazy a rovnicemi dělají zažatečníci někdy hrubé chyby, kterých se lze při alespoň minimální pozornosti vyhnout. Takové chyby jsou neomluvitelné. Uvedme typické příklady těchto chyb.

Výraz je nesmyslný kvůli rozměrům matic

Jako první příklad uvedme chyby, kdy výraz nemá smysl kvůli rozměrům matic a vektorů. Např.:

- Pokud $\mathbf{A} \in \mathbb{R}^{2 \times 3}$ a $\mathbf{B} \in \mathbb{R}^{3 \times 3}$, tak následující výrazy jsou chybné:

$$\mathbf{A} + \mathbf{B}, \quad \mathbf{A} = \mathbf{B}, \quad [\mathbf{A} \ \mathbf{B}], \quad \mathbf{A}^T \mathbf{B}, \quad \mathbf{A}^{-1}, \quad \det \mathbf{A}, \quad \mathbf{A}^2.$$

- Zcela odstrašující příklad je použití ‘zlomku’ pro matice, např. $\frac{\mathbf{A}}{\mathbf{B}}$.
- Inverze čtvercové, ale evidentně singulární matice. Např. $(\mathbf{w}\mathbf{w}^T)^{-1}$, kde $\mathbf{w} \in \mathbb{R}^3$.
- Předpoklad existence levé inverze široké matice. Ale široká matice má lineárně závislé sloupce, proto nemá levou inverzi. Podobně úzká matice matice nemá pravou inverzi. Napíšeme-li proto např. $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$, kde $\mathbf{Q} \in \mathbb{R}^{5 \times 3}$, je to chybné.

Použití neexistujících maticových identit

Pro manipulaci s maticovými výrazy je užitečné mít v paměti zásobu maticových identit. Ovšem nesmí být chybné. Typické příklady:

- $(\mathbf{AB})^T = \mathbf{A}^T \mathbf{B}^T$ (pokud v maticovém součinu $\mathbf{A}^T \mathbf{B}^T$ je vnitřní rozměr různý, je to chyba už kvůli rozměrům matic)
- $(\mathbf{AB})^{-1} = \mathbf{A}^{-1} \mathbf{B}^{-1}$ (pro nečtvercové matice je to také syntaktická chyba, pro čtvercové ale singulární matice je to také chyba kvůli rozměrům matic)
- $(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2$. Tato identita se opírá o velice užitečnou (avšak neexistující) identitu $\mathbf{AB} = \mathbf{BA}$. Správně je $(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + \mathbf{AB} + \mathbf{BA} + \mathbf{B}^2$.

Neekvivalentní úpravy (ne)rovnice

Zde student udělá chybný úsudek při *neekvivalentní úpravě* rovnice či nerovnice. Ekvivalentní a neekvivalentní úpravy skalárních rovnic známe již ze základní školy. Např. úprava ‘přičti k rovnici jedničku’ je ekvivalentní, neboť $(a + b) \Leftrightarrow (a + 1 = b + 1)$. Úprava ‘umocni rovnici na druhou’ je neekvivalentní, neboť sice $(a = b) \Rightarrow (a^2 = b^2)$, ale neplatí $(a = b) \Leftarrow (a^2 = b^2)$. Příklady:

- Úsudek, že $(\mathbf{a}^T \mathbf{x} = \mathbf{a}^T \mathbf{y}) \Rightarrow (\mathbf{x} = \mathbf{y})$ (není pravda, ani když vektor \mathbf{a} je nenulový).
- Úsudek, že pokud $\mathbf{A} \in \mathbb{R}^{3 \times 5}$ a $\mathbf{AX} = \mathbf{AY}$, pak platí $\mathbf{X} = \mathbf{Y}$ (není pravda, protože \mathbf{A} nemá lineárně nezávislé sloupce, tedy nemá levou inverzi).
- Úsudek, že $(\mathbf{A}^T \mathbf{A} = \mathbf{B}^T \mathbf{B}) \Rightarrow (\mathbf{A} = \mathbf{B})$ (není pravda dokonce ani pro skaláry).

Další nápady pro práci s maticemi

- Pod výrazy s maticemi a vektory si malujte obdélníčky s rozměry matic, abyste měli jasnou představu o jejich rozměrech.
- Vidíte-li maticovou rovnici či soustavu rovnic, spočítejte si skalární rovnice a neznámé.
- Pracujte nejen s papírem, ale i s Matlabem. Úpravy maticových výrazů lze často ověřit na náhodných maticích. Např. chceme-li ověřit rovnost $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$, zkusíme např. $\mathbf{A} = \text{randn}(5, 3)$; $\mathbf{B} = \text{randn}(3, 6)$; $(\mathbf{A} * \mathbf{B})' - \mathbf{B}' * \mathbf{A}'$. Samozřejmě to není důkaz.

2.9 Cvičení

2.1. Vyřešte tyto rovnice pro neznámou matici \mathbf{X} (předpokládejte, že každá potřebná inverze existuje):

- $\mathbf{AX} + \mathbf{B} = \mathbf{A}^2 \mathbf{X}$
- $\mathbf{X} - \mathbf{A} = \mathbf{XB}$
- $2\mathbf{X} - \mathbf{AX} + 2\mathbf{A} = \mathbf{0}$

2.2. Řešíme soustavu rovnic $\{\mathbf{b}_i = \mathbf{X}\mathbf{a}_i, i = 1, \dots, k\}$ pro neznámou matici $\mathbf{X} \in \mathbb{R}^{m \times n}$. Jaké musí být k , aby soustava měla stejný počet (skalárních) rovnic jako neznámých? Za jaké podmínky má soustava jediné řešení?

2.3. Vyřešte soustavu rovnic $\{\mathbf{Ax} = \mathbf{b}, \mathbf{x} = \mathbf{A}^T \mathbf{y}\}$, kde \mathbf{x}, \mathbf{y} jsou neznámé vektory. Předpokládejte, že matice \mathbf{AA}^T je regulární. Najděte jen vztah pro \mathbf{x} , vztah pro \mathbf{y} nás nezajímá. Ověřte v Matlabu pro náhodné zadání, které získáte příkazy $\mathbf{A}=\text{randn}(m,n)$; $\mathbf{b}=\text{randn}(m,1)$.

2.4. Mějme soustavu rovnic pro neznámé \mathbf{x} a \mathbf{y} :

$$\begin{aligned} \mathbf{Ax} + \mathbf{By} &= \mathbf{a} \\ \mathbf{Cx} + \mathbf{Dy} &= \mathbf{b} \end{aligned}$$

a) Vyjádřete soustavu ve tvaru $\mathbf{Pu} = \mathbf{q}$.

b) Je-li $\mathbf{a}, \mathbf{x} \in \mathbb{R}^m$, $\mathbf{b}, \mathbf{y} \in \mathbb{R}^n$, ukažte, že $\mathbf{x} = (\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1}(\mathbf{a} - \mathbf{BD}^{-1}\mathbf{b})$. Jakou to má výpočetní výhodu oproti počítání \mathbf{u} přímo ze soustavy $\mathbf{Pu} = \mathbf{q}$?

2.5. V následujících soustavách rovnic malá písmena značí vektory a velká matice. Jaké jsou nejobecnější rozměry matic a vektorů, aby rovnice byly syntakticky správně? Jaký je počet rovnic a neznámých v každé soustavě? Které z těchto soustav rovnic jsou lineární?

- $\mathbf{Ax} = \mathbf{b}$, neznámá \mathbf{x} .
- $\mathbf{x}^T \mathbf{Ax} = 1$, neznámá \mathbf{x} .
- $\mathbf{a}^T \mathbf{Xb} = 0$, neznámá \mathbf{X} .
- $\mathbf{AX} + \mathbf{XA}^T = \mathbf{C}$, neznámá \mathbf{X} .
- $\{\mathbf{X}^T \mathbf{Y} = \mathbf{A}, \mathbf{XY}^T = \mathbf{B}\}$, neznámé \mathbf{X}, \mathbf{Y} .

2.6. Zobrazení $\text{vec}: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{mn}$ ('vektORIZACE' matice, v Matlabu označeno $\mathbf{A}(:)$) je definováno tak, že $\text{vec} \mathbf{A}$ je matice \mathbf{A} přerovnaná po sloupcích do vektoru. *Kroneckerův součin matic* (v Matlabu $\text{kron}(\mathbf{A}, \mathbf{B})$) je definován jako

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}.$$

Pro libovolné matice (s kompatibilními velikostmi) platí

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec} \mathbf{B}. \quad (2.8)$$

Použijte tohoto vzorce pro převedení následujících soustav rovnic s neznámou maticí \mathbf{X} do tvaru $\mathbf{Pu} = \mathbf{q}$ s neznámým vektorem \mathbf{u} . Předpokládejte, že počet rovnic je rovný počtu neznámých. Předpokládejte, že matice a vektory mají nejobecnější možné rozměry, aby to dávalo smysl.

- $\{\mathbf{b}_i^T \mathbf{Xa}_i = 0, i = 1, \dots, k\}$
- $\mathbf{AX} + \mathbf{XA}^T = \mathbf{C}$

2.7. Dokažte:

- Matice \mathbf{AB} a \mathbf{BA} mají stejnou stopu.
- Rovnice $\mathbf{AB} - \mathbf{BA} = \mathbf{I}$ nemá řešení pro žádné \mathbf{A}, \mathbf{B} .

2.8. *Komutátorem* dvou matic rozumíme matici $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$. Dokažte, že platí *Jacobiho identita* $[\mathbf{A}, [\mathbf{B}, \mathbf{C}]] + [\mathbf{B}, [\mathbf{C}, \mathbf{A}]] + [\mathbf{C}, [\mathbf{A}, \mathbf{B}]] = \mathbf{0}$.

2.9. Dokažte *Sherman-Morrisonův vzorec*

$$(\mathbf{A} + \mathbf{uv}^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{uv}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}$$

(\mathbf{A} je čtvercová regulární a $\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u} \neq -1$) a *Sherman-Morrison-Woodburyho vzorec*

$$(\mathbf{A} + \mathbf{UV}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{I} + \mathbf{V}^T\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}^T\mathbf{A}^{-1}$$

(kde $\mathbf{I} + \mathbf{V}^T\mathbf{A}^{-1}\mathbf{U}$ je regulární).

2.10. Dokažte, že $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

2.11. Dokažte, že pro každou čtvercovou matici \mathbf{A}

- $\mathbf{A} + \mathbf{A}^T$ je symetrická
- $\mathbf{A} - \mathbf{A}^T$ je antisymetrická
- existuje symetrická \mathbf{B} a antisymetrická \mathbf{C} tak, že $\mathbf{A} = \mathbf{B} + \mathbf{C}$, přičemž \mathbf{B}, \mathbf{C} jsou dány jednoznačně.
- $\mathbf{A}^T \mathbf{A}$ je symetrická

2.12. Dokažte, že pro každé $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{B} \in \mathbb{R}^{n \times m}$ má matice

$$\mathbf{L} = \begin{bmatrix} \mathbf{I} - \mathbf{BA} & \mathbf{B} \\ 2\mathbf{A} - \mathbf{ABA} & \mathbf{AB} - \mathbf{I} \end{bmatrix}$$

vlastnost $\mathbf{L}^2 = \mathbf{I}$ (kde \mathbf{L}^2 je zkratka pro \mathbf{LL}). Matice s touto vlastností se nazývá *involve*.

2.13. Kdy je diagonální matice regulární? Co je inverzí diagonální matice?

2.14. Ukažte, že diagonální matice komutují (tj. $\mathbf{AB} = \mathbf{BA}$).

2.15. Dokažte, že pokud je $\mathbf{I} - \mathbf{A}$ regulární, pak $\mathbf{A}(\mathbf{I} - \mathbf{A})^{-1} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{A}$.

2.16. Dokažte, že pokud \mathbf{A}, \mathbf{B} a $\mathbf{A} + \mathbf{B}$ jsou regulární, pak

$$\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B} = \mathbf{B}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{A} = (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}.$$

2.17. (*) Necht' čtvercové matice $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ jsou takové, že \mathbf{AB}^T a \mathbf{CD}^T jsou symetrické a platí $\mathbf{AD}^T - \mathbf{BC}^T = \mathbf{I}$. Dokažte, že $\mathbf{A}^T \mathbf{D} - \mathbf{C}^T \mathbf{B} = \mathbf{I}$.

Nápověda a řešení

2.1.a) $\mathbf{X} = (\mathbf{A}^2 - \mathbf{A})^{-1}\mathbf{B} = (\mathbf{A} - \mathbf{I})^{-1}\mathbf{A}^{-1}\mathbf{B}$

2.1.b) $\mathbf{X} = \mathbf{A}(\mathbf{I} - \mathbf{B})^{-1}$

2.1.c) $\mathbf{X} = 2(\mathbf{A} - 2\mathbf{I})^{-1}\mathbf{A} = (\mathbf{A}/2 - \mathbf{I})^{-1}\mathbf{A}$

2.2. Neznámých je $m \times n$, rovnic je $m \times k$, tedy musí být $n = k$. Pro jediné řešení musí být vektory \mathbf{a}_i lineárně nezávislé.

2.3. $\mathbf{x} = \mathbf{A}^T(\mathbf{AA}^T)^{-1}\mathbf{b}$

2.4.a) $\mathbf{P} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$, $\mathbf{q} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}$

2.5.a) Rovnic je m , neznámých n , kde $\mathbf{A} \in \mathbb{R}^{m \times n}$. Je lineární.

2.5.b) Rovnice je jedna, neznámých je n , kde $\mathbf{x} \in \mathbb{R}^n$. Není lineární.

2.5.c) Rovnice je jedna, neznámých je mn , kde $\mathbf{X} \in \mathbb{R}^{m \times n}$. Je lineární.

2.5.d) Všechny tři matice $\mathbf{A}, \mathbf{C}, \mathbf{X}$ musí být čtvercové velikosti $n \times n$. Rovnic i neznámých je n^2 .

2.5.e) Rovnic je $m^2 + n^2$, neznámých je $2mn$, kde $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{m \times n}$. Není lineární.

Kapitola 3

Linearita

3.1 Lineární podprostory

Množina \mathbb{R}^n spolu s operacemi sčítání vektorů a násobení vektorů skalárem tvoří *lineární prostor* nad tělesem \mathbb{R} . Zopakujte si z lineární algebry pojem lineárního prostoru!

Lineární kombinace vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ je vektor

$$\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k \quad (3.1)$$

pro nějaké skaláry $\alpha_1, \dots, \alpha_k \in \mathbb{R}$. Vektory jsou **lineárně nezávislé**, když platí implikace

$$\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k = \mathbf{0} \implies \alpha_1 = \dots = \alpha_k = 0. \quad (3.2)$$

V opačném případě jsou **lineárně závislé**. Jsou-li vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$ lineárně nezávislé, potom koeficienty $\alpha_1, \dots, \alpha_k$ jsou vektorem $\mathbf{x} = \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$ určeny jednoznačně (dokažte!).

Lineární obal vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je množina

$$\text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} = \{ \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k \mid \alpha_1, \dots, \alpha_k \in \mathbb{R} \}$$

všech jejich lineárních kombinací (zde předpokládáme, že vektorů je konečný počet).

Množina $X \subseteq \mathbb{R}^n$ se nazývá **lineární podprostor** (nebo jen **podprostor**) lineárního prostoru \mathbb{R}^n , pokud libovolná lineární kombinace libovolné množiny vektorů z X leží v X (neboli že množina X je uzavřená vůči lineárním kombinacím):

$$\mathbf{x}_1, \dots, \mathbf{x}_k \in X, \alpha_1, \dots, \alpha_k \in \mathbb{R} \implies \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k \in X. \quad (3.3)$$

Lineární obal libovolné množiny vektorů je lineární podprostor.

Báze lineárního podprostoru $X \subseteq \mathbb{R}^n$ je lineárně nezávislá množina vektorů, jejíž lineární obal je X . Je-li $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ báze podprostoru a $\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k = \mathbf{x} \in X$, potom (jednoznačně určené) skaláry $\alpha_1, \dots, \alpha_k$ se nazývají **souřadnice** vektoru \mathbf{x} v dané bázi. Netriviální podprostor prostoru \mathbb{R}^n má nekonečný počet bází, každá báze má však stejný počet vektorů. Tento počet se nazývá **dimenze** lineárního podprostoru, kterou značíme $\dim X$.

Příklad 3.1. Množina $X = \text{span}\{(1, 2, 3)\} = \{ \alpha(1, 2, 3) \mid \alpha \in \mathbb{R} \} \subseteq \mathbb{R}^3$ je podprostor \mathbb{R}^3 dimenze 1. Je to přímka procházející počátkem. Její báze je např. množina $\{(1, 2, 3)\}$, jiná báze je množina $\{(2, 4, 6)\}$. \square

Příklad 3.2. Množina $X = \text{span}\{(1, 2, 3, 0), (0, 1, 2, -1), (0, 2, 4, -2)\} \subseteq \mathbb{R}^4$ je podprostor \mathbb{R}^4 dimenze 2. Všimněte si, že vektory jsou lineárně závislé. Báze podprostoru X je např. množina $\{(1, 2, 3, 0), (0, 1, 2, -1)\}$. \square

Příklad 3.3. Všechny možné podprostory prostoru \mathbb{R}^3 jsou tyto: počátek $\mathbf{0}$ (dimenze 0), všechny přímky procházející počátkem (dimenze 1), všechny roviny procházející počátkem (dimenze 2), a konečně celý prostor \mathbb{R}^3 (dimenze 3). \square

Příklad 3.4. Množina $X = \{ (1 + \alpha, \alpha) \mid \alpha \in \mathbb{R} \} \subseteq \mathbb{R}^2$ (přímka neprocházející počátkem) není podprostor \mathbb{R}^2 , protože např. $(1, 0) \in X$ ale $2(1, 0) = (2, 0) \notin X$. \square

3.2 Lineární zobrazení

Zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je **lineární**, jestliže pro každé $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ a $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ platí

$$\mathbf{f}(\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k) = \alpha_1 \mathbf{f}(\mathbf{x}_1) + \dots + \alpha_k \mathbf{f}(\mathbf{x}_k), \quad (3.4)$$

tedy jestliže ‘zobrazení lineární kombinace je rovno lineární kombinaci zobrazení’.

Věta 3.1. Zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je lineární právě když existuje matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ taková, že

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}. \quad (3.5)$$

Matice \mathbf{A} je navíc zobrazením \mathbf{f} určena jednoznačně.

Ríkáme proto, že matice \mathbf{A} *reprezentuje* lineární zobrazení \mathbf{f} .

Důkaz. Důkaz jedné implikace je snadný: zobrazení (3.5) je lineární, neboť

$$\mathbf{f}(\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k) = \mathbf{A}(\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k) = \alpha_1 \mathbf{A}\mathbf{x}_1 + \dots + \alpha_k \mathbf{A}\mathbf{x}_k = \alpha_1 \mathbf{f}(\mathbf{x}_1) + \dots + \alpha_k \mathbf{f}(\mathbf{x}_k).$$

Dokažme opačnou implikaci. Necht' $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je lineární zobrazení. Necht' $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{R}^n$ je standardní báze prostoru \mathbb{R}^n . Pro každé $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ máme $\mathbf{x} = x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n$. Z (3.4) plyne

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n) = x_1 \mathbf{f}(\mathbf{e}_1) + \dots + x_n \mathbf{f}(\mathbf{e}_n) = [\mathbf{f}(\mathbf{e}_1) \ \dots \ \mathbf{f}(\mathbf{e}_n)] \mathbf{x}.$$

Nyní označíme $[\mathbf{f}(\mathbf{e}_1) \ \dots \ \mathbf{f}(\mathbf{e}_n)] = \mathbf{A}$. Tedy \mathbf{A} je matice se sloupci $\mathbf{f}(\mathbf{e}_1), \dots, \mathbf{f}(\mathbf{e}_n)$.

Dokažme jednoznačnost matice \mathbf{A} . Platí-li $\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x}$ pro každé $\mathbf{x} \in \mathbb{R}^n$, pak samozřejmě $\mathbf{A} = \mathbf{B}$. Stačí dosadit za \mathbf{x} postupně vektory $\mathbf{e}_1, \dots, \mathbf{e}_n$. \square

Příklad 3.5. Zobrazení $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definované jako $\mathbf{f}(x_1, x_2) = (x_1 + x_2, x_1 - x_2, 2x_1)$ je lineární. To bychom dokázali ověřením podmínky (3.4). Ovšem je to patrné na první pohled, protože jej lze vyjádřit ve tvaru (3.5):

$$\mathbf{f}(x_1, x_2) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \\ 2x_1 \end{bmatrix} = (x_1 + x_2, x_1 - x_2, 2x_1). \quad \square$$

Pokud $m = 1$, lineární zobrazení je funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ tvaru

$$f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} = a_1 x_1 + \cdots + a_n x_n, \quad (3.6)$$

kde $\mathbf{a} \in \mathbb{R}^n$. Této funkci se též říká *lineární forma*.

Podívejme se blíže na vzorec (3.5). Výraz \mathbf{Ax} je maticový součin matice $m \times n$ maticí $n \times 1$ (viz §2.6). Označíme-li $\mathbf{y} = \mathbf{Ax}$, je tedy podle (2.1)

$$y_i = \sum_{j=1}^n a_{ij} x_j \quad (3.7)$$

neboli

$$\begin{aligned} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + \cdots + a_{mn}x_n. \end{aligned}$$

Dále, vyjádříme-li matici \mathbf{A} pomocí sloupců, máme

$$\mathbf{Ax} = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n, \quad (3.8)$$

tedy vektor \mathbf{Ax} je *lineární kombinace sloupců* matice \mathbf{A} . Naopak, vyjádříme-li matici \mathbf{A} pomocí řádků, máme

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{a}_1^T \mathbf{x} \\ \vdots \\ \mathbf{a}_m^T \mathbf{x} \end{bmatrix}, \quad (3.9)$$

tedy složky vektoru \mathbf{Ax} jsou *skalární součiny řádků* matice \mathbf{A} a vektoru \mathbf{x} . Všimněte si, že (3.8) a (3.9) jsou speciální případy (2.6) a (2.7).

Složení lineárních zobrazení je opět lineární zobrazení. Pro $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ a $\mathbf{g}(\mathbf{y}) = \mathbf{By}$ máme

$$(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{B}(\mathbf{Ax}) = (\mathbf{BA})\mathbf{x},$$

tedy \mathbf{BA} je matice složeného zobrazení $\mathbf{g} \circ \mathbf{f}$. Tedy matice složeného zobrazení je součinem matic jednotlivých zobrazení. Toto je hlavní důvod, proč je rozumné definovat maticové násobení jako (2.1): *násobení matic odpovídá skládání lineárních zobrazení* reprezentovanými těmito maticemi.

3.3 Prostor obrazů a hodnota matice

S lineárním zobrazením jsou spjaty dva lineární podprostory, prostor obrazů a nulový prostor (jádro). Je-li zobrazení reprezentováno maticí jako $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$, hovoříme o prostoru obrazů a nulovém prostoru matice \mathbf{A} .

Prostor obrazů matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ je množina

$$\text{rng } \mathbf{A} = \{ \mathbf{Ax} \mid \mathbf{x} \in \mathbb{R}^n \} \subseteq \mathbb{R}^m. \quad (3.10)$$

Interpretace prostoru obrazů:

- Je to množina $\mathbf{f}(\mathbb{R}^n)$ všech hodnot, jichž může zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ nabýt.
- Je to množina všech vektorů \mathbf{y} , pro které má lineární soustava $\mathbf{Ax} = \mathbf{y}$ řešení.
- Je to lineární obal sloupců matice \mathbf{A} . Tedy je to lineární podprostor \mathbb{R}^m .

Dimenze lineárního obalu sloupců je **hodnota** (angl. *rank*) matice, značíme ji

$$\text{rank } \mathbf{A} = \dim \text{rng } \mathbf{A}. \quad (3.11)$$

Věta 3.2. *Dimenze lineárního obalu sloupců je rovna dimenzi lineárního obalu řádků, tedy*

$$\text{rank } \mathbf{A} = \dim \text{rng } \mathbf{A} = \dim \text{rng}(\mathbf{A}^T) = \text{rank}(\mathbf{A}^T). \quad (3.12)$$

Důkaz. Necht' $\mathbf{A} \in \mathbb{R}^{m \times n}$. Zvolme libovolnou bázi prostoru $\text{rng } \mathbf{A}$. Necht' vektory této báze tvoří sloupce matice $\mathbf{B} \in \mathbb{R}^{m \times r}$, kde $r = \dim \text{rng } \mathbf{A}$. Nyní j -tý sloupec \mathbf{a}_j matice \mathbf{A} je lineární kombinací sloupců matice \mathbf{B} , neboli existuje vektor \mathbf{c}_j tak že $\mathbf{a}_j = \mathbf{Bc}_j$. Je tedy

$$\mathbf{A} = \mathbf{BC}, \quad (3.13)$$

kde matice $\mathbf{C} \in \mathbb{R}^{r \times n}$ má sloupce $\mathbf{c}_1, \dots, \mathbf{c}_n$.

Rovnost (3.13) zároveň říká, že i -tý řádek \mathbf{A} je lineární kombinací řádků \mathbf{C} , kde koeficienty této lineární kombinace tvoří i -tý řádek \mathbf{B} . Protože \mathbf{C} má r řádků, musí být¹

$$\dim \text{rng}(\mathbf{A}^T) \leq r = \dim \text{rng } \mathbf{A}. \quad (3.14)$$

Dokázanou nerovnost (3.14) nyní použijeme na matici \mathbf{A}^T , což dá $\dim \text{rng } \mathbf{A} \leq \dim \text{rng}(\mathbf{A}^T)$. Obě nerovnosti dohromady dávají (3.12). \square

Mimoděk jsme dokázali, že každá matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ hodnosti r jde rozložit na součin (3.13), kde $\mathbf{B} \in \mathbb{R}^{m \times r}$ a $\mathbf{C} \in \mathbb{R}^{r \times n}$. To je známo jako **rozklad matice podle hodnot** (angl. *rank factorization*).

Je jasné, že

$$\text{rank } \mathbf{A} \leq \min\{m, n\}. \quad (3.15)$$

Když $\text{rank } \mathbf{A} = \min\{m, n\}$, říkáme, že matice má **plnou hodnotu**. Je $\text{rank } \mathbf{A} = n$ právě když \mathbf{A} má lineárně nezávislé sloupce, a $\text{rank } \mathbf{A} = m$ právě když \mathbf{A} má lineárně nezávislé řádky. Čtvercová matice má plnou hodnotu právě když je regulární.

Věta 3.3. *Následující tvrzení jsou ekvivalentní:*

1. $\text{rng } \mathbf{A} = \mathbb{R}^m$
2. $\text{rank } \mathbf{A} = m$
3. Zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ je surjektivní (viz §1.1.2).
4. Soustava $\mathbf{Ax} = \mathbf{y}$ má řešení pro každé \mathbf{y} .
5. Řádky matice \mathbf{A} jsou lineárně nezávislé.
6. Matice \mathbf{A} má pravou inverzi, tj. existuje \mathbf{B} tak, že $\mathbf{AB} = \mathbf{I}$.

¹Zde tvrdíme, že je-li každý z vektorů $\mathbf{a}_1, \dots, \mathbf{a}_m$ lineární kombinací vektorů $\mathbf{c}_1, \dots, \mathbf{c}_r$, pak podprostor $\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$ nemůže mít větší dimenzi než podprostor $\text{span}\{\mathbf{c}_1, \dots, \mathbf{c}_r\}$. Toto zřejmé tvrzení dále nedokazujeme; přesně jej lze dokázat pomocí Věty 3.4.

Důkaz.

- $1 \Leftrightarrow 2 \Leftrightarrow 3 \Leftrightarrow 4$ plyne přímo z definic.
- $2 \Leftrightarrow 5$ plyne z (3.11) a (3.12).
- $4 \Rightarrow 6$ platí, neboť soustava $\mathbf{A}\mathbf{b}_i = \mathbf{e}_i$ má řešení \mathbf{b}_i pro každé i (kde \mathbf{e}_i resp. \mathbf{b}_i je i -tý sloupec matice \mathbf{I} resp. \mathbf{B}). Pro důkaz $6 \Rightarrow 4$ položíme $\mathbf{x} = \mathbf{B}\mathbf{y}$. \square

Věta 3.4. Pro libovolné matice \mathbf{A}, \mathbf{B} platí

- $\text{rng}(\mathbf{AB}) \subseteq \text{rng } \mathbf{A}$.
- $\text{rng}(\mathbf{AB}) = \text{rng } \mathbf{A}$, jestliže řádky \mathbf{B} jsou lineárně nezávislé.

Důkaz. První tvrzení říká, že jestliže soustava $\mathbf{z} = \mathbf{AB}\mathbf{x}$ má řešení, pak i soustava $\mathbf{z} = \mathbf{A}\mathbf{y}$ má řešení. To je ale jasné, protože můžeme položit $\mathbf{y} = \mathbf{B}\mathbf{x}$. Druhé tvrzení navíc říká, že jestliže $\mathbf{z} = \mathbf{A}\mathbf{y}$ má řešení, pak i $\mathbf{z} = \mathbf{AB}\mathbf{x}$ má řešení. To plyne z implikace $3 \Rightarrow 4$ ve Větě 3.3. \square

3.4 Nulový prostor matice

Nulový prostor matice \mathbf{A} (také se nazývá *jádro* zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$) je množina

$$\text{null } \mathbf{A} = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0} \} \subseteq \mathbb{R}^n. \quad (3.16)$$

Interpretace nulového prostoru:

- Je to množina všech vektorů, které se zobrazí na nulový vektor.
- Podle (3.9) je to množina všech vektorů, které jsou kolmé na každý řádek matice \mathbf{A} . Z toho je vidět, že je to lineární podprostor \mathbb{R}^n .

Věta 3.5. Následující tvrzení jsou ekvivalentní:

1. $\text{null } \mathbf{A} = \{ \mathbf{0} \}$ (tj. nulový prostor je triviální).
2. Sloupce matice \mathbf{A} jsou lineárně nezávislé.
3. $\text{rank } \mathbf{A} = n$.
4. Zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ je injektivní (viz §1.1.2).
5. Matice \mathbf{A} má levou inverzi, tj. existuje \mathbf{B} tak, že $\mathbf{BA} = \mathbf{I}$.

Důkaz.

- Ekvivalence $1 \Leftrightarrow 3$ plyne z definice lineární nezávislosti (3.2), tj. $(\mathbf{A}\mathbf{x} = \mathbf{0}) \Rightarrow (\mathbf{x} = \mathbf{0})$.
- Ekvivalence $2 \Leftrightarrow 3$ plyne z definice hodnosti (3.11).
- Tvrzení 4 říká, že pro každé \mathbf{x}, \mathbf{y} je $\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{y} \Rightarrow \mathbf{x} = \mathbf{y}$, neboli $\mathbf{A}(\mathbf{x} - \mathbf{y}) = \mathbf{0} \Rightarrow \mathbf{x} - \mathbf{y} = \mathbf{0}$. Ale to je definice (3.2) lineární nezávislosti sloupců \mathbf{A} . Tedy platí $1 \Leftrightarrow 4$.
- Tvrzení 5 je ekvivalentní tomu, že matice \mathbf{A}^T má pravou inverzi, tj. $\mathbf{A}^T\mathbf{B}^T = \mathbf{I}$. Tedy $3 \Leftrightarrow 5$ plyne z Věty 3.3. \square

Věta 3.6. Pro libovolné matice \mathbf{A}, \mathbf{B} platí:

- $\text{null}(\mathbf{BA}) \supseteq \text{null } \mathbf{A}$.

- $\text{null}(\mathbf{BA}) = \text{null } \mathbf{A}$, jestliže sloupce \mathbf{B} jsou lineárně nezávislé.

Důkaz. První tvrzení platí, protože $\mathbf{A}\mathbf{x} = \mathbf{0}$ implikuje $\mathbf{BA}\mathbf{x} = \mathbf{0}$ (po vynásobení maticí \mathbf{B} zleva). Druhé tvrzení platí, neboť $\text{null } \mathbf{B} = \{ \mathbf{0} \}$, což je implikace $3 \Rightarrow 1$ ve Větě 3.5. \square

Následující věta dává do vztahu dimenze prostoru obrazů, nulového prostoru a prostoru \mathbb{R}^n .

Věta 3.7. Pro každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ platí

$$\dim \text{rng } \mathbf{A} + \dim \text{null } \mathbf{A} = n. \quad (3.17)$$

Důkaz. (*) Necht' báze prostoru $\text{null } \mathbf{A}$ tvoří sloupce matice $\mathbf{C} \in \mathbb{R}^{n \times k}$, kde $k = \dim \text{null } \mathbf{A}$. Necht' matice $\mathbf{B} \in \mathbb{R}^{n \times (n-k)}$ je taková, že matice $[\mathbf{B} \ \mathbf{C}] \in \mathbb{R}^{n \times n}$ je regulární. Ukážeme, že sloupce matice $\mathbf{AB} \in \mathbb{R}^{n \times k}$ tvoří bázi prostoru $\text{rng } \mathbf{A}$. Tím větu dokážeme.

Jelikož sloupce $[\mathbf{B} \ \mathbf{C}]$ tvoří bázi \mathbb{R}^n , každý vektor $\mathbf{x} \in \mathbb{R}^n$ lze psát jako $\mathbf{x} = \mathbf{B}\mathbf{y} + \mathbf{C}\mathbf{z}$. Z toho $\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{B}\mathbf{y} + \mathbf{A}\mathbf{C}\mathbf{z} = \mathbf{A}\mathbf{B}\mathbf{y}$, neboť $\mathbf{A}\mathbf{C}\mathbf{z} = \mathbf{0}$ protože $\mathbf{C}\mathbf{z} \in \text{null } \mathbf{A}$. Tedy $\text{rng}(\mathbf{AB}) = \text{rng } \mathbf{A}$.

Zbývá dokázat, že matice \mathbf{AB} má lineárně nezávislé sloupce, tedy že $(\mathbf{A}\mathbf{B}\mathbf{y} = \mathbf{0}) \Rightarrow (\mathbf{y} = \mathbf{0})$. Necht' $\mathbf{A}\mathbf{B}\mathbf{y} = \mathbf{0}$. Z toho plyne $\mathbf{B}\mathbf{y} \in \text{null } \mathbf{A}$, tedy $\mathbf{B}\mathbf{y} = \mathbf{C}\mathbf{z}$ pro nějaké \mathbf{z} . Ale protože $[\mathbf{A} \ \mathbf{B}]$ má lineárně nezávislé sloupce, dle (3.2) rovnost $\mathbf{B}\mathbf{y} = \mathbf{C}\mathbf{z}$ implikuje $\mathbf{y} = \mathbf{0}$ a $\mathbf{z} = \mathbf{0}$. \square

Neformální interpretace věty: Číslo $\dim \text{rng } \mathbf{A} = \text{rank } \mathbf{A}$ je dimenze podprostoru všech hodnot, které se mohou objevit na výstupu zobrazení. Číslo $\dim \text{null } \mathbf{A}$ je dimenze podprostoru vstupních hodnot, které zobrazení 'smáčkne' do nulového vektoru. Věta říká, že každá vstupní dimenze se buď 'smáčkne' do nulového vektoru nebo se objeví na výstupu. Jiná interpretace: počet lineárně nezávislých vektorů, které řeší homogenní soustavu $\mathbf{A}\mathbf{x} = \mathbf{0}$, je právě $n - \text{rank } \mathbf{A}$.

3.5 Afinní podprostor a zobrazení

Afinní kombinace vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ je lineární kombinace $\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k$, ve které koeficienty kombinace splňují

$$\alpha_1 + \dots + \alpha_k = 1.$$

Afinní obal vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je množina všech jejich afinních kombinací.

Příklad 3.6. Afinní obal dvou bodů $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ je množina

$$\text{aff}\{\mathbf{x}_1, \mathbf{x}_2\} = \{ \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 \mid \alpha_1 + \alpha_2 = 1 \}.$$

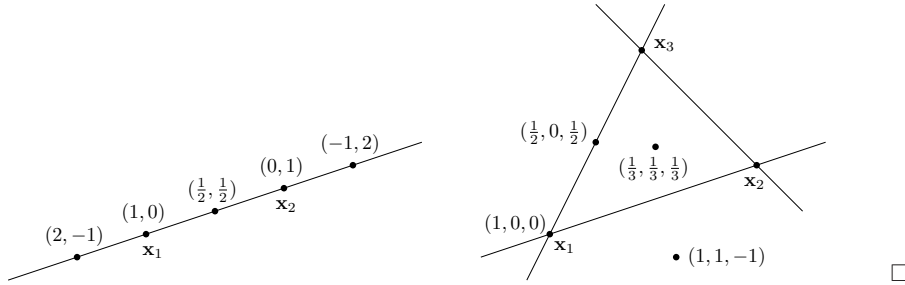
Pokud $\mathbf{x}_1 \neq \mathbf{x}_2$, je touto množinou přímka procházející body $\mathbf{x}_1, \mathbf{x}_2$. Na obrázku dole vlevo je několik bodů na přímce a příslušné vektory (α_1, α_2) .

Naproti tomu, lineární obal dvou vektorů $\mathbf{x}_1, \mathbf{x}_2$ (pokud jsou lineárně nezávislé) je rovina procházející těmito dvěma body a počátkem $\mathbf{0}$.

Afinní obal tří bodů $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \in \mathbb{R}^n$ je množina

$$\text{aff}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\} = \{ \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_3\mathbf{x}_3 \mid \alpha_1 + \alpha_2 + \alpha_3 = 1 \}.$$

Pokud body $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ neleží v přímce, je touto množinou rovina jimi procházející. Na obrázku vpravo je několik bodů v této rovině a příslušné vektory $(\alpha_1, \alpha_2, \alpha_3)$.



Vektory $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ nazveme **afinně nezávislé**, jestliže žádný z nich není roven afinní kombinaci ostatních. To je ekvivalentní tomu (rozmyslete!), že (srov. se (3.2))

$$\alpha_1 + \dots + \alpha_k = 0, \quad \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k = \mathbf{0} \implies \alpha_1 = \dots = \alpha_k = 0. \quad (3.18)$$

Afinní nezávislost lze ekvivalentně definovat i dvěma dalšími způsoby, jak říká tato věta:

Věta 3.8. Následující tvrzení jsou ekvivalentní:

1. Platí implikace (3.18).
2. Vektory $\mathbf{x}_2 - \mathbf{x}_1, \mathbf{x}_3 - \mathbf{x}_1, \dots, \mathbf{x}_k - \mathbf{x}_1$ jsou lineárně nezávislé.
3. Vektory $\begin{bmatrix} \mathbf{x}_1 \\ 1 \end{bmatrix}, \dots, \begin{bmatrix} \mathbf{x}_k \\ 1 \end{bmatrix}$ jsou lineárně nezávislé.

Důkaz. Ponecháváme jako cvičení. □

Příklad 3.7. Dva body $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ jsou afinně nezávislé právě tehdy, když $\mathbf{x}_1 \neq \mathbf{x}_2$. Tři body $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \in \mathbb{R}^n$ jsou afinně nezávislé právě když neleží v jedné přímce (neboli nejsou *kolinéární*). Viz Příklad 3.6. Čtyři body $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4 \in \mathbb{R}^n$ jsou afinně nezávislé právě když neleží v jedné rovině (neboli nejsou *koplanární*). □

Afinní podprostor² lineárního prostoru \mathbb{R}^n (též známý jako *lineární varieta*) je množina $A \subseteq \mathbb{R}^n$, která je uzavřená vůči afinním kombinacím (tedy každá afinní kombinace vektorů z A leží v A). Následující věta charakterizuje afinní podprostory. Ve větě pro množinu $X \subseteq \mathbb{R}^n$ a vektor $\mathbf{y} \in \mathbb{R}^n$ označíme

$$X + \mathbf{y} = \{ \mathbf{x} + \mathbf{y} \mid \mathbf{x} \in X \}. \quad (3.19)$$

Věta 3.9.

- Je-li A afinní podprostor \mathbb{R}^n a $\mathbf{x}_0 \in A$, pak množina $A - \mathbf{x}_0$ je lineární podprostor \mathbb{R}^n .
- Je-li X lineární podprostor \mathbb{R}^n a $\mathbf{x}_0 \in \mathbb{R}^n$, pak množina $X + \mathbf{x}_0$ je afinní podprostor \mathbb{R}^n .

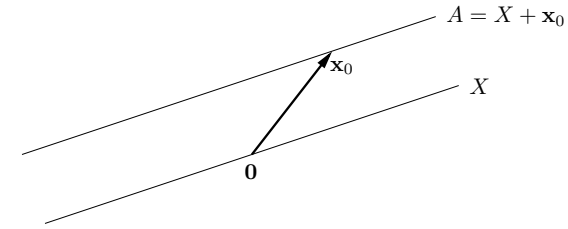
Důkaz. Dokážeme jen první tvrzení, druhé se dokáže podobně. Chceme dokázat, že libovolná lineární kombinace vektorů z množiny $A - \mathbf{x}_0$ leží v $A - \mathbf{x}_0$. To znamená, že pro $\mathbf{x}_1, \dots, \mathbf{x}_k \in A$ a $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ musí být $\alpha_1(\mathbf{x}_1 - \mathbf{x}_0) + \dots + \alpha_k(\mathbf{x}_k - \mathbf{x}_0) \in A - \mathbf{x}_0$, tedy

$$\alpha_1(\mathbf{x}_1 - \mathbf{x}_0) + \dots + \alpha_k(\mathbf{x}_k - \mathbf{x}_0) + \mathbf{x}_0 = \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k + (1 - \alpha_1 - \dots - \alpha_k) \mathbf{x}_0 \in A.$$

To je ale pravda, neboť $\alpha_1 + \dots + \alpha_k + (1 - \alpha_1 - \dots - \alpha_k) = 1$ a tedy poslední výraz je afinní kombinace vektorů z A , která podle předpokladu leží v A . □

²Všimněte si, že zde definujeme afinní *podprostor* lineárního prostoru, ale už ne afinní *prostor* sám o sobě. Definice afinního prostoru bez odkazu k nějakému lineárnímu prostoru existuje, ale neuvádíme ji.

Věta ukazuje, že afinní podprostor není nic jiného než ‘posunutý’ lineární podprostor (tedy nemusí procházet počátkem, na rozdíl od lineárního podprostoru). Viz obrázek:



Dimenze afinního podprostoru je dimenze tohoto lineárního podprostoru. Afinnímu podprostoru \mathbb{R}^n dimenze 0, 1, 2 a $n - 1$ se říká po řadě **bod**, **přímka**, **rovina** a **nadrovina**.

Příklad 3.8. Dokážeme, že množina $A = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b} \}$ všech řešení nehomogenní lineární soustavy je afinní podprostor \mathbb{R}^n . Necht' $\mathbf{x}_1, \dots, \mathbf{x}_k \in A$ a $\alpha_1 + \dots + \alpha_k = 1$. Pak

$$\mathbf{A}(\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k) = \alpha_1 \mathbf{A}\mathbf{x}_1 + \dots + \alpha_k \mathbf{A}\mathbf{x}_k = \alpha_1 \mathbf{b} + \dots + \alpha_k \mathbf{b} = \mathbf{b}.$$

Necht' nyní $X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0} \} = \text{null } \mathbf{A}$ a necht' $\mathbf{x}_0 \in A$ je libovolné (tzv. *partikulární*) řešení soustavy. Platí

$$X + \mathbf{x}_0 = \{ \mathbf{x} + \mathbf{x}_0 \mid \mathbf{A}\mathbf{x} = \mathbf{0} \} = \{ \mathbf{x}' \mid \mathbf{A}(\mathbf{x}' - \mathbf{x}_0) = \mathbf{0} \} = \{ \mathbf{x}' \mid \mathbf{A}\mathbf{x}' - \mathbf{b} = \mathbf{0} \} = A. \quad \square$$

Zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ nazveme **afinní**, pokud (3.4) platí pro všechna $\alpha_1 + \dots + \alpha_k = 1$, tedy zobrazení afinní kombinace je rovno afinní kombinaci zobrazení. Lze ukázat (proved'te!), že zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je afinní právě tehdy, když existuje matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ a vektor $\mathbf{b} \in \mathbb{R}^m$ tak, že

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}. \quad (3.20)$$

Příklad 3.9. Zobrazení $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definované jako $\mathbf{f}(x_1, x_2) = (x_1 + x_2 + 1, x_1 - x_2, 2x_1)$ je afinní. To bychom mohli dokázat ověřením podmínek (3.4) pro $\alpha_1 + \dots + \alpha_k = 1$. Ale je to patrné i z toho, že zobrazení lze vyjádřit ve tvaru (3.20):

$$\mathbf{f}(x_1, x_2) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \quad \square$$

Pro $m = 1$ je zobrazení (3.20) **afinní funkce**³ $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a má tvar

$$f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b = a_1 x_1 + \dots + a_n x_n + b, \quad (3.21)$$

kde $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$.

³ V lineární algebře znamená slovo ‘lineární’ něco jiného než v matematické analýze. Např. funkci jedné proměnné $f(x) = ax + b$ znáte ze základní školy jako lineární, v lineární algebře však lineární není – je afinní. Ovšem soustavě rovnic $\mathbf{A}\mathbf{x} = \mathbf{b}$ se říká ‘lineární’ i v lineární algebře.

3.6 Cvičení

3.1. Rozhodněte, zda následující množiny tvoří lineární nebo afinní podprostor \mathbb{R}^n a případně určete jeho dimenzi:

- $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = 0 \}$ pro dané $\mathbf{a} \in \mathbb{R}^n$
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b \}$ pro dané $\mathbf{a} \in \mathbb{R}^n, b \in \mathbb{R}$
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{x} = 1 \}$
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a} \mathbf{x}^T = \mathbf{I} \}$ pro dané $\mathbf{a} \in \mathbb{R}^n$
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 0 \}$

3.2. Je dáno zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{x} \times \mathbf{y}$, kde $\mathbf{y} \in \mathbb{R}^3$ je pevný vektor a \times označuje vektorový součin. Jde tedy o zobrazení $\mathbb{R}^3 \rightarrow \mathbb{R}^3$. Je toto zobrazení lineární? Pokud ano, najděte matici \mathbf{A} tak, aby $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$. Čemu je rovno \mathbf{A}^T ? Jakou hodnotu má \mathbf{A} ?

3.3. Máme zobrazení $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definované jako $\mathbf{f}(x, y) = (x + y, 2x - 1, x - y)$. Je toto zobrazení lineární? Pokud ano, napište ho ve formě (3.5). Je toto zobrazení afinní? Pokud ano, napište ho ve formě (3.20). Obě odpovědi dokažte z definic.

3.4. Mějme nehomogenní lineární soustavu

$$\begin{aligned} x + 2y + z &= 1 \\ -x + y + 2z &= 2 \end{aligned}$$

dvou rovnic o třech neznámých. Napište množinu řešení soustavy jako $X + \mathbf{x}_0$, kde $X \subseteq \mathbb{R}^3$ je lineární podprostor a $\mathbf{x}_0 \in \mathbb{R}^3$.

3.5. Najděte bázi prostoru obrazů a bázi nulového prostoru následujících lineárních zobrazení:

- $\mathbf{f}(x_1, x_2, x_3) = (x_1 - x_2, x_2 - x_3 + 2x_1)$
- $\mathbf{f}(x_1, x_2) = (2x_1 + x_2, x_1 - x_2, 2x_2 + x_1)$

3.6. Máte matice \mathbf{A} a \mathbf{B} se stejným počtem řádků. Jak byste ověřili, zda $\text{rng } \mathbf{A} = \text{rng } \mathbf{B}$, umíte-li spočítat hodnotu libovolné matice?

3.7. Které z těchto výroků jsou pravdivé? Každý výrok dokažte nebo najděte protipříklad. Některé výroky mohou platit jen pro určité rozměry matic – najděte tedy co nejobecnější podmínky na rozměry matic, aby výroky byly pravdivé.

- Pokud $\mathbf{A}\mathbf{B}$ má plnou hodnotu, pak \mathbf{A} a \mathbf{B} mají plnou hodnotu.
- Pokud \mathbf{A} a \mathbf{B} mají plnou hodnotu, pak $\mathbf{A}\mathbf{B}$ má plnou hodnotu.
- Pokud \mathbf{A} a \mathbf{B} mají triviální nulový prostor, pak $\mathbf{A}\mathbf{B}$ má triviální nulový prostor.
- (*) Pokud \mathbf{A} a \mathbf{B} jsou úzké s plnou hodnotou a platí $\mathbf{A}^T \mathbf{B} = \mathbf{0}$, pak matice $[\mathbf{A} \ \mathbf{B}]$ je úzká s plnou hodnotou.
- (*) Pokud matice $\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$ má plnou hodnotu, pak \mathbf{A} i \mathbf{B} mají plnou hodnotu.

3.8. Necht' X je lineární podprostor \mathbb{R}^n a $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je lineární zobrazení. Je množina $\mathbf{f}(X)$ lineární podprostor \mathbb{R}^m ? Odpověď dokažte.

3.9. Najdi afinní zobrazení $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ které zobrazí trojúhelník s vrcholy $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3 \in \mathbb{R}^2$ na trojúhelník s vrcholy $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3 \in \mathbb{R}^2$. Zobrazení má zobrazit bod \mathbf{p}_1 do bodu \mathbf{q}_1 atd.

Obečněji, máme k dvojic bodů $(\mathbf{p}_i, \mathbf{q}_i), \dots, (\mathbf{p}_k, \mathbf{q}_k)$, kde $\mathbf{p}_i \in \mathbb{R}^n$ a $\mathbf{q}_i \in \mathbb{R}^m$. Kolik dvojic potřebujeme, aby existovalo jediné afinní zobrazení \mathbf{f} tak, že $\mathbf{f}(\mathbf{p}_i) = \mathbf{q}_i, i = 1, \dots, k$?

3.10. Zjistí, zda existuje lineární funkce f splňující tyto podmínky:

- $f(1, 2) = 2, f(3, 4) = 3.$
- $f(1, 2) = 2, f(3, 4) = 3, f(5, 6) = 4.$
- $f(1, 0, 1) = -1, f(0, 1, 2) = 1, f(1, 1, 3) = 2.$

3.11. K dané matici \mathbf{A} hledáme libovolný nenulový vektor \mathbf{x} takový, že $\mathbf{A}\mathbf{x} = \mathbf{0}$.

- Navrhnete verzi Gaussovy eliminační metody, která takové \mathbf{x} najde, příp. pozná, že takové \mathbf{x} neexistuje.
- Na základě toho navrhnete algoritmus, který najde bázi nulového prostoru matice \mathbf{A} .

3.12. Dokažte, že každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ hodnosti r lze psát jako $\mathbf{A} = \mathbf{B}\mathbf{C}$, kde $\mathbf{B} \in \mathbb{R}^{m \times r}$ a $\mathbf{C} \in \mathbb{R}^{r \times n}$.

3.13. Dokažte, že pro libovolné dvě matice platí $\text{rank}(\mathbf{A}\mathbf{B}) \leq \min\{\text{rank } \mathbf{A}, \text{rank } \mathbf{B}\}$.

3.14. (*) Mějme vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$. Vektor \mathbf{x}_i nazveme *klíčový*, je-li lineárně nezávislý na předchozích vektorech $\mathbf{x}_1, \dots, \mathbf{x}_{i-1}$. Dokažte, že množina klíčových vektorů je báze podprostoru $\text{span } \mathbf{x}_1, \dots, \mathbf{x}_k$.

Nápověda a řešení

3.1.a) Lineární podprostor, dimenze $n - 1$ pro $\mathbf{a} \neq \mathbf{0}$ a n pro $\mathbf{a} = \mathbf{0}$.

3.1.b) Afinní podprostor dimenze $n - 1$ pro $\mathbf{a} \neq \mathbf{0}$ a n pro $\mathbf{a} = \mathbf{0}, b = 0$. Pro $\mathbf{a} = \mathbf{0}, b \neq 0$ je množina prázdná (tedy není afinní podprostor).

3.1.c) Není lineární ani afinní podprostor (je to sféra).

3.1.d) Pro $n = 1$ je množinou jediný bod, tedy afinní podprostor prostoru \mathbb{R} . Pro $n > 1$ je množina prázdná (tedy není afinní podprostor), protože soustava $\mathbf{a}\mathbf{x}^T = \mathbf{I}$ nemá řešení pro žádné \mathbf{a}, \mathbf{x} (možný důkaz: je $\text{rank } \mathbf{I} = n$, ale $\text{rank}(\mathbf{a}\mathbf{x}^T) \leq 1$).

3.1.e) Lineární podprostor dimenze $n - 1$.

3.2. Je lineární, $\mathbf{A} = \begin{bmatrix} 0 & y_3 & -y_2 \\ -y_3 & 0 & y_1 \\ y_2 & -y_1 & 0 \end{bmatrix}$ je antisymetrická hodnosti 2.

3.3. Je afinní. Je $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, kde $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 1 & -1 \end{bmatrix}$ $\mathbf{b} = (0, -1, 0) = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{x} = (x, y) = \begin{bmatrix} x \\ y \end{bmatrix}$.

3.4. Např. $(1, -1, 2) + \text{span}\{(1, -1, 1)\} = \{(1 + \alpha, -1 - \alpha, 2 + \alpha) \mid \alpha \in \mathbb{R}\}$.

3.5.a) $\text{rng } \mathbf{A} = \mathbb{R}^2$ a tedy báze $\text{rng } \mathbf{A}$ je např. \mathbf{I}_2 . Báze $\text{null } \mathbf{A}$ je např. $(1, 1, 3)$.

3.5.b) Báze $\text{rng } \mathbf{A}$ je např. $(2, 1, 1), (1, -1, 2)$. Báze $\text{null } \mathbf{A}$ je např. $(0, 0)$.

3.6. $\text{rng } \mathbf{A} = \text{rng } \mathbf{B}$ je ekvivalentní $\text{rank } \mathbf{A} = \text{rank}[\mathbf{A} \ \mathbf{B}] = \text{rank } \mathbf{B}$.

3.9. Řešíme soustavu $\mathbf{q}_i = \mathbf{A}\mathbf{p}_i + \mathbf{b}, i = 1, \dots, k$, pro \mathbf{A}, \mathbf{b} .

3.10.a) Je $f(x_1, x_2) = a_1 x_1 + a_2 x_2$. Řešíme soustavu $a_1 + 2a_2 = 2, 3a_1 + 4a_2 = 3$. Tato soustava má řešení, tedy lineární funkce existuje.

3.10.b) Ano.

3.10.c) Ne.

3.11.b) Hledejte vektory báze jeden po druhém.

- 3.12. Necht' sloupce \mathbf{B} tvoří bázi podprostoru $\text{rng } \mathbf{A}$. Pak každý sloupec \mathbf{a}_j matice \mathbf{A} je lineární kombinací sloupců \mathbf{B} , tedy $\mathbf{a}_j = \mathbf{B}\mathbf{c}_j$. Vektory \mathbf{c}_j tvoří sloupce matice \mathbf{C} .
- 3.13. Z Věty 3.4 je $\text{rank}(\mathbf{AB}) \leq \text{rank } \mathbf{A}$. Dle (3.12) je $\text{rank}(\mathbf{AB}) = \text{rank}((\mathbf{AB})^T) = \text{rank}(\mathbf{B}^T \mathbf{A}^T) \leq \text{rank}(\mathbf{B}^T) = \text{rank } \mathbf{B}$.

Kapitola 4

Ortogonalita

4.1 Standardní skalární součin

Prostor \mathbb{R}^n je přirozeně vybaven **standardním skalárním součinem**

$$\mathbf{x}^T \mathbf{y} = x_1 y_1 + \dots + x_n y_n = \mathbf{y}^T \mathbf{x}.$$

Skalární součin splňuje **Cauchy-Schwarzovu nerovnost** $(\mathbf{x}^T \mathbf{y})^2 \leq (\mathbf{x}^T \mathbf{x})(\mathbf{y}^T \mathbf{y})$.

Standardní skalární součin indukuje **eukleidovskou normu**

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}} = (x_1^2 + \dots + x_n^2)^{1/2}, \quad (4.1)$$

Norma splňuje **trojúhelníkovou nerovnost** $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$, která snadno plyne z Cauchyovy-Schwarzovy nerovnosti (umocněte a roznásobte). Norma měří *délku* vektoru \mathbf{x} .

Úhel φ dvojice vektorů se spočítá jako

$$\cos \varphi = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}. \quad (4.2)$$

Speciálně, vektory jsou **ortogonální** neboli **kolmé**, jestliže $\mathbf{x}^T \mathbf{y} = 0$, což značíme také $\mathbf{x} \perp \mathbf{y}$.

Eukleidovská norma indukuje **eukleidovskou metriku**

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|, \quad (4.3)$$

která měří *vzdálenost* bodů \mathbf{x} a \mathbf{y} .

Protože pro $n = 3$ takto definované pojmy délky, úhlu a vzdálenosti dobře modelují prostor, ve kterém žijeme, prostoru \mathbb{R}^n se standardním skalárním součinem se často říká **Eukleidovský prostor**.

4.2 Ortogonální podprostory

Vektor $\mathbf{y} \in \mathbb{R}^n$ je **ortogonální** na podprostor $X \subseteq \mathbb{R}^n$ (což značíme $\mathbf{y} \perp X$ nebo $X \perp \mathbf{y}$), je-li $\mathbf{y} \perp \mathbf{x}$ pro každé $\mathbf{x} \in X$. Pro testování této podmínky stačí ověřit, že \mathbf{y} je kolmý na každý bázový vektor podprostoru X , neboť (dokažte!)

$$\mathbf{y} \perp \text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \iff \mathbf{y} \perp \mathbf{x}_1, \dots, \mathbf{y} \perp \mathbf{x}_k. \quad (4.4)$$

Podprostory $X, Y \subseteq \mathbb{R}^n$ jsou **ortogonální** (značíme $X \perp Y$), je-li $\mathbf{x} \perp \mathbf{y}$ pro každé $\mathbf{x} \in X$ a $\mathbf{y} \in Y$, neboli $\mathbf{y} \perp X$ pro každé $\mathbf{y} \in Y$. Platí

$$X \perp Y \implies X \cap Y = \{\mathbf{0}\}, \quad (4.5)$$

neboť jediný vektor kolmý sám na sebe je $\mathbf{0}$.

Ortogonalní doplněk podprostoru $X \subseteq \mathbb{R}^n$ je množina

$$X^\perp = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{y} \perp X\} \quad (4.6)$$

všech vektorů z \mathbb{R}^n kolmých na podprostor X . Množina (4.6) je podprostor \mathbb{R}^n (dokažte!).

Věta 4.1. Pro dva podprostory $X, Y \subseteq \mathbb{R}^n$ platí

$$Y = X^\perp \iff X \perp Y, \quad \dim X + \dim Y = n. \quad (4.7)$$

Z této věty snadno plyne, že $(Y = X^\perp) \Leftrightarrow (X = Y^\perp)$. Proto má smysl říkat, že podprostory X a Y jsou ortogonálním doplňkem jeden druhého. Z posledního dále plyne

$$(X^\perp)^\perp = X.$$

Příklad 4.1. Dvě na sebe kolmé přímky v \mathbb{R}^3 procházející počátkem jsou ortogonální podprostory, nejsou ale ortogonální doplněk jeden druhého. Ortogonální doplněk k přímce v \mathbb{R}^3 procházející počátkem je rovina procházející počátkem, která je na tuto přímku kolmá. \square

Příklad 4.2. Mějme podprostor $X = \text{span}\{(1, 2, 3), (0, 1, -1)\}$. Jeho ortogonální doplněk X^\perp je množina všech vektorů (x_1, x_2, x_3) splňujících soustavu rovnic

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ x_2 - x_3 &= 0. \end{aligned} \quad \square$$

4.2.1 Čtyři podprostory generované maticí

Vraťme se do §3.2. Každá matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ generuje čtyři **základní podprostory**:

- $\text{rng } \mathbf{A} = \{\mathbf{Ax} \mid \mathbf{x} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$ je lineární obal sloupců \mathbf{A} .
- $\text{null } \mathbf{A} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{0}\} \subseteq \mathbb{R}^n$ je množina všech vektorů kolmých na řádky \mathbf{A} .
- $\text{rng}(\mathbf{A}^T) = \{\mathbf{A}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^m\} \subseteq \mathbb{R}^n$ je lineární obal řádků \mathbf{A} .
- $\text{null}(\mathbf{A}^T) = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{A}^T\mathbf{x} = \mathbf{0}\} \subseteq \mathbb{R}^m$ je množina všech vektorů kolmých na sloupce \mathbf{A} .

Věta 4.2. Pro každou matici \mathbf{A} platí

$$(\text{rng } \mathbf{A})^\perp = \text{null}(\mathbf{A}^T). \quad (4.8a)$$

$$(\text{null } \mathbf{A})^\perp = \text{rng}(\mathbf{A}^T), \quad (4.8b)$$

Důkaz. Rovnost (4.8a) plyne ihned z (4.4): jsou-li $\mathbf{a}_1, \dots, \mathbf{a}_n$ sloupce matice \mathbf{A} , je

$$(\text{rng } \mathbf{A})^\perp = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{x} \perp \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}\} = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{a}_1^T\mathbf{x} = \dots = \mathbf{a}_n^T\mathbf{x} = 0\} = \text{null}(\mathbf{A}^T).$$

Rovnost (4.8b) se získá použitím rovnosti (4.8a) na matici \mathbf{A}^T a rovnosti $(X^\perp)^\perp = X$. \square

4.3 Ortonormální množina vektorů

Vektor \mathbf{u} nazveme **normalizovaný**, pokud má jednotkovou délku ($\|\mathbf{u}\| = 1 = \mathbf{u}^T\mathbf{u}$). Množinu vektorů $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ nazveme **ortonormální**, pokud každý vektor z této množiny je normalizovaný a každá dvojice vektorů z této množiny je ortogonální, tedy

$$\mathbf{u}_i^T\mathbf{u}_j = \delta_{ij} = \begin{cases} 0 & \text{když } i \neq j, \\ 1 & \text{když } i = j. \end{cases} \quad (4.9)$$

Ortonormální množina vektorů je lineárně nezávislá. Pro důkaz vynásobme levou stranu implikace (3.2) skalárně vektorem \mathbf{u}_i , což dá

$$0 = \mathbf{u}_i^T\mathbf{0} = \alpha_1\mathbf{u}_i^T\mathbf{u}_1 + \dots + \alpha_n\mathbf{u}_i^T\mathbf{u}_n = \alpha_i\mathbf{u}_i^T\mathbf{u}_i = \alpha_i.$$

tedy $\alpha_i = 0$. Když toto uděláme pro každé i , máme $\alpha_1 = \dots = \alpha_n = 0$.

Nechť sloupce matice $\mathbf{U} \in \mathbb{R}^{m \times n}$ tvoří ortonormální množinu vektorů. Protože ortonormální vektory jsou lineárně nezávislé, je nutné $m \geq n$. Podmínku ortonormality (4.9) sloupců matice \mathbf{U} lze psát stručně jako

$$\mathbf{U}^T\mathbf{U} = \mathbf{I}. \quad (4.10)$$

Lineární zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{U}\mathbf{x}$ (zobrazení z \mathbb{R}^n do \mathbb{R}^m) zachovává skalární součin, neboť

$$\mathbf{f}(\mathbf{x})^T\mathbf{f}(\mathbf{y}) = (\mathbf{U}\mathbf{x})^T(\mathbf{U}\mathbf{y}) = \mathbf{x}^T\mathbf{U}^T\mathbf{U}\mathbf{y} = \mathbf{x}^T\mathbf{y}. \quad (4.11)$$

Pro $\mathbf{x} = \mathbf{y}$ dostaneme, že se zachovává také eukleidovská norma, $\|\mathbf{f}(\mathbf{x})\| = \|\mathbf{U}\mathbf{x}\| = \|\mathbf{x}\|$. Tedy zobrazení \mathbf{f} zachovává délky a úhly. Taková zobrazení se nazývají **isometrie**.

Pokud je matice \mathbf{U} čtvercová ($m = n$), následující podmínky jsou ekvivalentní:

$$\mathbf{U}^T\mathbf{U} = \mathbf{I} \iff \mathbf{U}^T = \mathbf{U}^{-1} \iff \mathbf{U}\mathbf{U}^T = \mathbf{I}. \quad (4.12)$$

Důkaz není těžký. Protože sloupce \mathbf{U} jsou ortonormální, jsou lineárně nezávislé a \mathbf{U} je regulární. Vynásobením levé rovnice maticí \mathbf{U}^{-1} zprava získáme prostřední rovnici. Vynásobením prostřední rovnice maticí \mathbf{U} zleva získáme pravou rovnici. Zbylé implikace dokážeme analogicky.

Ekvivalence (4.12) říká, že má-li čtvercová matice ortonormální sloupce, má ortonormální i řádky. Dále říká, že inverze takové matice se spočítá jednoduše transpozicí. Čtvercové matici splňující podmínky (4.12) se říká **ortogonální matice**.

Zdůrazněme, že pokud \mathbf{U} je obdélníková s ortonormálními sloupci, neplatí $\mathbf{U}\mathbf{U}^T = \mathbf{I}$. Dále, pokud má \mathbf{U} ortogonální (ne však ortonormální) sloupce, nemusí mít ortogonální řádky¹.

Nechť \mathbf{U} je ortogonální matice. Vezmeme-li determinant obou stran rovnice $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, máme $\det(\mathbf{U}^T\mathbf{U}) = \det(\mathbf{U}^T)\det\mathbf{U} = (\det\mathbf{U})^2 = 1$. Tedy $\det\mathbf{U}$ může nabývat dvou hodnot:

- Pokud $\det\mathbf{U} = 1$, matici se říká **speciální ortogonální** nebo také **rotační**, protože zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{U}\mathbf{x}$ (což je zobrazení z \mathbb{R}^n do sebe) znamená *otočení* vektoru \mathbf{x} okolo počátku. Každou rotaci v prostoru \mathbb{R}^n lze jednoznačně reprezentovat rotační maticí.
- Pokud $\det\mathbf{U} = -1$, zobrazení \mathbf{f} je složením otočení a *zrcadlení* (reflexe) kolem nadroviny procházející počátkem.

¹To je možná důvod, proč se čtvercové matici s ortonormálními sloupci (tedy i řádky) neříká 'ortonormální' ale 'ortogonální'. Obdélníková matice s ortonormálními sloupci a čtvercová matice s ortogonálními (ne však ortonormálními) sloupci zvláštní jména nemají.

Příklad 4.3. Všechny rotační matice 2×2 lze napsat jako

$$\mathbf{U} = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$$

pro nějaké φ . Násobení vektoru touto maticí odpovídá otočení vektoru v rovině o úhel φ . Zkontrolujte si, že $\mathbf{U}^T \mathbf{U} = \mathbf{I} = \mathbf{U} \mathbf{U}^T$ a $\det \mathbf{U} = 1$. \square

Příklad 4.4. Zrcadlení (neboli reflexe) v \mathbb{R}^2 kolem přímky procházející počátkem a směrnicí $\tan(\varphi/2)$ je reprezentováno ortogonální maticí

$$\mathbf{U} = \begin{bmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{bmatrix}. \quad \square$$

Příklad 4.5. Permutační matice je čtvercová matice, jejíž sloupce jsou permutované vektory standardní báze. Např.

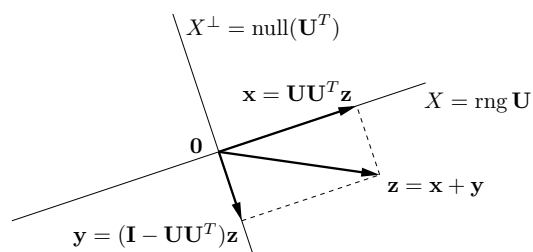
$$[\mathbf{e}_3 \ \mathbf{e}_1 \ \mathbf{e}_2] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Permutační matice je ortogonální (dokažte!) a její determinant je rovný znaménku permutace. \square

Rozdělíme-li sloupce ortogonální matice do dvou bloků, generuje to rozklad prostora na podprostor a jeho ortogonální doplněk. Necht' $[\mathbf{U} \ \bar{\mathbf{U}}] \in \mathbb{R}^{n \times n}$ je ortogonální matice složená ze dvou bloků $\mathbf{U} \in \mathbb{R}^{n \times n}$ a $\bar{\mathbf{U}} \in \mathbb{R}^{n \times (n-m)}$. Pak samozřejmě $\mathbf{U}^T \mathbf{U} = \mathbf{I}$ a $\bar{\mathbf{U}}^T \bar{\mathbf{U}} = \mathbf{I}$. Dále $\mathbf{U}^T \bar{\mathbf{U}} = \mathbf{0}$, neboli $\text{rng } \mathbf{U} \perp \text{rng } \bar{\mathbf{U}}$. Ale navíc platí $(\text{rng } \mathbf{U})^\perp = \text{rng } \bar{\mathbf{U}}$, což ihned plyne z (4.7).

4.4 Ortogonální projekce

Necht' X je lineární podprostor \mathbb{R}^n . Ukážeme, že pro každý vektor $\mathbf{z} \in \mathbb{R}^n$ existuje právě jedno $\mathbf{x} \in X$ a právě jedno $\mathbf{y} \in X^\perp$ tak, že $\mathbf{z} = \mathbf{x} + \mathbf{y}$. Vektor \mathbf{x} nazýváme **ortogonální projekce** vektoru \mathbf{z} na podprostor X a, symetricky, vektor \mathbf{y} je ortogonální projekce vektoru \mathbf{z} na podprostor X^\perp . Všimněte si, že podmínku $\mathbf{y} \in X^\perp$ lze psát jako $(\mathbf{z} - \mathbf{x}) \in X^\perp$ neboli $\mathbf{y} \perp X$. Tedy vektor \mathbf{x} je určen podmínkami $\mathbf{x} \in X$ a $(\mathbf{z} - \mathbf{x}) \perp X$.



Protože (jak dokážeme v §4.5) každý podprostor má ortonormální bázi, můžeme bez újmy na obecnosti reprezentovat podprostor X ortonormální bází. Necht' tato báze tvoří sloupce matice $\mathbf{U} \in \mathbb{R}^{n \times m}$, tedy $X = \text{rng } \mathbf{U}$ a $\mathbf{U}^T \mathbf{U} = \mathbf{I}$. Pak $\mathbf{x} \in X$ právě tehdy, je-li $\mathbf{x} = \mathbf{U} \mathbf{t}$ pro

nějaké $\mathbf{t} \in \mathbb{R}^m$. Podmínku $(\mathbf{z} - \mathbf{x}) \perp X$ lze nyní psát jako $\mathbf{U}^T (\mathbf{z} - \mathbf{U} \mathbf{t}) = \mathbf{0}$. Tato rovnice má právě jedno řešení $\mathbf{t} = \mathbf{U}^T \mathbf{z}$ (ověřte!). Tedy pro každý vektor $\mathbf{z} \in \mathbb{R}^n$ existuje právě jedna ortogonální projekce na podprostor $\text{rng } \mathbf{U}$, a to

$$\mathbf{x} = \mathbf{U} \mathbf{U}^T \mathbf{z} = (\mathbf{u}_1^T \mathbf{z}) \mathbf{u}_1 + \cdots + (\mathbf{u}_m^T \mathbf{z}) \mathbf{u}_m. \quad (4.13)$$

Zde jsme zároveň uvedli i vektorovou formu vzorce, kde $\mathbf{u}_1, \dots, \mathbf{u}_m$ jsou řádky matice \mathbf{U} (ověřte druhou rovnost v (4.13)!). Ortogonální projekce na podprostor $\text{rng } \mathbf{U}$ je tedy lineární zobrazení reprezentované maticí

$$\mathbf{P} = \mathbf{U} \mathbf{U}^T, \quad (4.14)$$

které se proto říká ortogonální **projektor** na podprostor X . Máme-li projekci na podprostor X , projekce na podprostor X^\perp je

$$\mathbf{y} = \mathbf{z} - \mathbf{x} = \mathbf{z} - \mathbf{P} \mathbf{z} = (\mathbf{I} - \mathbf{P}) \mathbf{z}. \quad (4.15)$$

Tedy je-li \mathbf{P} projektor na X , tak $\mathbf{I} - \mathbf{P}$ je projektor na X^\perp .

Je zajímavé ukázat ještě jiný pohled na (4.15). Je-li $[\mathbf{U} \ \bar{\mathbf{U}}] \in \mathbb{R}^{n \times n}$ ortogonální matice složená ze dvou bloků $\mathbf{U} \in \mathbb{R}^{n \times m}$ a $\bar{\mathbf{U}} \in \mathbb{R}^{n \times (n-m)}$, potom $(\text{rng } \mathbf{U})^\perp = \text{rng } \bar{\mathbf{U}}$ (viz §4.3). Navíc

$$[\mathbf{U} \ \bar{\mathbf{U}}] [\mathbf{U} \ \bar{\mathbf{U}}]^T = \mathbf{U} \mathbf{U}^T + \bar{\mathbf{U}} \bar{\mathbf{U}}^T = \mathbf{I}, \quad (4.16)$$

což plyne z násobení blokových matic (viz §2.7) a z (4.12). Projektor na podprostor $X = \text{rng } \mathbf{U}$ je $\mathbf{P} = \mathbf{U} \mathbf{U}^T$ a tedy projektor na podprostor $X^\perp = \text{rng } \bar{\mathbf{U}}$ je $\bar{\mathbf{U}} \bar{\mathbf{U}}^T = \mathbf{I} - \mathbf{U} \mathbf{U}^T = \mathbf{I} - \mathbf{P}$.

Když $n = 1$, matice \mathbf{U} má jediný sloupec \mathbf{u} splňující $\mathbf{u}^T \mathbf{u} = 1$ a podprostor X je přímka. Projektor je tedy $\mathbf{P} = \mathbf{u} \mathbf{u}^T$, což je matice hodnoty jedna (viz §2.6). Vzoreček

$$\mathbf{x} = \mathbf{P} \mathbf{z} = \mathbf{u} \mathbf{u}^T \mathbf{z} = (\mathbf{u}^T \mathbf{z}) \mathbf{u} \quad (4.17)$$

pro² průmět vektoru \mathbf{z} na normalizovaný vektor \mathbf{u} máte znát ze střední školy: skalární součin $\alpha = \mathbf{u}^T \mathbf{z}$ je délka (se znaménkem) průmětu a $\alpha \mathbf{u}$ je vektor o této délce ve směru vektoru \mathbf{u} .

Uved'me některé vlastnosti ortogonálního projektoru. Z (4.14) plyne $\mathbf{P}^2 = \mathbf{P}$ a $\mathbf{P} = \mathbf{P}^T$. Z Vět 3.4 a 3.6 plyne $\text{rng } \mathbf{P} = X$ a $\text{null } \mathbf{P} = X^\perp$, což je ovšem geometricky jasné přímo z definice ortogonální projekce.

Poznámka. Obecnou *projekcí* se v lineární algebře rozumí každé lineární zobrazení $\mathbf{f}(\mathbf{z}) = \mathbf{P} \mathbf{z}$, které splňuje $\mathbf{f}(\mathbf{f}(\mathbf{z})) = \mathbf{f}(\mathbf{z})$, tedy $\mathbf{P}^2 = \mathbf{P}$ (takové zobrazení se nazývá *idempotentní*). To vyjadřuje vlastnost, že když jednou vektor promítneme, tak další promítnutí na stejný podprostor ho již nezmění. Projekce obecně nemusí být ortogonální, může být šikmá – pak promítáme ve směru podprostoru $\text{null } \mathbf{P}$ na podprostor $\text{rng } \mathbf{P}$. Projekce je ortogonální právě když³ $\text{null } \mathbf{P} \perp \text{rng } \mathbf{P}$. To nastane právě když kromě $\mathbf{P}^2 = \mathbf{P}$ platí navíc $\mathbf{P}^T = \mathbf{P}$ (důkaz tohoto tvrzení vynecháme). \square

Ortogonální projekce má důležitou vlastnost: \mathbf{x} je bod v podprostoru X nejbližší bodu \mathbf{z} :

Věta 4.3. Je-li $\mathbf{x} \in X$ ortogonální projekce vektoru \mathbf{z} na podprostor X , pak pro každé $\mathbf{x}' \in X$ platí $\|\mathbf{z} - \mathbf{x}\| \leq \|\mathbf{z} - \mathbf{x}'\|$.

²Závorka ve výrazu $(\mathbf{u}^T \mathbf{z}) \mathbf{u}$ je nutná, protože součin výrazů $\mathbf{u}^T \mathbf{z}$ a \mathbf{u} není maticový součin, ale násobení vektoru skalárem. Viz poznámka v §2.1.

³Pro obecnou čtvercovou maticí samozřejmě neplatí, že její nulový prostor a prostor obrazů jsou navzájem ortogonální, tím méně ortogonálním doplňkem. Neplést se vztahy (4.8)!

Důkaz. Protože X je podprostor a $\mathbf{x}, \mathbf{x}' \in X$, je $\mathbf{x} - \mathbf{x}' \in X$. Protože $(\mathbf{z} - \mathbf{x}) \perp X$, platí $(\mathbf{z} - \mathbf{x}) \perp (\mathbf{x} - \mathbf{x}')$. Z Pythagorovy věty tedy $\|\mathbf{z} - \mathbf{x}'\|^2 = \|\mathbf{z} - \mathbf{x}\|^2 + \|\mathbf{x} - \mathbf{x}'\|^2 \geq \|\mathbf{z} - \mathbf{x}\|^2$. \square

Délka projekce bodu \mathbf{z} na podprostor $\text{rng } \mathbf{U}$ je

$$\|\mathbf{x}\| = \|\mathbf{U}\mathbf{U}^T\mathbf{z}\| = \|\mathbf{U}^T\mathbf{z}\|,$$

kde rovnost plyne z toho, že zobrazení reprezentované maticí \mathbf{U} zachovává eukleidovskou normu (viz §4.3). Tato délka je zároveň rovna vzdálenosti bodu \mathbf{z} od podprostoru $(\text{rng } \mathbf{U})^\perp = \text{null}(\mathbf{U}^T)$.

4.5 Gram-Schmidtova ortonormalizace

Gram-Schmidtova ortonormalizace je algoritmus, který pro dané lineárně nezávislé vektory $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ najde vektory $\mathbf{q}_1, \dots, \mathbf{q}_n \in \mathbb{R}^m$ takové, že

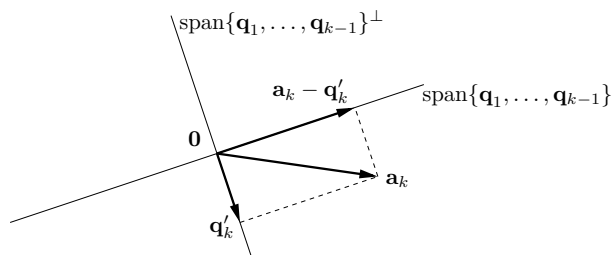
- $\mathbf{q}_1, \dots, \mathbf{q}_n$ jsou ortonormální,
- pro každé $k = 1, \dots, n$ platí $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_k\} = \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$.

Všimněte si, že algoritmus dokazuje, že každý podprostor má ortonormální bázi.

Myšlenka algoritmu je následující. Předpokládejme, že již máme vektory $\mathbf{q}_1, \dots, \mathbf{q}_{k-1}$ s požadovanými vlastnostmi. Vektor \mathbf{q}_k spočítáme tak, že spočítáme ortogonální projekci vektoru \mathbf{a}_k na podprostor $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}^\perp$ a výsledek znormalizujeme. Tedy

$$\mathbf{q}'_k = \mathbf{a}_k - \sum_{i=1}^{k-1} (\mathbf{q}_i^T \mathbf{a}_k) \mathbf{q}_i, \quad \mathbf{q}_k = \frac{\mathbf{q}'_k}{\|\mathbf{q}'_k\|}. \quad (4.18)$$

Protože vektor \mathbf{q}'_k je projekce na $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}^\perp$, je kolmý na vektory $\mathbf{q}_1, \dots, \mathbf{q}_{k-1}$. Dle předpokladu máme $\mathbf{a}_k \notin \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{k-1}\} = \text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}$, proto $\mathbf{q}'_k \neq \mathbf{0}$ a $\|\mathbf{q}'_k\| \neq 0$. Tedy vektor \mathbf{q}_k existuje a má požadované vlastnosti. Viz obrázek:



Celý algoritmus provede iteraci (4.18) pro $k = 1, \dots, n$. Zde je běh algoritmu pro $n = 3$:

$$\begin{aligned} \mathbf{q}'_1 &= \mathbf{a}_1, & \mathbf{q}_1 &= \mathbf{q}'_1 / \|\mathbf{q}'_1\| \\ \mathbf{q}'_2 &= \mathbf{a}_2 - (\mathbf{q}_1^T \mathbf{a}_2) \mathbf{q}_1, & \mathbf{q}_2 &= \mathbf{q}'_2 / \|\mathbf{q}'_2\| \\ \mathbf{q}'_3 &= \mathbf{a}_3 - (\mathbf{q}_1^T \mathbf{a}_3) \mathbf{q}_1 - (\mathbf{q}_2^T \mathbf{a}_3) \mathbf{q}_2, & \mathbf{q}_3 &= \mathbf{q}'_3 / \|\mathbf{q}'_3\| \end{aligned}$$

Algoritmus v uvedeném tvaru není stabilní vůči zaokrouhlovacím chybám. Lze jej ale poměrně snadnou změnou (kterou neuvádíme) učinit stabilní.

4.5.1 QR rozklad

Věta 4.4. Každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ kde $m \geq n$ lze rozložit na součin

$$\mathbf{A} = \mathbf{Q}\mathbf{R}, \quad (4.19)$$

kde matice $\mathbf{Q} \in \mathbb{R}^{m \times n}$ má ortonormální sloupce a matice $\mathbf{R} \in \mathbb{R}^{n \times n}$ je horní trojúhelníková.

Věta se dokáže z Gram-Schmidtovy ortogonalizace. Jelikož jsme ale Gram-Schmidtovu ortogonalizaci uvedli jen pro případ lineárně nezávislých vektorů, větu dokážeme jen pro případ, kdy matice \mathbf{A} má lineárně nezávislé sloupce. Existuje ale modifikace Gram-Schmidtova algoritmu, která spočítá QR rozklad i pro lineárně závislé vstupní vektory.

Důkaz. Rovnosti (4.18) lze napsat jako

$$\mathbf{a}_k = \sum_{i=1}^{k-1} r_{ik} \mathbf{q}_i + r_{kk} \mathbf{q}_k = \sum_{i=1}^k r_{ik} \mathbf{q}_i, \quad (4.20)$$

kde $r_{ik} = \mathbf{q}_i^T \mathbf{a}_k$ pro $i < k$ a $r_{kk} = \|\mathbf{q}'_k\|$. Soustava rovnic (4.20) pro $k = 1, \dots, n$ se v maticovém tvaru napíše jako $\mathbf{A} = \mathbf{Q}\mathbf{R}$, kde $\mathbf{a}_1, \dots, \mathbf{a}_n$ jsou sloupce matice \mathbf{A} , $\mathbf{q}_1, \dots, \mathbf{q}_n$ jsou sloupce matice \mathbf{Q} a r_{ik} jsou prvky matice \mathbf{R} . \square

Matici \mathbf{Q} lze přidáním sloupců doplnit na čtvercovou ortogonální matici $[\mathbf{Q} \ \bar{\mathbf{Q}}] \in \mathbb{R}^{m \times m}$. Toto doplnění lze udělat např. vylepšením Gram-Schmidtova algoritmu, které zde ale nepopíšeme. Matici \mathbf{R} pak doplníme přidáním nulových řádků na odpovídající rozměr. Tedy

$$\mathbf{A} = [\mathbf{Q} \ \bar{\mathbf{Q}}] \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix} = \mathbf{Q}'\mathbf{R}'. \quad (4.21)$$

Rozkladu $\mathbf{A} = \mathbf{Q}\mathbf{R}$ se pak říká *redukovaný QR rozklad* a rozkladu $\mathbf{A} = \mathbf{Q}'\mathbf{R}'$ *plný QR rozklad*. Jelikož matice $[\mathbf{Q} \ \bar{\mathbf{Q}}]$ je ortogonální, je $(\text{rng } \mathbf{Q})^\perp = \text{rng } \bar{\mathbf{Q}}$. V Matlabu je plný QR rozklad implementován příkazem $[\mathbf{Q}, \mathbf{R}] = \text{qr}(\mathbf{A})$ a redukovaný příkazem $[\mathbf{Q}, \mathbf{R}] = \text{qr}(\mathbf{A}, 0)$.

QR rozklad má mnoho aplikací. Typické je užití na řešení lineárních soustav. Řešíme soustavu $\mathbf{A}\mathbf{x} = \mathbf{b}$. Necht' $\mathbf{A} = \mathbf{Q}\mathbf{R}$ je plný QR rozklad. Vynásobením soustavy zleva maticí \mathbf{Q}^T dá

$$\mathbf{R}\mathbf{x} = \mathbf{Q}^T\mathbf{b}. \quad (4.22)$$

Toto je ekvivalentní úprava, neboť \mathbf{Q} je regulární. Ale protože je \mathbf{R} trojúhelníková, soustavu můžeme řešit zpětnou substitucí.

4.6 Skalární součin a norma matic

Je přirozené definovat **skalární součin matic** $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ jako

$$\mathbf{A} \cdot \mathbf{B} = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij}, \quad (4.23)$$

což je vlastně skalární součin vektorů vytvořených přerovnáním prvků každé matice do vektoru. Zjevně $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = \mathbf{A}^T \cdot \mathbf{B}^T$. Také je $\mathbf{A} \cdot \mathbf{B} = \mathbf{a}_1^T \mathbf{b}_1 + \dots + \mathbf{a}_n^T \mathbf{b}_n$, kde $\mathbf{a}_i, \mathbf{b}_i \in \mathbb{R}^m$ jsou sloupce matic \mathbf{A}, \mathbf{B} (totéž lze napsat pro řádky). Platí také (ověřte!)

$$\mathbf{A} \cdot \mathbf{B} = \text{tr}(\mathbf{A}^T \mathbf{B}), \quad (4.24)$$

proto se značení $\mathbf{A} \cdot \mathbf{B}$ často neuvádí a skalární součin se značí přímo $\text{tr}(\mathbf{A}^T \mathbf{B})$.

Skalární součin (4.23) indukuje normu matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ jako

$$\|\mathbf{A}\| = \sqrt{\mathbf{A} \cdot \mathbf{A}} = \sqrt{\text{tr}(\mathbf{A}^T \mathbf{A})}. \quad (4.25)$$

Podobně jako pro vektory, skalární součin matic je invariantní vůči transformaci isometrií: jsou-li \mathbf{U}, \mathbf{V} matice s ortonormálními sloupci, je

$$\begin{aligned} (\mathbf{UAV}^T) \cdot (\mathbf{UBV}^T) &= \text{tr}[(\mathbf{UAV}^T)^T (\mathbf{UBV}^T)] = \text{tr}(\mathbf{VA}^T \mathbf{U}^T \mathbf{UBV}^T) \\ &= \text{tr}(\mathbf{VA}^T \mathbf{BV}^T) = \text{tr}(\mathbf{A}^T \mathbf{BV}^T \mathbf{V}) = \text{tr}(\mathbf{A}^T \mathbf{B}) \\ &= \mathbf{A} \cdot \mathbf{B}, \end{aligned}$$

kde ve čtvrté rovnosti plyne z rovnosti $\text{tr}(\mathbf{CD}) = \text{tr}(\mathbf{DC})$ (viz §2.3). Tedy i norma (4.25) je invariantní vůči isometrii, $\|\mathbf{UAV}^T\| = \|\mathbf{A}\|$.

Existují i jiné maticové normy a maticové skalární součiny. Maticová norma (4.25) se nazývá **Frobeniova norma** a obvykle se značí $\|\cdot\|_F$ místo pouhého $\|\cdot\|$. Skalární součin (4.23) se také někdy nazývá Frobeniův.

4.7 Cvičení

- 4.1. Najděte bázi ortogonálního doplňku prostoru $\text{span}\{(0, 1, 1), (1, 2, 3)\}$.
- 4.2. Najděte ortonormální bázi podprostoru $\text{span}\{(1, 1, 1, -1), (2, -1, -1, 1), (-1, 2, 2, 1)\}$ pomocí QR rozkladu (použijte Matlab).
- 4.3. Dokažte, že součin ortogonálních matic je ortogonální matice.
- 4.4. Pro jaké n je matice $\text{diag}(-\mathbf{1}_n)$ (tedy diagonální matice se samými mínus jedničkami na diagonále) rotační?
- 4.5. Počet nezávislých parametrů (stupňů volnosti) ortogonální matice $n \times n$ se získá jako rozdíl počtu prvků matice (n^2) a počtu nezávislých rovnic v podmínce $\mathbf{U}^T \mathbf{U} = \mathbf{I}$. Neformálně řečeno udává, kolika 'knoflíky' můžeme nezávisle 'krotit' při rotaci v n -rozměrném prostoru. Jaké je toto číslo pro $n = 2, 3, 4$? Najděte vzorec pro obecné n .
- 4.6. Zobrazení $\mathbf{F}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ je dané vzorcem $\mathbf{F}(\mathbf{A}) = (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A})^{-1}$. Předpokládejte, že \mathbf{A} je taková, že $\mathbf{I} + \mathbf{A}$ je regulární. Dokažte, že:
 - a) Pro každou \mathbf{A} je $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A}) = (\mathbf{I} + \mathbf{A})(\mathbf{I} - \mathbf{A})$.
 - b) Pro každou \mathbf{A} je $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A})^{-1} = (\mathbf{I} + \mathbf{A})^{-1}(\mathbf{I} - \mathbf{A})$.
 - c) Pro každou antisymetrickou \mathbf{A} je $\mathbf{F}(\mathbf{A})$ ortogonální.
 - d) Pro každou ortogonální \mathbf{A} je $\mathbf{F}(\mathbf{A})$ antisymetrická.
 - e) Zobrazení \mathbf{F} je inverzí sama sebe, tedy $\mathbf{F}(\mathbf{F}(\mathbf{A})) = \mathbf{A}$ pro každé \mathbf{A} .

Před důkazy na papíře se přesvědčte v Matlabu, že tvrzení platí pro náhodné matice.

- 4.7. Pokud $\|\mathbf{x}\| = \|\mathbf{y}\|$, dokažte, že $(\mathbf{x} + \mathbf{y}) \perp (\mathbf{x} - \mathbf{y})$. V prostoru \mathbb{R}^2 nakreslete vektory $\mathbf{x} + \mathbf{y}$ a $\mathbf{x} - \mathbf{y}$.
- 4.8. Mějme vektory $\mathbf{x}_1 = (1, -1, 0, 2)$, $\mathbf{x}_2 = (1, 1, 1, 0)$, $\mathbf{x}_3 = (-1, -1, 2, 0)$. Ověřte, že vektory jsou vzájemně ortogonální. Najděte nenulový vektor \mathbf{x}_4 , který je ortogonální na vektory $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$.

4.9. Najděte dva ortogonální vektory \mathbf{x}, \mathbf{y} takové, že $\text{span}\{\mathbf{x}, \mathbf{y}\} = \text{span}\{(0, 1, 1), (1, 2, 3)\}$.

$$4.10. \text{ Spočítejte co nejjednodušším způsobem inverzi matice } \frac{1}{3} \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}.$$

4.11. Necht' X, Y jsou podprostory \mathbb{R}^n . Definujme $X + Y = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in X, \mathbf{y} \in Y\}$. Dokažte:

- a) $X \subseteq Y \implies X^\perp \supseteq Y^\perp$.
- b) Necht' $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ je báze podprostoru X a $\{\mathbf{y}_1, \dots, \mathbf{y}_l\}$ je báze podprostoru Y . Pak $X + Y = \text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_l\}$.
- c) $(X + Y)^\perp = X^\perp \cap Y^\perp$.
- d) $(X \cap Y)^\perp = X^\perp + Y^\perp$.

4.12. Jak byste levně spočítali absolutní hodnotu determinantu matice z jejího QR rozkladu?

4.13. Pro $\mathbf{V}^T \mathbf{V} = \mathbf{I}$ reprezentuje matice $\mathbf{H} = \mathbf{I} - 2\mathbf{V}\mathbf{V}^T$ zrcadlení (reflexi) vzhledem k podprostoru $X = (\text{rng } \mathbf{V})^\perp$, proto této matici budeme říkat *reflektor*.

- a) Odvoďte vzorec pro reflektor geometrickou úvahou, podobnou jako pro projektor.
- b) Dokažte, že $\mathbf{H} = \mathbf{H}^T = \mathbf{H}^{-1}$.
- c) Co je $\mathbf{H}\mathbf{x}$ když $\mathbf{x} \in X$? Ukažte algebraicky a odůvodněte geometricky.
- d) Co je $\mathbf{H}\mathbf{x}$ když $\mathbf{x} \perp X$? Ukažte algebraicky a odůvodněte geometricky.

Poznamenejme, že známější je případ, kdy \mathbf{V} má jediný sloupec, tedy $\mathbf{H} = \mathbf{I} - 2\mathbf{v}\mathbf{v}^T$ pro $\mathbf{v}^T \mathbf{v} = 1$. Tato matice reprezentuje zrcadlení kolem nadroviny s normálovým vektorem \mathbf{v} a je známa jako *elementární reflektor* nebo *Householderova matice*.

4.14. (*) *RQ rozklad* rozloží matici $\mathbf{A} = \mathbf{R}\mathbf{Q}$, kde \mathbf{R} je horní trojúhelníková a \mathbf{Q} je ortogonální. Jak byste spočítali RQ rozklad z QR rozkladu?

4.15. Matice \mathbf{A} se nazývá *normální*, pokud $\mathbf{A}^T \mathbf{A} = \mathbf{A}\mathbf{A}^T$. Příkladem je symetrická nebo antisymetrická matice. Dokažte, že pro normální matice platí $\text{rng } \mathbf{A} \perp \text{null } \mathbf{A}$.

4.16. Necht' $\mathbf{U}^T \mathbf{U} = \mathbf{I}$. Najděte co nejjednodušší vzorec pro

- a) vzdálenost bodu \mathbf{x} od podprostoru $\{\mathbf{x} \mid \mathbf{U}^T \mathbf{x} = \mathbf{0}\} = \text{null}(\mathbf{U}^T)$,
- b) vzdálenost počátku $\mathbf{0}$ od afinního podprostoru $\{\mathbf{x} \mid \mathbf{U}^T \mathbf{x} = \mathbf{b}\}$,
- c) vzdálenost bodu \mathbf{x} od afinního podprostoru $\{\mathbf{x} \mid \mathbf{U}^T \mathbf{x} = \mathbf{b}\}$.

4.17. Necht' $\mathbf{U}^T \mathbf{U} = \mathbf{I}$. Dokažte, že

- a) Prvky matice \mathbf{U} splňují $|u_{ij}| \leq 1$.
- b) Pro každý řádek \mathbf{u}^T matice \mathbf{U} platí $\|\mathbf{u}\| \leq 1$.

4.18. Necht' \mathbf{P} je ortogonální projektor. Dokažte, že

- a) Prvky matice \mathbf{P} splňují $|p_{ij}| \leq 1$.
- b) Diagonální prvky matice \mathbf{P} jsou nezáporné.

Nápověda a řešení

4.1. Např. $(1, 1, -1)$

4.2. Báze je $\{(1, 1, 1, -1)/2, (3, -1, -1, 1)/\sqrt{12}, (0, 1, 1, 2)/\sqrt{6}\}$

- 4.4. Musí být $\det \text{diag}(-\mathbf{1}_n) = (-1)^n > 0$, tedy pro sudá n .
- 4.5. $\binom{n}{2} = n(n-1)/2$
- 4.8. $\mathbf{x}_4 = (-1, 1, 0, 1)$.
- 4.9. Zvolíme $\mathbf{y} = (1, 2, 3) - r\mathbf{x}$, kde $r \in \mathbb{R}$ spočítáme z $\mathbf{x}^T \mathbf{y} = 0$. Tedy $r = \frac{5}{2}$ a $\mathbf{y} = (1, -\frac{1}{2}, \frac{1}{2})$. (V duchu Gram-Smidtovy ortogonalizace.)
- 4.10. Není matice náhodou ortogonální?
- 4.11.c) Plyne z (b).
- 4.11.d) Plyne z (c) s použitím $(X^\perp)^\perp = X$.
- 4.13.c) $\mathbf{H}\mathbf{x} = -\mathbf{x}$. Plyne z toho, že $\mathbf{V}\mathbf{V}^T \mathbf{x} = \mathbf{x}$, neboť $\mathbf{V}\mathbf{V}^T$ je projektor na X .
- 4.13.d) $\mathbf{H}\mathbf{x} = \mathbf{x}$, neboť $\mathbf{x} \in X = (\text{rng } \mathbf{V})^\perp = \text{null}(\mathbf{V}^T)$ a tedy $\mathbf{V}^T \mathbf{x} = \mathbf{0}$.
- 4.15. Použijte (5.4) a (4.8): $\text{rng } \mathbf{A} = \text{rng}(\mathbf{A}\mathbf{A}^T) = \text{rng}(\mathbf{A}^T \mathbf{A}) = \text{rng}(\mathbf{A}^T) = (\text{null } \mathbf{A})^\perp$.
- 4.16.a) $\|\mathbf{U}^T \mathbf{x}\|$
- 4.16.b) $\|\mathbf{b}\|$
- 4.16.c) $\|\mathbf{U}^T \mathbf{x} - \mathbf{b}\|$, což plyne buď z 4.16.a posunutím afinního podprostoru do počátku, nebo z 4.16.b posunutím bodu \mathbf{x} do počátku.
- 4.17.b) Doplňme matici \mathbf{U} na ortogonální matici $\mathbf{W} = [\mathbf{U} \ \mathbf{V}]$. Řádek této matice je $\mathbf{w}^T = [\mathbf{u}^T \ \mathbf{v}^T]$. Z ortogonality \mathbf{W} je však $\mathbf{w}^T \mathbf{w} = \mathbf{u}^T \mathbf{u} + \mathbf{v}^T \mathbf{v} = 1$. Z toho $\mathbf{u}^T \mathbf{u} \leq 1$.
- 4.18.a) Z $\mathbf{P} = \mathbf{U}\mathbf{U}^T$ máme $p_{ij} = \mathbf{u}_i^T \mathbf{u}_j$, kde \mathbf{u}_i je i -tý řádek \mathbf{U} . Protože $\|\mathbf{u}_i\| \leq 1$, musí být $|\mathbf{u}_i^T \mathbf{u}_j| \leq 1$.
- 4.18.b) Máme $p_{ii} = \mathbf{u}_i^T \mathbf{u}_i \geq 0$

Kapitola 5

Nehomogenní lineární soustavy

Mějme soustavu m lineárních rovnic o n neznámých

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad (5.1)$$

kde $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$. Soustava má (aspoň jedno) řešení právě tehdy, když $\mathbf{b} \in \text{rng } \mathbf{A}$ (tedy \mathbf{b} je lineární kombinací sloupců \mathbf{A}), což lze psát také jako $\text{rank}[\mathbf{A} \ \mathbf{b}] = \text{rank } \mathbf{A}$ (Frobeniova věta). Pokud je množina řešení soustavy neprázdná, je to afinní podprostor \mathbb{R}^n (viz Příklad 3.8).

Soustava je **homogenní** pokud $\mathbf{b} = \mathbf{0}$ a **nehomogenní** pokud $\mathbf{b} \neq \mathbf{0}$. V této kapitole se zaměříme pouze na nehomogenní soustavy. Rozlišíme tři případy:

- Soustava nemá řešení. To nastane právě tehdy, když $\mathbf{b} \notin \text{rng } \mathbf{A}$. Taková soustava se nazývá **přeuročená**. V tom případě můžeme chtít řešit soustavu přibližně, což je tématem §5.1.
- Soustava má právě jedno řešení. To nastane právě tehdy, když $\mathbf{b} \in \text{rng } \mathbf{A}$ a matice \mathbf{A} má lineárně nezávislé sloupce (tedy její nulový prostor je triviální).
- Soustava má nekonečně mnoho řešení. To nastane právě tehdy, když $\mathbf{b} \in \text{rng } \mathbf{A}$ a matice \mathbf{A} má lineárně závislé sloupce. Taková soustava se nazývá **nedourčená**. V tom případě můžeme chtít z množiny řešení vybrat jediné, což je tématem §5.2.

5.1 Přibližné řešení ve smyslu nejmenších čtverců

Pokud soustava (5.1) nemá řešení, řešme ji přibližně (což můžeme značit $\mathbf{A}\mathbf{x} \approx \mathbf{b}$). Hledjeme takové \mathbf{x} , aby eukleidovská norma vektoru $\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}$ zbytků (neboli *reziduí*) byla co nejmenší. Úloha se nezmění (proč?), když místo eukleidovské normy budeme minimalizovat její čtverec $\|\mathbf{r}\|^2 = \mathbf{r}^T \mathbf{r} = r_1^2 + \dots + r_m^2$. Tedy řešíme úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2. \quad (5.2)$$

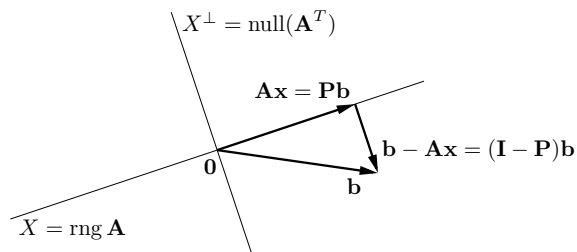
Protože minimalizujeme součet čtverců reziduí, mluvíme o přibližném řešení soustavy **ve smyslu nejmenších čtverců** (*least squares solution*).

Příklad 5.1. Soustava třech rovnic o dvou neznámých

$$\begin{aligned} x + 2y &= 6 \\ -x + y &= 3 \\ x + y &= 4 \end{aligned}$$

je přeuročená. Její přibližné řešení ve smyslu nejmenších čtverců znamená najít taková čísla x, y , která minimalizují číslo $(x + 2y - 6)^2 + (-x + y - 3)^2 + (x + y - 4)^2$. \square

Úlohu (5.2) vyřešíme následující úvahou. Pokud $\|\mathbf{Ax} - \mathbf{b}\|$ (tedy vzdálenost bodů \mathbf{Ax} a \mathbf{b}) má být minimální, musí být vektor $\mathbf{b} - \mathbf{Ax}$ kolmý na prostor $\text{rng } \mathbf{A}$, tedy na každý sloupec matice \mathbf{A} . Obrázek ukazuje situaci:



Tuto podmínku lze zapsat jako $\mathbf{A}^T(\mathbf{Ax} - \mathbf{b}) = \mathbf{0}$, tedy

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}. \quad (5.3)$$

Soustava (5.3) se proto nazývá **normální rovnice** (normála = kolmice). Je to soustava n rovnic o n neznámých.

Abychom mohli zkoumat řešitelnost této soustavy, uvedeme poměrně překvapivou větu.

Věta 5.1. Pro libovolnou matici \mathbf{A} platí

$$\text{null}(\mathbf{A}^T \mathbf{A}) = \text{null } \mathbf{A}, \quad (5.4a)$$

$$\text{rng}(\mathbf{A}^T \mathbf{A}) = \text{rng}(\mathbf{A}^T). \quad (5.4b)$$

Důkaz. Rovnost (5.4a) říká, že $(\mathbf{A}^T \mathbf{Ax} = \mathbf{0}) \Leftrightarrow (\mathbf{Ax} = \mathbf{0})$. Implikace \Leftarrow je jasná, vynásobením $\mathbf{Ax} = \mathbf{0}$ zleva maticí \mathbf{A}^T . Implikace \Rightarrow se dokáže takto:

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{0} \implies \mathbf{x}^T \mathbf{A}^T \mathbf{Ax} = (\mathbf{Ax})^T (\mathbf{Ax}) = 0 \implies \mathbf{Ax} = \mathbf{0},$$

neboť pro libovolný vektor \mathbf{y} máme $(\mathbf{y}^T \mathbf{y} = 0) \Rightarrow (\mathbf{y} = \mathbf{0})$ (napište si součin $\mathbf{y}^T \mathbf{y}$ ve složkách!).

Z (5.4a) a (3.17) je $\dim \text{rng}(\mathbf{A}^T \mathbf{A}) = \dim \text{rng}(\mathbf{A}^T)$. Z Věty 3.4 je $\text{rng}(\mathbf{A}^T \mathbf{A}) \subseteq \text{rng}(\mathbf{A}^T)$. Ale pokud je podprostor podmnožinou jiného podprostoru a oba mají stejnou dimenzi, musejí být stejné. To je jasné: libovolná báze podprostoru $\text{rng}(\mathbf{A}^T \mathbf{A})$ leží také v podprostoru $\text{rng}(\mathbf{A}^T)$, a protože oba podprostory mají stejnou dimenzi, je to také báze $\text{rng}(\mathbf{A}^T)$. \square

Speciálně, ze (5.4b) a (3.12) plyne

$$\text{rank}(\mathbf{A}^T \mathbf{A}) = \text{rank}(\mathbf{A}^T) = \text{rank } \mathbf{A} = \text{rank}(\mathbf{A} \mathbf{A}^T). \quad (5.5)$$

Nyní snadno vidíme, že soustava (5.3) má řešení pro libovolné \mathbf{A} a \mathbf{b} , neboť $\mathbf{A}^T \mathbf{b} \in \text{rng}(\mathbf{A}^T) = \text{rng}(\mathbf{A}^T \mathbf{A})$, kde $\text{rng}(\mathbf{A}^T \mathbf{A}) = \text{rng}(\mathbf{A}^T)$.

Dle (5.5) je matice $\mathbf{A}^T \mathbf{A}$ regulární právě tehdy, když \mathbf{A} má lineárně nezávislé sloupce. V tom případě můžeme soustavu (5.3) řešit pomocí inverze. Řešením je vektor $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$, kde

$$\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T. \quad (5.6)$$

Matice (5.6) se nazývá **pseudoinverze** matice \mathbf{A} s lineárně nezávislými sloupci. Je to jedna z levých inverzí matice \mathbf{A} , neboť $\mathbf{A}^+ \mathbf{A} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{A} = \mathbf{I}$.

Má-li matice \mathbf{A} lineárně závislé sloupce, vzorec (5.6) nelze použít. V tom případě soustava (5.3), a tedy i úloha (5.2), mají nekonečně mnoho (afinní podprostor) řešení (pozor, to je něco jiného, než že soustava (5.1) má nekonečně mnoho řešení!).

Pokud \mathbf{x} je řešení normální rovnice, vektor \mathbf{Ax} je ortogonální projekcí vektoru \mathbf{b} na podprostor $X = \text{rng } \mathbf{A}$ (viz obrázek výše). Pokud \mathbf{A} má lineárně nezávislé sloupce (tj. tyto sloupce tvoří bázi podprostoru X), z (5.6) máme $\mathbf{Ax} = \mathbf{Pb}$, kde

$$\mathbf{P} = \mathbf{A} \mathbf{A}^+ = \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T. \quad (5.7)$$

Toto je tedy projektor na podprostor $X = \text{rng } \mathbf{A}$ v případě, že \mathbf{A} má lineárně nezávislé (avšak ne nutně ortonormální) sloupce. Pokud \mathbf{A} má ortonormální sloupce, je $\mathbf{A}^T \mathbf{A} = \mathbf{I}$ a (5.7) se redukuje na (4.14). Projekce na X^\perp má přirozenou úlohu v problému (5.2): hodnota jeho minima je $\|\mathbf{b} - \mathbf{Ax}\|^2 = \|\mathbf{b} - \mathbf{Pb}\|^2 = \|(\mathbf{I} - \mathbf{P})\mathbf{b}\|^2$.

5.1.1 Řešení pomocí QR rozkladu

I když má matice \mathbf{A} lineárně nezávislé sloupce, řešení pomocí pseudoinverze (5.6) nemusí být vhodné pro numerické výpočty, kdy nezbytně používáme aritmetiku s konečnou přesností.

Příklad 5.2. Řešme soustavu $\mathbf{Ax} = \mathbf{b}$ pro

$$\mathbf{A} = \begin{bmatrix} 3 & 6 \\ 1 & 2.01 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 9 \\ 3.01 \end{bmatrix}.$$

Matice \mathbf{A} je regulární. Dejme tomu, že používáme aritmetiku s pohyblivou řádovou čárkou s přesností na 3 platné cifry. Gaussova eliminace najde přesné řešení soustavy $\mathbf{x} = (1, 1)$. Pokud ovšem v této aritmetice zformulujeme normální rovnici $\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}$, dostaneme

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix}, \quad \mathbf{A}^T \mathbf{b} = \begin{bmatrix} 30 \\ 60.1 \end{bmatrix}.$$

I když v přesné aritmetice je matice $\mathbf{A}^T \mathbf{A}$ regulární, v naší přibližné aritmetice došlo v součinu $\mathbf{A}^T \mathbf{A}$ došlo k zaokrouhlení a výsledná matice je singulární. Tedy soustava $\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}$ nemá řešení. \square

Numericky vhodnější způsob je řešit normální rovnici bez explicitního výpočtu součinu $\mathbf{A}^T \mathbf{A}$. To lze udělat pomocí redukovaného QR rozkladu $\mathbf{A} = \mathbf{QR}$. Po dosazení do normální rovnice máme $\mathbf{R}^T \mathbf{Q}^T \mathbf{Q} \mathbf{R} \mathbf{x} = \mathbf{R}^T \mathbf{Q}^T \mathbf{b}$ neboli $\mathbf{R}^T \mathbf{R} \mathbf{x} = \mathbf{R}^T \mathbf{Q}^T \mathbf{b}$. Jestliže \mathbf{A} má lineárně nezávislé sloupce, matice \mathbf{R} je regulární. Vynásobením maticí \mathbf{R}^{-T} zleva (což je tedy ekvivalentní úprava) máme

$$\mathbf{R} \mathbf{x} = \mathbf{Q}^T \mathbf{b}. \quad (5.8)$$

To je formálně stejný vzorec jako (4.22), rozdíl je ale v tom, že v (4.22) jsme použili *plný* QR rozklad, zatímco zde redukovaný. Tedy soustava (5.8) není ekvivalentní původní soustavě $\mathbf{Ax} = \mathbf{b}$ (protože \mathbf{Q} není čtvercová).

Jestliže sloupce \mathbf{A} jsou lineárně závislé, postup je trochu složitější, ale také stojí na QR rozkladu. V Matlabu je řešení nehomogenní lineární soustavy implementováno v operátoru `\` (*zpětné lomítko*). Pokud je soustava přeurčená, výsledkem je přibližné řešení ve smyslu nejmenších čtverců, přičemž použitý algoritmus používá QR rozklad. Pochopte všechny funkce operátorů *lomítka* a *zpětné lomítka* pomocí studia příkazů `help mrdivide` a `help mldivide`!

5.1.2 Lineární regrese

Regrese je modelování závislosti proměnné $y \in \mathbb{R}$ na proměnné $t \in T$ regresní funkcí

$$y = f(t, \mathbf{x}),$$

která je známa až na parametry $\mathbf{x} \in \mathbb{R}^n$. Je dán soubor dvojic (t_i, y_i) , $i = 1, \dots, m$, kde měření $y_i \in \mathbb{R}$ jsou zatížena chybou. Úkolem je najít parametry \mathbf{x} , aby $y_i \approx f(t_i, \mathbf{x})$ pro všechna i . Minimalizujeme součet čtverců reziduí, tedy řešíme úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^n} \sum_{i=1}^m (y_i - f(t_i, \mathbf{x}))^2. \quad (5.9)$$

Zvolme regresní funkci tak, aby pro každé t byla lineární funkcí parametrů \mathbf{x} . V to případě mluvíme o **lineární regresi**. Taková funkce je lineární kombinací

$$f(t, \mathbf{x}) = x_1 \varphi_1(t) + \dots + x_n \varphi_n(t) = \boldsymbol{\varphi}(t)^T \mathbf{x} \quad (5.10)$$

nějakých daných funkcí¹ $\varphi_j: T \rightarrow \mathbb{R}$. Pak

$$\sum_{i=1}^m (y_i - f(t_i, \mathbf{x}))^2 = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2,$$

kde $\mathbf{y} = (y_1, \dots, y_m)$ a prvky matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ jsou $a_{ij} = \varphi_j(t_i)$ (odvod'te!). Tedy vyjádřili jsme úlohu (5.9) ve tvaru (5.2).

Příklad 5.3. *Polynomiální regrese*². Necht' $T = \mathbb{R}$ a $\varphi_j(t) = t^{j-1}$. Pak regresní funkce je polynom stupně $n - 1$,

$$f(t, \mathbf{x}) = x_1 + x_2 t + x_3 t^2 + \dots + x_n t^{n-1}.$$

Matice

$$\mathbf{A} = \begin{bmatrix} 1 & t_1 & t_1^2 & \dots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \dots & t_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_m & t_m^2 & \dots & t_m^{n-1} \end{bmatrix}$$

je známá jako *Vandermondova matice*.

Speciálně pro $n = 1$ úloha (5.9) je $\min_x \sum_i (y_i - x)^2$. Řešením je aritmetický průměr $x = \frac{1}{m} \sum_{i=1}^m y_i$ (ověřte!). \square

5.2 Řešení s nejmenší normou

Předpokládejme nyní, že soustava (5.1) je nedourčená, neboli má nekonečně mnoho řešení. Je často užitečné z této množiny řešení vybrat jediné podle nějakého kritéria. Přírozeným kritériem je minimalizovat euklidovskou normu (tedy vzdálenost od počátku) řešení, což vede na úlohu

$$\min \{ \|\mathbf{x}\|^2 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b} \}. \quad (5.11)$$

Místo normy $\|\mathbf{x}\|$ opět minimalizujeme její čtverec. Tato úloha je známa jako řešení nehomogenní lineární soustavy s **nejmenší normou** (*least norm solution*). Podotkněme, že někdy je vhodné použít jiná kritéria než nejmenší eukleidovskou normu, viz např. Cvičení 9.25.

Příklad 5.4. Soustava dvou rovnic o třech neznámých

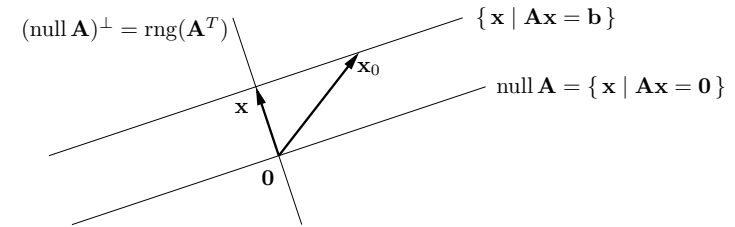
$$\begin{aligned} x + 2y + z &= 1 \\ -x + y + 2z &= 2 \end{aligned}$$

je nedourčená, tj. má nekonečně mnoho řešení. Její řešení s nejmenší normou je takové řešení, které minimalizuje číslo $x^2 + y^2 + z^2$. \square

Množinu řešení soustavy (5.1) lze psát (viz Příklad 3.8) jako

$$\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b} \} = \text{null } \mathbf{A} + \mathbf{x}_0, \quad (5.12)$$

kde \mathbf{x}_0 je libovolné (partikulární) řešení soustavy, tedy $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$. Množina (5.12) je afinní podprostor \mathbb{R}^m , je to lineární podprostor $\text{null } \mathbf{A}$ posunutý o vektor \mathbf{x}_0 . Viz obrázek:



Vektory \mathbf{x} a \mathbf{x}_0 jsou dvě různá řešení soustavy, ale pouze \mathbf{x} má nejmenší normu. Řešení \mathbf{x} má nejmenší normu právě tehdy, když $\mathbf{x} \perp \text{null } \mathbf{A}$, neboli $\mathbf{x} \in (\text{null } \mathbf{A})^\perp = \text{rng}(\mathbf{A}^T)$, kde poslední rovnost je (4.8b). Neboli musí existovat vektor $\mathbf{y} \in \mathbb{R}^m$ tak, že $\mathbf{x} = \mathbf{A}^T \mathbf{y}$. Pro vyřešení úlohy (5.11) tedy musíme vyřešit soustavu rovnic

$$\mathbf{A}^T \mathbf{y} = \mathbf{x}, \quad (5.13a)$$

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \quad (5.13b)$$

To je soustava $m + n$ rovnic o $m + n$ neznámých (\mathbf{x}, \mathbf{y}) .

Vyřešíme tuto soustavu. Dosadíme \mathbf{x} do druhé rovnice, $\mathbf{A}\mathbf{A}^T \mathbf{y} = \mathbf{b}$. Předpokládejme, že matice $\mathbf{A}\mathbf{A}^T$ má plnou hodnotu, což dle (5.5) nastane právě když \mathbf{A} má lineárně nezávislé řádky. Potom $\mathbf{y} = (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{b}$. Dosazením do první rovnice dostaneme $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$, kde

$$\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \quad (5.14)$$

se nazývá **pseudoinverze** matice \mathbf{A} s lineárně nezávislými řádky. Je to jedna z pravých inverzí matice \mathbf{A} (ověřte!).

Je poučné odvodit tento výsledek i trochu jinou úvahou. Z obrázku je patrné, že řešení \mathbf{x} má nejmenší normu právě tehdy, když je ortogonální projekcí vektoru \mathbf{x}_0 na podprostor $(\text{null } \mathbf{A})^\perp = \text{rng}(\mathbf{A}^T)$. Ortogonální projektor na podprostor reprezentovaný svou bází je dán vztahem (5.7), zde ovšem promítáme na $\text{rng}(\mathbf{A}^T)$ a tedy musíme vzorec použít s \mathbf{A}^T místo s \mathbf{A} . Tedy

$$\mathbf{x} = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{A}\mathbf{x}_0 = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{b} = \mathbf{A}^+ \mathbf{b}. \quad (5.15)$$

¹ Funkce φ_j se často nazývají *bázové funkce* (pokud jsou ovšem lineárně závislé).

² Nedejte se zmást tím, že polynom není lineární funkce a přesto jde o lineární regresi. Důležité je, že regresní funkce (5.10) je lineární v parametrech \mathbf{x} .

5.3 Cvičení

5.1. Máme soustavu $\mathbf{Ax} = \mathbf{b}$, kde $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{b} \neq \mathbf{0}$. Jsou tyto výroky pravdivé? Odpovědi dokažte.

- Pokud $m < n$, pak soustava má vždy řešení.
- Pokud $m > n$, pak soustava nemá nikdy řešení.
- Pokud $m < n$ a \mathbf{A} má plnou hodnotu, pak soustava má vždy nekonečně mnoho řešení.

5.2. Vyřešte (možno použít počítač) soustavu

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 1 & 1 & -3 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

přibližně ve smyslu nejmenších čtverců pomocí (a) pseudoinverze, (b) QR rozkladu.

5.3. Formulujte jako přibližné řešení soustavy $\mathbf{Pu} = \mathbf{q}$ ve smyslu nejmenších čtverců, tedy jako úlohu $\min_{\mathbf{u}} \|\mathbf{Pu} - \mathbf{q}\|^2$. Jako výsledek napište matice \mathbf{P} , \mathbf{q} , \mathbf{u} . Pokud existuje jednoduchý vzorec pro řešení (jak pro optimální hodnotu tak optimální argument), napište je.

- Hledá se bod $\mathbf{x} \in \mathbb{R}^m$, který minimalizuje součet čtverců vzdáleností k daným bodům $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$, tj. výraz $\sum_{i=1}^n \|\mathbf{a}_i - \mathbf{x}\|^2$.
- Hledá se vzdálenost bodu $\mathbf{y} \in \mathbb{R}^n$ od přímky $\{\mathbf{a} + t\mathbf{s} \mid t \in \mathbb{R}\}$ kde $\mathbf{a}, \mathbf{s} \in \mathbb{R}^n$.
- Hledá se příčka (nejkratší spojnice) dvou mimoběžných přímek $\{\mathbf{a}_1 + t\mathbf{s}_1 \mid t \in \mathbb{R}\}$ a $\{\mathbf{a}_2 + t\mathbf{s}_2 \mid t \in \mathbb{R}\}$ v \mathbb{R}^n , kde $\mathbf{a}_1, \mathbf{a}_2, \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^n$.
- Máme množinu m přímek v \mathbb{R}^n , kde i -tá přímka je množina $\{\mathbf{a}_i + t\mathbf{s}_i \mid t \in \mathbb{R}\}$ pro dané $\mathbf{a}_i, \mathbf{s}_i \in \mathbb{R}^n$. Hledá se bod $\mathbf{y} \in \mathbb{R}^n$, jehož součet čtverců vzdáleností k přímkám je minimální.
- Máme m nadrovin v prostoru \mathbb{R}^n , kde i -tá nad rovina má rovnici $\mathbf{a}_i^T \mathbf{x} = b_i$ pro dané $\mathbf{a}_i \in \mathbb{R}^n$ a $b_i \in \mathbb{R}$. Hledá se bod $\mathbf{y} \in \mathbb{R}^n$, který minimalizuje součet čtverců vzdáleností od jednotlivých nadrovin.
- V prknu je n děr o souřadnicích $x_1, \dots, x_n \in \mathbb{R}$, všechny v jedné přímce. Naměříme metrem vzdálenosti $d_{ij} = x_j - x_i$ pro vybrané dvojice $(i, j) \in E$, kde množina $E \subseteq \{1, \dots, n\} \times \{1, \dots, n\}$ je dána. Přitom dvojice jsou vybrané tak, že vždy $x_j > x_i$. Ze vzdáleností d_{ij} chceme spočítat souřadnice x_1, \dots, x_n . Odpovězte dále na otázky:
 - Kolik řešení má soustava $\mathbf{Pu} = \mathbf{q}$? Dokažte algebraicky a objasněte fyzikální význam.
 - Jsou sloupce \mathbf{P} lineárně nezávislé?

Diskutujte obě otázky pro případ, že měření jsou přesná, a pro případ, že měření jsou zatížená nepřesnostmi.

5.4. V problému *vážených nejmenších čtverců* chceme najít $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ minimalizující funkci

$$f(\mathbf{x}) = \sum_{i=1}^m w_i \left(\sum_{j=1}^n a_{ij} x_j - b_i \right)^2$$

kde w_i jsou nezáporné váhy. Napište funkci v maticovém tvaru, k čemuž zaved'te diagonální matici $\mathbf{W} = \text{diag}(w_1, \dots, w_m)$. Napište normální rovnici a pseudoinverzi pro tento případ.

5.5. Máme vektory $\mathbf{u} = (2, 1, -3)$ a $\mathbf{v} = (1, -1, 1)$. Najdi ortogonální projekci vektoru $(2, 0, 1)$ na podprostor (a) $\text{span}\{\mathbf{u}\}$, (b) $(\text{span}\{\mathbf{u}\})^\perp$, (c) $\text{span}\{\mathbf{u}, \mathbf{v}\}$, (d) $(\text{span}\{\mathbf{u}, \mathbf{v}\})^\perp$.

5.6. Necht' $X = \text{span}\{(-\frac{3}{5}, 0, \frac{4}{5}, 0), (0, 0, 0, 1), (\frac{4}{5}, 0, \frac{3}{5}, 0)\}$. Najdi projektory na podprostor X a podprostor X^\perp .

5.7. Máme $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 1 & 2 & 0 \end{bmatrix}$. Najdi ortogonální projekci vektoru $(1, 1, 1)$ na podprostory (a) $\text{rng } \mathbf{A}$, (b) $\text{null } \mathbf{A}$, (c) $\text{rng}(\mathbf{A}^T)$, (d) $\text{null}(\mathbf{A}^T)$.

5.8. Nulový prostor projektoru je typicky netriviální, tedy projektor \mathbf{P} je singularní matice. Kdy je \mathbf{P} regulární? Jaká je v tom případě matice \mathbf{A} ve vzorci (5.7) a podprostor $X = \text{rng } \mathbf{A}$? Jaký je geometrický význam této situace?

5.9. Dokažte následující vlastnosti pseudoinverze ze vztahů (5.6) a (5.14) pro libovolné (úzké, široké nebo čtvercové) matice plné hodnoty:

- $\mathbf{A}^+ = \mathbf{A}^{-1}$ když \mathbf{A} je čtvercová
- $(\mathbf{A}^+)^+ = \mathbf{A}$
- $(\mathbf{A}^T)^+ = (\mathbf{A}^+)^T$
- $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$, $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$
- $(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+$, $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}$
- $\mathbf{A}^T = \mathbf{A}^T\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+\mathbf{A}\mathbf{A}^T$
- $(\mathbf{A}^T\mathbf{A})^+ = \mathbf{A}^+(\mathbf{A}^T)^+$, $(\mathbf{A}\mathbf{A}^T)^+ = (\mathbf{A}^T)^+\mathbf{A}^+$

5.10. Spočítejte pseudoinverzi nenulového skaláru (matice s jedním řádkem a jedním sloupcem), nenulového sloupcového vektoru (matice s jedním sloupcem) a nenulového řádkového vektoru (matice s jedním řádkem).

5.11. Uvažujme projektor (5.7). Báze podprostoru X , na který promítáme, je tvořena sloupci matice \mathbf{A} . Projektor \mathbf{P} se nesmí změnit, vezmeme-li jinou bázi podprostoru. Různé báze podprostoru jsou dány sloupci matice $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{C}$ pro různé regulární matice $\mathbf{C} \in \mathbb{R}^{n \times n}$ (tedy \mathbf{C} je matice přechodu k jiné bázi).

5.12. Jaké bude řešení normálních rovnic (5.3) v případě, že $\mathbf{b} \perp \text{rng } \mathbf{A}$? Vyřešte geometrickou úvahou a pak zkuste dokázat algebraicky.

5.13. Hledáme vektor \mathbf{x} , který minimalizuje $\|\mathbf{Ax} - \mathbf{b}\|^2 + \mu\|\mathbf{x}\|^2$. Převed'te na tvar (5.2).

5.14. Najděte co nejjednodušší vzorec pro

- vzdálenost počátku $\mathbf{0}$ od nadroviny $\mathbf{a}^T \mathbf{x} = b$,
- vzdálenost počátku $\mathbf{0}$ od afinního podprostoru $\{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}\}$ (kde \mathbf{A} má lineárně nezávislé řádky)
- vzdálenost bodu \mathbf{x} od nadroviny $\mathbf{a}^T \mathbf{x} = b$

5.15. Dokažte, že pokud $\mathbf{A}^T \mathbf{A} = \mathbf{B}^T \mathbf{B}$, pak existuje \mathbf{C} tak, že $\mathbf{A} = \mathbf{C}\mathbf{B}$.

5.16. Neformálně je jasné, že nedourčená soustava bude mít vždy jen jedno řešení s nejmenší normou. Dokažte toto formálně, tj. dokažte, že jestliže soustava $\mathbf{Ax} = \mathbf{b}$ má alespoň jedno řešení pak soustava (5.13) má právě jedno řešení.

Nápověda a řešení

- 5.1.a) Neplatí. Příklad: $m = 1, n = 2, \mathbf{A} = \begin{bmatrix} 0 & 0 \end{bmatrix}, \mathbf{b} = 1$.
- 5.1.b) Neplatí.
- 5.1.c) Platí.
- 5.2. $(x_1, x_2, x_3) = (2, 1, 0)/3$
- 5.3.a) Minimum se nabývá v těžišti $\mathbf{x} = \frac{1}{n} \sum_{i=1}^n \mathbf{a}_i$.
- 5.3.c) Řešíme přeúčenou soustavu $\mathbf{a}_1 + t_1 \mathbf{s}_1 = \mathbf{a}_2 + t_2 \mathbf{s}_2$ o n rovnicích a 2 neznámých, tedy řešíme úlohu $\min_{t_1, t_2} \|(\mathbf{a}_1 + t_1 \mathbf{s}_1) - (\mathbf{a}_2 + t_2 \mathbf{s}_2)\|^2$. Tedy $\mathbf{P} = \begin{bmatrix} \mathbf{s}_1 & -\mathbf{s}_2 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}, \mathbf{q} = \mathbf{a}_2 - \mathbf{a}_1$.
- 5.3.d) Minimalizujte přes proměnné $\mathbf{y}, t_1, \dots, t_m$.
- 5.3.e) Nejprve si vzpomeňte či odvoďte, jak se spočítá vzdálenost bodu \mathbf{y} od nadroviny $\mathbf{a}^T \mathbf{x} = b$.
- 5.4. $f(\mathbf{x}) = (\mathbf{A}\mathbf{x} - \mathbf{b})^T \mathbf{W}(\mathbf{A}\mathbf{x} - \mathbf{b})$.
- 5.5. (a) $(2, 1, -3)/14$, (b) $(26, -1, 17)/14$, (c) $(62, -35, 17)/38$, (d) $(14, 35, 21)/38$
- 5.6. Nejsou náhodou vektory ortonormální?
- 5.6. Projektor na X je $\mathbf{P} = \text{diag}(1, 0, 1, 1)$. Projektor na X^\perp je $\mathbf{P} = \text{diag}(0, 1, 0, 0)$.
- 5.7. (a) $(1, 1, 1)$, (b) $(0.4, -0.2, 0)$, (c) $(0.6, 1.2, 1)$, (d) $(0, 0, 0)$. Pozor, \mathbf{A} nemá plnou hodnot.
- 5.8. \mathbf{A} je regulární, tedy $X = \mathbb{R}^m$. Projektor je identita.
- 5.11. $\tilde{\mathbf{A}}(\tilde{\mathbf{A}}^T \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^T = \mathbf{A}\mathbf{C}(\mathbf{C}^T \mathbf{A}^T \mathbf{A}\mathbf{C})^{-1} \mathbf{C}^T \mathbf{A}^T = \mathbf{A}\mathbf{C}\mathbf{C}^{-1}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{C}^{-T} \mathbf{C}^T \mathbf{A}^T = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$
- 5.12. $\mathbf{x} = \mathbf{0}$
- 5.14.a) $|b|/\|\mathbf{a}\|$
- 5.14.b) Čtverec vzdálenosti je roven optimální hodnotě úlohy (5.11), tedy $(\mathbf{A}^+ \mathbf{b})^T (\mathbf{A}^+ \mathbf{b}) = \mathbf{b}^T (\mathbf{A}\mathbf{A}^T)^{-T} \mathbf{A}\mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{b} = \mathbf{b}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{b}$.
- 5.14.c) $|\mathbf{a}^T \mathbf{x} - b|/\|\mathbf{a}\|$
- 5.15. Dle (5.4b) je $\text{rng}(\mathbf{A}^T) = \text{rng}(\mathbf{B}^T)$. Tedy každý řádek \mathbf{A} je lineární kombinací řádků \mathbf{B} . To jde napsat jako $\mathbf{A} = \mathbf{C}\mathbf{B}$ pro nějaké \mathbf{C} (viz Cvičení 3.12).

Kapitola 6

Spektrální rozklad a kvadratické funkce

Ze základní školy znáte polynomy jedné proměnné, co jsou ale polynomy více proměnných? **Monom** k -tého stupně n proměnných je výraz

$$x_1^{k_1} \dots x_n^{k_n},$$

kde $a_{k_1 \dots k_n} \in \mathbb{R}, k_1, \dots, k_n \in \{0, 1, \dots, k\}, k_1 + \dots + k_n = k$. **Polynom** n proměnných je lineární kombinace monomů, přičemž **stupeň polynomu** je stupeň jeho monomu nejvyššího stupně. Např. funkce

$$f(x, y) = x^2 y + xy - 2x + 1 \quad (6.1)$$

je polynom dvou proměnných třetího stupně, kde např. $x^2 y$ je monom třetího stupně a xy je monom druhého stupně. Polynom je *homogenní*, pokud stupně všech jeho monomů jsou stejné. Polynom (6.1) není homogenní, ale např. $f(x, y) = x^2 y - 5y^3$ je homogenní stupně tři.

Vidíme, že afinní funkce (3.21) je jen jiný název pro polynom prvního stupně a lineární funkce (3.6) (také zvaná lineární forma) je jiný název pro homogenní polynom prvního stupně¹. Polynom druhého stupně se nazývá *kvadratická funkce* a homogenní polynom druhého stupně *kvadratická forma*². Cílem této kapitoly je porozumět extrémům kvadratických forem a funkcí.

6.1 Vlastní čísla a vektory

Nechť pro čtvercovou matici $\mathbf{A} \in \mathbb{R}^{n \times n}$, nenulový vektor $\mathbf{v} \in \mathbb{C}^n$ a skalár $\lambda \in \mathbb{C}$ platí

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}. \quad (6.2)$$

Pak λ se nazývá **vlastní číslo** matice a \mathbf{v} **vlastní vektor** matice příslušný vlastnímu číslu λ . Vlastní čísla a vektory mohou být obecně komplexní.

Rovnici (6.2) lze přepsat jako

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}. \quad (6.3)$$

To je soustava homogenních lineárních rovnic pro \mathbf{v} , která má netriviální řešení právě tehdy, když matice $\mathbf{A} - \lambda\mathbf{I}$ je singulární. Tedy vlastní čísla jsou kořeny polynomu

$$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}), \quad (6.4)$$

¹Tohle samozřejmě platí jen pro funkce na vektorovém prostoru \mathbb{R}^n . Pokud uvažujeme funkce na abstraktním (tedy definovaným axiomou) vektorovém prostoru, který znáte z lineární algebry, pak např. lineární forma není polynom jednoduše proto, že 'polynom' na tomto prostoru nelze snadno definovat.

²Názvosloví není zcela konzistentní, což je opět dáno tím, že některá jména pocházejí z lineární algebry a některá z matematické analýzy.

který se nazývá **charakteristický polynom** matice \mathbf{A} .

Vlastní vektory příslušné vlastnímu číslu λ pak spočítáme ze soustavy (6.3). Vlastní vektor není svým vlastním číslem určen jednoznačně, vlastní vektory příslušné vlastnímu číslu λ tvoří celý podprostor $\text{null}(\mathbf{A} - \lambda\mathbf{I})$ (kromě počátku $\mathbf{0}$).

Příklad 6.1. Najděte vlastní čísla matice $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Charakteristická rovnice je

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det \begin{bmatrix} 1 - \lambda & 2 \\ 3 & 4 - \lambda \end{bmatrix} = (1 - \lambda)(4 - \lambda) - 3 \cdot 2 = \lambda^2 - 5\lambda - 2 = 0.$$

Tato kvadratická rovnice má dva kořeny $\lambda = (5 \pm \sqrt{33})/2$. To jsou vlastní čísla matice \mathbf{A} . Vlastní vektory příslušné každému λ najdeme řešením homogenní lineární soustavy

$$\begin{bmatrix} 1 - \lambda & 2 \\ 3 & 4 - \lambda \end{bmatrix} \mathbf{v} = \mathbf{0}. \quad \square$$

Z definice determinantu (2.4) plyne (promyslete!), že charakteristický polynom má stupeň n . Podle *základní věty algebry* má tedy právě n komplexních kořenů, počítáme-li násobnost. Označíme-li kořeny $\lambda_1, \dots, \lambda_n$, platí

$$p_{\mathbf{A}}(\lambda) = \prod_{i=1}^n (\lambda_i - \lambda).$$

V tomto smyslu má matice právě n vlastních čísel, z nichž některá mohou být stejná. Tomuto seznamu vlastních čísel se někdy říká **spektrum** matice.

Nechť $\lambda_1, \dots, \lambda_n$ jsou vlastní čísla matice a $\mathbf{v}_1, \dots, \mathbf{v}_n$ jsou jim příslušné vlastní vektory. Rovnice (6.2) lze pro ně napsat jako jedinou maticovou rovnici (rozmyslete!)

$$\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Lambda}, \quad (6.5)$$

kde diagonální matice $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$ má na diagonále vlastní čísla a sloupce čtvercové matice $\mathbf{V} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n] \in \mathbb{R}^{n \times n}$ jsou vlastní vektory.

Vlastní vektory mohou být lineárně závislé. Tato otázka není jednoduchá a podrobně ji zde diskutovat nebudeme. Řekneme jen, že existuje dobrý důvod vlastní vektory vybrat tak, aby hodnost matice \mathbf{V} byla největší možná.

Jak se počítají vlastní čísla a vektory? Charakteristický polynom je hlavně teoretický nástroj a přímé hledání jeho kořenů není vhodné pro numerický výpočet³. Pro větší matice se používají numerické iterační algoritmy, přičemž pro matice různého typu jsou vhodné různé algoritmy. Matlabská funkce `[V,D]=eig(A)` spočítá matice \mathbf{V} a $\mathbf{\Lambda}$ splňující (6.5).

6.1.1 Spektrální rozklad

Pokud je \mathbf{V} regulární (tj. existuje n lineárně nezávislých vlastních vektorů), je invertovatelná a (6.5) lze psát jako

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}. \quad (6.6)$$

Vztahu (6.6) se pak říká **rozklad matice podle vlastních čísel** nebo **spektrální rozklad**. V tom případě je matice \mathbf{A} podobná diagonální matici (neboli **diagonalizovatelná**), protože z (6.6) plyne $\mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \mathbf{\Lambda}$.

³Naopak, hledání kořenů libovolného polynomu lze převést na hledání vlastních čísel matice, která se nazývá *doprovodná matice (companion matrix)* polynomu.

Věta 6.1. Necht' $\mathbf{A} \in \mathbb{R}^{n \times n}$. Následující dvě tvrzení jsou ekvivalentní:

- Matice \mathbf{A} je symetrická.
- Všechna vlastní čísla matice \mathbf{A} jsou reálná a existuje ortonormální množina n jejích vlastních vektorů.

Této větě se někdy říká **spektrální věta**. Podle ní pro každou symetrickou \mathbf{A} je v (6.5) matice $\mathbf{\Lambda}$ reálná a \mathbf{V} může být zvolena jako ortogonální ($\mathbf{V}^{-1} = \mathbf{V}^T$). Tedy

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^T. \quad (6.7)$$

Zároveň jsme vpravo uvedli i druhou formu rozkladu jako součet dyád (viz §2.6), tzv. **dyadický rozvoj** (přesvědčte se, že druhá rovnost v (6.7) platí!).

Všimněme si, že dle Věty 3.4 (použité dvakrát na výraz $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$) máme

$$r = \text{rank } \mathbf{\Lambda} = \text{rank } \mathbf{A}, \quad (6.8)$$

kde r je počet nenulových vlastních čísel. Když $r < n$ (tedy některá vlastní čísla jsou nulová), můžeme vynechat sloupce+řádky matice $\mathbf{\Lambda}$ a sloupce matice \mathbf{V} odpovídající nulovým vlastním číslům, tedy v (6.7) bude $\mathbf{\Lambda} \in \mathbb{R}^{r \times r}$ diagonální regulární a $\mathbf{V} \in \mathbb{R}^{n \times r}$ bude mít ortonormální sloupce ($\mathbf{V}^T\mathbf{V} = \mathbf{I}$). V tom případě suma dyád bude mít jen r nenulových členů.

Příklad 6.2. Zde je spektrální rozklad matice 3×3 s jedním vlastním číslem nulovým:

$$\begin{aligned} \begin{bmatrix} 2 & 3 & -2 \\ 3 & 2 & 2 \\ -2 & 2 & -8 \end{bmatrix} &= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{18} & -2/3 \\ 1/\sqrt{2} & -1/\sqrt{18} & 2/3 \\ 0 & 4/\sqrt{18} & 1/3 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ -2/3 & 2/3 & 1/3 \end{bmatrix} \\ &= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{18} \\ 1/\sqrt{2} & -1/\sqrt{18} \\ 0 & 4/\sqrt{18} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & -9 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \end{bmatrix} \quad \square \end{aligned}$$

Vlastní čísla a vektory jsou rozsáhlé téma, které jsme zde zdaleka nevyčerпали. To ale není ani třeba, protože dále budeme potřebovat jen spektrální rozklad symetrické matice.

6.2 Kvadratická forma

Kvadratická forma je homogenní polynom $f: \mathbb{R}^n \rightarrow \mathbb{R}$ druhého stupně. Je pohodlné ji zapsat v maticovém tvaru

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} x_i x_j \quad (6.9)$$

pro nějakou matici $\mathbf{A} \in \mathbb{R}^{n \times n}$. Protože $x_i x_j = x_j x_i$ (násobení reálných čísel je komutativní), máme

$$\sum_{j=1}^n \sum_{i=1}^n a_{ij} x_i x_j = \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n (a_{ij} + a_{ji}) x_i x_j.$$

Vidíme, že funkce f závisí jen na součtech $a_{ij} + a_{ji}$. Je proto zvykem předpokládat $a_{ij} = a_{ji}$, neboli že matice \mathbf{A} je symetrická ($\mathbf{A}^T = \mathbf{A}$).

Příklad 6.3. Příkladem kvadratické formy dvou proměnných je funkce

$$f(x, y) = 2x^2 - 2xy + y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Všimněte si, že první matice není symetrická a druhá ano. \square

Jako užitečné cvičení toto odvodíme i v maticové formě. Každou čtvercovou matici můžeme psát jako součet symetrické a antisymetrické matice:

$$\mathbf{A} = \underbrace{\frac{1}{2}(\mathbf{A} + \mathbf{A}^T)}_{\text{symetrická}} + \underbrace{\frac{1}{2}(\mathbf{A} - \mathbf{A}^T)}_{\text{antisymetrická}}$$

(viz Cvičení 2.11). Ale

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \frac{1}{2} \mathbf{x}^T (\mathbf{A} + \mathbf{A}^T) \mathbf{x} + \frac{1}{2} \underbrace{\mathbf{x}^T (\mathbf{A} - \mathbf{A}^T) \mathbf{x}}_0,$$

nebot' $\mathbf{x}^T (\mathbf{A} - \mathbf{A}^T) \mathbf{x} = \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{A}^T \mathbf{x} = \mathbf{x}^T \mathbf{A} \mathbf{x} - (\mathbf{x}^T \mathbf{A} \mathbf{x})^T = 0$, kde jsme použili skutečnost, že transpozice skaláru je tentýž skalár. Tedy když \mathbf{A} není symetrická, můžeme ji nahradit její symetrickou částí a kvadratická forma se nezmění.

Definice 6.1. Symetrická matice \mathbf{A} je

- **pozitivně [negativně] semidefinitní**, když pro každé \mathbf{x} platí $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$ [$\mathbf{x}^T \mathbf{A} \mathbf{x} \leq 0$]
- **pozitivně [negativně] definitní**, když pro každé $\mathbf{x} \neq \mathbf{0}$ platí $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ [$\mathbf{x}^T \mathbf{A} \mathbf{x} < 0$]
- **indefinitní**, když existuje \mathbf{x} a \mathbf{y} tak, že $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ a $\mathbf{y}^T \mathbf{A} \mathbf{y} < 0$.

Matice může mít i několik těchto vlastností najednou. Např. pozitivně definitní matice je zároveň pozitivně semidefinitní. Nulová matice je zároveň pozitivně i negativně semidefinitní.

I když definice dává smysl pro libovolné čtvercové matice, je zvykem hovořit o těchto vlastnostech jen pro symetrické matice. Někdy se tyto vlastnosti definují ne pro matici, ale abstraktněji pro kvadratickou formu.

Z Definice 6.1 je jasné, má-li kvadratická forma extrém a případně jaký:

Věta 6.2. Necht' funkce f je dána jako $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$.

- Je-li \mathbf{A} pozitivně [negativně] semidefinitní, pak f v bodě $\mathbf{0}$ nabývá minimum [maximum].
- Je-li \mathbf{A} pozitivně [negativně] definitní, pak f v bodě $\mathbf{0}$ nabývá ostré minimum [maximum].
- Je-li \mathbf{A} indefinitní, pak f nemá minimum ani maximum.

Důkaz. Je-li \mathbf{A} pozitivně semidefinitní, kvadratická forma nemůže být záporná a zároveň pro $\mathbf{x} = \mathbf{0}$ je nulová, proto v bodě $\mathbf{x} = \mathbf{0}$ (i když možná i jinde) nabývá svého minima. Je-li \mathbf{A} indefinitní a např. $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$, bod \mathbf{x} nemůže být maximum, protože $(2\mathbf{x})^T \mathbf{A} (2\mathbf{x}) > \mathbf{x}^T \mathbf{A} \mathbf{x}$, a nemůže to být minimum, protože pro nějaké \mathbf{y} je $\mathbf{y}^T \mathbf{A} \mathbf{y} < 0$. \square

Věta 6.3. Symetrická matice je

- **pozitivně [negativně] semidefinitní**, právě když má všechna vlastní čísla nezáporná [nekladná]
- **pozitivně [negativně] definitní**, právě když má všechna vlastní čísla kladná [záporná]

- **indefinitní**, právě když má alespoň jedno kladné a alespoň jedno záporné vlastní číslo.

Důkaz. Z rozkladu podle vlastních čísel (6.7) máme

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \mathbf{x} = \mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2. \quad (6.10)$$

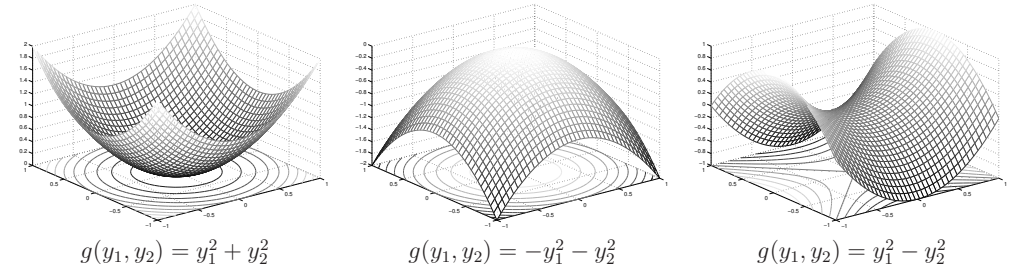
Substituce $\mathbf{x} = \mathbf{V} \mathbf{y}$ tedy diagonalizovala matici kvadratické formy. Protože \mathbf{V} je regulární, definitnost matice \mathbf{A} je stejná jako definitnost matice $\mathbf{\Lambda}$. Např.

$$(\forall \mathbf{x})(\mathbf{x}^T \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \mathbf{x} \geq 0) \iff (\forall \mathbf{y})(\mathbf{y}^T \mathbf{\Lambda} \mathbf{y} \geq 0).$$

Ale protože $\mathbf{\Lambda}$ je diagonální, její definitnost je okamžitě patrná ze znamének čísel λ_i . Např. výraz (6.10) je nezáporný pro každé \mathbf{y} právě tehdy, když všechna λ_i jsou nezáporná. \square

Označíme-li $g(\mathbf{y}) = \mathbf{y}^T \mathbf{\Lambda} \mathbf{y}$, máme $f(\mathbf{x}) = g(\mathbf{V}^T \mathbf{x})$. Protože matice \mathbf{V} je ortogonální, transformace $\mathbf{y} = \mathbf{V}^T \mathbf{x}$ je isometrie, tedy funkce f a g se liší jen otočením (příp. zrcadlením) v prostoru vzorů.

Pro případ dvou proměnných ($n = 2$) si grafy diagonální formy g snadno představíme. Jsou-li obě vlastní čísla kladná, funkce g vypadá jako 'dolík'. Jsou-li obě vlastní čísla záporná, funkce g vypadá jako 'kopec'. Mají-li vlastní čísla opačná znaménka, tvarem je 'sedlo':



6.3 Kvadratická funkce

Kvadratická funkce je polynom (ne nutně homogenní) druhého stupně. Lze jej psát v maticovém tvaru

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c, \quad (6.11)$$

kde $\mathbf{A}^T = \mathbf{A}$. Oproti⁴ kvadratické formě tedy přibýly lineární a konstantní členy. Všimněte si, že pro $n = 1$ je (6.11) známá kvadratická funkce jedné proměnné $f(x) = ax^2 + bx + c$.

Jak nalézt extrémy kvadratické funkce? Extrémy lze hledat mechanicky pomocí derivací, to však ukážeme až v pozdější kapitole. Jiný způsob je převést kvadratickou funkci na kvadratickou formu posunutím počátku. Tento způsob popíšeme nyní.

Někdy lze najít vektor $\mathbf{x}_0 \in \mathbb{R}^n$ a skalár y_0 takové, že

$$\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = (\mathbf{x} - \mathbf{x}_0)^T \mathbf{A} (\mathbf{x} - \mathbf{x}_0) + y_0. \quad (6.12)$$

⁴Pokud $\mathbf{A} = \mathbf{0}$, bude f pouhá afinní funkce. Je věcí konvence, zda afinní funkci máme nazývat kvadratickou či nikoliv, tedy zda máme zakázat případ $\mathbf{A} = \mathbf{0}$.

Výraz na pravé straně je kvadratická forma s počátkem posunutým do bodu \mathbf{x}_0 , plus konstanta. Této úpravě se říká **doplnění na čtverec**. Znáte ji pro případ $n = 1$, neboť tak se na základní škole odvozuje vzorec pro kořeny kvadratické rovnice jedné proměnné. Spočítáme \mathbf{x}_0, y_0 z daných $\mathbf{A}, \mathbf{b}, c$. Roznásobením pravé strany dostaneme

$$\begin{aligned} (\mathbf{x} - \mathbf{x}_0)^T \mathbf{A} (\mathbf{x} - \mathbf{x}_0) + y_0 &= \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{A} \mathbf{x}_0 - \mathbf{x}_0^T \mathbf{A} \mathbf{x} + \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + y_0 \\ &= \mathbf{x}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}_0^T \mathbf{A} \mathbf{x} + \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + y_0. \end{aligned}$$

Porovnáním členů stejného stupně máme

$$\mathbf{b} = -2\mathbf{A}\mathbf{x}_0, \quad (6.13a)$$

$$c = \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + y_0, \quad (6.13b)$$

z čehož spočítáme \mathbf{x}_0 a y_0 . Pokud soustava (6.13a) není řešitelná, doplnění na čtverec není možné.

Pokud je doplnění na čtverec možné, vyšetření extrémů kvadratické funkce se neliší od vyšetření extrémů kvadratické formy, protože rozdíl je jen v posunutí \mathbf{x}_0 . Pokud doplnění na čtverec možné není, kvadratická funkce extrém nemá.

Příklad 6.4. Máme kvadratickou funkci

$$f(x, y) = 2x^2 - 2xy + y^2 - 2y + 3 = \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ -2 \end{bmatrix}^T \begin{bmatrix} x \\ y \end{bmatrix} + 3.$$

Její doplnění na čtverec je

$$f(x, y) = 2(x-1)^2 - 2(x-1)(y-2) + (y-2)^2 + 1 = \begin{bmatrix} x-1 \\ y-2 \end{bmatrix}^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x-1 \\ y-2 \end{bmatrix} + 1,$$

tedy máme $\mathbf{x}_0 = (1, 2), y_0 = 1$. Jelikož matice \mathbf{A} je pozitivně definitní (ověřte!), má kvadratická funkce minimum v bodě \mathbf{x}_0 . \square

Příklad 6.5. Kvadratická funkce

$$f(x, y) = x^2 - y = \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix}^T \begin{bmatrix} x \\ y \end{bmatrix}$$

doplnit na čtverec nejde. Funkce tedy nemá na \mathbb{R}^2 extrém. \square

Příklad 6.6. Řešme znovu úlohu (5.2). Hledáme minimum kvadratické funkce

$$\begin{aligned} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 &= (\mathbf{A}\mathbf{x} - \mathbf{b})^T (\mathbf{A}\mathbf{x} - \mathbf{b}) \\ &= (\mathbf{x}^T \mathbf{A}^T - \mathbf{b}^T) (\mathbf{A}\mathbf{x} - \mathbf{b}) \\ &= \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{A}^T \mathbf{b} - \mathbf{b}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{b} \\ &= \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} - 2\mathbf{b}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{b}, \end{aligned} \quad (6.14)$$

kde jsme použili skutečnost, že skalár je roven své transpozici a tedy $\mathbf{b}^T \mathbf{A} \mathbf{x} = (\mathbf{b}^T \mathbf{A} \mathbf{x})^T = \mathbf{x}^T \mathbf{A}^T \mathbf{b}$. Extrém této kvadratické funkce můžeme najít doplněním na čtverec (viz §6.3). Soustava (6.13a) bude mít tvar $\mathbf{A}^T \mathbf{A} \mathbf{x}_0 = \mathbf{A}^T \mathbf{b}$ (pozor: \mathbf{A}, \mathbf{b} znamená něco jiného v (6.14) a v (6.13a)), tedy dostali jsme normální rovnici (5.3). Zároveň je jasné, že matice $\mathbf{A}^T \mathbf{A}$ je pozitivně semidefinitní, neboť pro každé $\mathbf{x} \in \mathbb{R}^n$ máme

$$\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} = (\mathbf{A}\mathbf{x})^T \mathbf{A} \mathbf{x} = \|\mathbf{A}\mathbf{x}\|^2 \geq 0. \quad (6.15)$$

Tedy v bodě \mathbf{x}_0 bude minimum. \square

6.3.1 Kvadrika

Vrstevnice kvadratické funkce se nazývá **kvadrika**. Tedy kvadrika je množina

$$\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = 0\}. \quad (6.16)$$

Když $\mathbf{b} = \mathbf{0}$, kvadrika má střed v počátku. Když \mathbf{A} je diagonální, kvadrika má osy rovnoběžné se souřadnicovými osami. Když $\mathbf{A} = \mathbf{0}$, množina (6.16) je pouhá nadrovina. Když \mathbf{A} nemá plnou hodnotu, kvadrika je degenerovaná. Množina (6.16) může být i prázdná.

Jestliže kvadratická funkce dovoluje doplnění na čtverec, můžeme množinu (6.16) psát jako

$$\{\mathbf{x} \in \mathbb{R}^n \mid (\mathbf{x} - \mathbf{x}_0)^T \mathbf{A} (\mathbf{x} - \mathbf{x}_0) + y_0 = 0\},$$

což je vlastně vrstevnice kvadratické formy posunutá o \mathbf{x}_0 . V tom případě je typ kvadriky určen jednoduše znaménky vlastních čísel matice \mathbf{A} . Speciálně, když všechna vlastní čísla jsou kladná (tedy \mathbf{A} je pozitivně definitní), jde o **elipsoid**⁵. Když některá vlastní čísla jsou nulová (tedy \mathbf{A} nemá plnou hodnotu), kvadrika je degenerovaná. Toto ale nevyčerpává všechny typy degenerace: další typy degenerace nastanou v případě, kdy \mathbf{A} nemá plnou hodnotu a funkce nedovoluje doplnění na čtverec.

Pro $n = 2$ se kvadrika nazývá **kuželosečka** (angl. *conic*).

6.4 Spektrální rozklad jako optimalizační úloha

Necht' $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ je spektrální rozklad symetrické matice $\mathbf{A} \in \mathbb{R}^{n \times n}$. Předpokládáme přitom, že $\lambda_1 \geq \dots \geq \lambda_n$, což lze vždy zařadit vhodnou permutací sloupců $\mathbf{v}_1, \dots, \mathbf{v}_n$ matice \mathbf{V} a diagonálních prvků $\lambda_1, \dots, \lambda_n$ matice $\mathbf{\Lambda}$.

Ukážeme, že

$$\min\{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\} = \lambda_n \quad (6.17)$$

a minimální hodnota se nabývá pro $\mathbf{x} = \mathbf{v}_n$, tedy pro vlastní vektor příslušný nejmenšímu vlastního číslu. To snadno dokážeme. Pohleďme na rovnost (6.10). Protože \mathbf{V} je ortogonální a $\mathbf{y} = \mathbf{V}^T \mathbf{x}$, je $\mathbf{y}^T \mathbf{y} = \mathbf{x}^T \mathbf{V} \mathbf{V}^T \mathbf{x} = \mathbf{x}^T \mathbf{x}$ a tedy podmínka $\mathbf{x}^T \mathbf{x} = 1$ je ekvivalentní podmínce $\mathbf{y}^T \mathbf{y} = 1$. Tedy úloha (6.17) má stejnou minimální hodnotu jako úloha

$$\min\{\mathbf{y}^T \mathbf{\Lambda} \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^n, \mathbf{y}^T \mathbf{y} = 1\} = \min\{\lambda_1 y_1^2 + \dots + \lambda_n y_n^2 \mid \mathbf{y} \in \mathbb{R}^n, y_1^2 + \dots + y_n^2 = 1\}.$$

Je jasné, že minimum se nabývá pro

$$y_i = \begin{cases} 0 & \text{když } i < n, \\ 1 & \text{když } i = n. \end{cases}$$

tedy pro $\mathbf{y} = \mathbf{e}_n$. To odpovídá $\mathbf{x} = \mathbf{V} \mathbf{e}_n = \mathbf{v}_n$ a $\mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = \lambda_n$.

Zobecníme nyní úlohu (6.17):

Věta 6.4. *Necht' $k \leq n$. Platí*

$$\min\{\text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) \mid \mathbf{X} \in \mathbb{R}^{n \times k}, \mathbf{X}^T \mathbf{X} = \mathbf{I}\} = \lambda_{n-k+1} + \dots + \lambda_n, \quad (6.18)$$

přičemž minimum se nabývá pro $\mathbf{X} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_k]$, tedy pro sloupce \mathbf{X} rovné vlastním vektorům příslušným k nejmenším vlastním číslům matice \mathbf{A} .

⁵Někdy se přesněji nazývá *elipsoidní povrch* a elipsoidem se myslí množina i s vnitřkem. Rozdíl je stejný jako mezi sférou a koulí.

Všimněte si, že $\text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) = \mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 + \dots + \mathbf{x}_k^T \mathbf{A} \mathbf{x}_k$, kde $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ jsou sloupce matice \mathbf{X} . Podmínka $\mathbf{X}^T \mathbf{X} = \mathbf{I}$ pak znamená, že množina vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je ortonormální.

Důkaz. Máme (viz §2.3)

$$\text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) = \text{tr}(\mathbf{X}^T \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \mathbf{X}) = \text{tr}(\mathbf{Y}^T \mathbf{\Lambda} \mathbf{Y}) = \text{tr}(\mathbf{Y} \mathbf{Y}^T \mathbf{\Lambda}) = \text{tr}(\mathbf{P} \mathbf{\Lambda}) = \lambda_1 p_{11} + \dots + \lambda_n p_{nn},$$

kde $\mathbf{Y} = \mathbf{V}^T \mathbf{X}$ a p_{11}, \dots, p_{nn} jsou diagonální prvky matice $\mathbf{P} = \mathbf{Y} \mathbf{Y}^T$. Protože \mathbf{V} je ortogonální, podmínka $\mathbf{X}^T \mathbf{X} = \mathbf{I}$ je ekvivalentní podmínce $\mathbf{Y}^T \mathbf{Y} = \mathbf{I}$.

Diagonální prvky matice \mathbf{P} splňují rovnost

$$p_{11} + \dots + p_{nn} = \text{tr} \mathbf{P} = \text{tr}(\mathbf{Y} \mathbf{Y}^T) = \text{tr}(\mathbf{Y}^T \mathbf{Y}) = \text{tr} \mathbf{I}_k = k.$$

Navíc splňují nerovnosti $0 \leq p_{ii} \leq 1$, viz Cvičení 4.18. Řešme nyní úlohu

$$\min\{\lambda_1 p_{11} + \dots + \lambda_n p_{nn} \mid 0 \leq p_{11}, \dots, p_{nn} \leq 1, p_{11} + \dots + p_{nn} = k\}.$$

Úvahou snadno vidíme (viz Cvičení 11.2), že optimální hodnota je $\lambda_{n-k+1} + \dots + \lambda_n$ (tedy součet k nejmenších vlastních hodnot matice \mathbf{A}) a optimální argument je

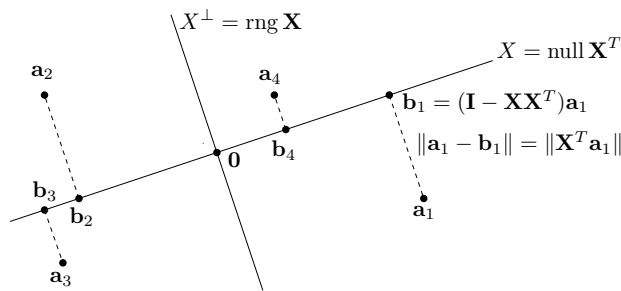
$$p_{ii} = \begin{cases} 0 & \text{když } i \leq k, \\ 1 & \text{když } i > n - k. \end{cases}$$

Ale matici $\mathbf{P} = \mathbf{Y} \mathbf{Y}^T$ s těmito diagonálními prvky lze realizovat volbou $\mathbf{Y} = \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_k \end{bmatrix}$. Z toho $\mathbf{X} = \mathbf{V} \mathbf{Y} = [\mathbf{v}_{n-k+1} \ \dots \ \mathbf{v}_n]$. \square

Věta netvrdí, že úloha (6.18) má jediný optimální argument. Opravdu, jestliže \mathbf{X} je optimální argument úlohy, pak také $\mathbf{X} \mathbf{U}$ je optimální argument pro každou ortogonální $\mathbf{U} \in \mathbb{R}^{k \times k}$, neboť $\text{tr}[(\mathbf{X} \mathbf{U})^T \mathbf{A} (\mathbf{X} \mathbf{U})] = \text{tr}(\mathbf{U}^T \mathbf{X}^T \mathbf{A} \mathbf{X} \mathbf{U}) = \text{tr}(\mathbf{U} \mathbf{U}^T \mathbf{X}^T \mathbf{A} \mathbf{X}) = \text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X})$.

6.4.1 Optimální proložení bodů podprostorem

Hledejme lineární podprostor $X \subseteq \mathbb{R}^n$ dimenze $n - k$, který minimalizuje součet čtverců vzdáleností k daným bodům⁶ $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$. Viz obrázek:



⁶ Ve statistice se této úloze říká *rozvoj podle hlavních komponent (principal component analysis, PCA)* nebo *Karhunen-Loewův rozvoj*.

Podotkneme, že tuto úlohu nelze nijak převést na úlohu nejmenších čtverců z (5.2).

Místo podprostoru X uvažujme jeho ortogonální doplněk X^\perp , který reprezentujeme bází tvořící sloupce matice $\mathbf{X} \in \mathbb{R}^{n \times k}$. Tedy $X^\perp = \text{rng} \mathbf{X}$ a $X = \text{null}(\mathbf{X}^T)$. Má-li matice \mathbf{X} ortonormální sloupce, vzdálenost bodu $\mathbf{a} \in \mathbb{R}^n$ k podprostoru X je $\|\mathbf{X}^T \mathbf{a}\|$ (viz §4.4). Součet čtverců vzdáleností všech bodů k podprostoru X je

$$\|\mathbf{X}^T \mathbf{a}_1\|^2 + \dots + \|\mathbf{X}^T \mathbf{a}_m\|^2 = \|\mathbf{A} \mathbf{X}\|^2,$$

kde $\mathbf{A} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_m]^T \in \mathbb{R}^{m \times n}$ je matice s řádky $\mathbf{a}_1^T, \dots, \mathbf{a}_m^T$. Musíme tedy vyřešit úlohu

$$\min\{\|\mathbf{A} \mathbf{X}\| \mid \mathbf{X} \in \mathbb{R}^{n \times k}, \mathbf{X}^T \mathbf{X} = \mathbf{I}\}. \quad (6.19)$$

Jelikož $\|\mathbf{A} \mathbf{X}\|^2 = \text{tr}(\mathbf{X}^T \mathbf{A}^T \mathbf{A} \mathbf{X})$, podle Věty 6.4 jsou sloupce optimální matice \mathbf{X} vlastní vektory příslušné k nejmenším vlastním číslům matice $\mathbf{A}^T \mathbf{A}$.

Příklad 6.7. Jsou dány body $\mathbf{a}_1, \dots, \mathbf{a}_m$ v prostoru \mathbb{R}^3 , jež tvoří řádky matice $\mathbf{A} \in \mathbb{R}^{m \times 3}$. Necht' $\mathbf{A}^T \mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ je spektrální rozklad matice \mathbf{A} , kde $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ jsou sloupce matice $\mathbf{V} \in \mathbb{R}^{3 \times 3}$ a $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$. Přímka procházející počátkem, která minimalizuje součet čtverců vzdáleností k bodům, je množina

$$\text{span}\{\mathbf{v}_1\} = (\text{span}\{\mathbf{v}_2, \mathbf{v}_3\})^\perp = \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{v}_2^T \mathbf{x} = \mathbf{v}_3^T \mathbf{x} = 0\}.$$

Rovina procházející počátkem, která minimalizuje součet čtverců vzdáleností k bodům, je množina

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2\} = (\text{span}\{\mathbf{v}_3\})^\perp = \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{v}_3^T \mathbf{x} = 0\}. \quad \square$$

Zobecníme nyní úlohu: místo lineárního podprostoru hledáme *afinní* podprostor dimenze $n - k$, který minimalizuje součet čtverců vzdáleností k bodům $\mathbf{a}_1, \dots, \mathbf{a}_m$. Hledaný afinní podprostor je množina $\{\mathbf{a} \in \mathbb{R}^n \mid \mathbf{X}^T \mathbf{a} = \mathbf{y}\}$ pro nějakou matici $\mathbf{X} \in \mathbb{R}^{n \times k}$ s ortonormálními sloupci a nějaký vektor $\mathbf{y} \in \mathbb{R}^m$. Vzdálenost bodu $\mathbf{a} \in \mathbb{R}^n$ od tohoto podprostoru je $\|\mathbf{X}^T \mathbf{a} - \mathbf{y}\|$ (viz Cvičení 4.16). Součet čtverců vzdáleností všech bodů k podprostoru je tedy

$$\|\mathbf{X}^T \mathbf{a}_1 - \mathbf{y}\|^2 + \dots + \|\mathbf{X}^T \mathbf{a}_m - \mathbf{y}\|^2 = \|\mathbf{A} \mathbf{X} - \mathbf{1} \mathbf{y}^T\|^2.$$

Abychom našli optimální afinní podprostor, musíme tento výraz minimalizovat přes proměnné \mathbf{X} a \mathbf{y} za podmínky $\mathbf{X}^T \mathbf{X} = \mathbf{I}$. Pokud je \mathbf{X} pevné a minimalizujeme pouze přes \mathbf{y} , optimum se nabývá v bodě $\mathbf{y} = \mathbf{X}^T \bar{\mathbf{a}}$, kde

$$\bar{\mathbf{a}} = \frac{1}{m}(\mathbf{a}_1 + \dots + \mathbf{a}_m)$$

je *těžiště* daných bodů (viz Cvičení 5.3.a). Tedy afinní podprostor je $\{\mathbf{a} \in \mathbb{R}^n \mid \mathbf{X}^T \mathbf{a} = \mathbf{X}^T \bar{\mathbf{a}}\}$. Vidíme, že tento podprostor prochází těžištěm $\bar{\mathbf{a}}$.

Nyní je řešení úlohy jasné: stačí nejprve posunout body $\mathbf{a}_1, \dots, \mathbf{a}_m$ tak, aby jejich těžiště leželo v počátku, a potom najít lineární podprostor, který minimalizuje součet čtverců vzdáleností k posunutým bodům.

6.4.2 Nejbližší matice nižší hodnosti

Ukážeme, že úloha (6.19) je ekvivalentní následující úloze: k dané matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ hledáme nejbližší (ve smyslu Frobeniovy normy) matici \mathbf{B} hodnosti nejvýše $n - k$. Tedy

$$\min\{\|\mathbf{A} - \mathbf{B}\| \mid \mathbf{B} \in \mathbb{R}^{m \times n}, \text{rank} \mathbf{B} \leq n - k\}. \quad (6.20)$$

Tato úloha je angl. známa jako *low rank approximation*.

Dle (3.17) má matice \mathbf{B} hodnost nejvýše $n - k$ právě tehdy, když její nulový prostor má hodnost nejméně k . To nastane právě tehdy, když existuje matice $\mathbf{X} \in \mathbb{R}^{n \times k}$ s ortonormálními sloupci tak, že $\mathbf{B}\mathbf{X} = \mathbf{0}$. Tedy úloha (6.20) je ekvivalentní úloze

$$\min\{\|\mathbf{A} - \mathbf{B}\| \mid \mathbf{B} \in \mathbb{R}^{m \times n}, \mathbf{X} \in \mathbb{R}^{n \times k}, \mathbf{X}^T \mathbf{X} = \mathbf{I}, \mathbf{B}\mathbf{X} = \mathbf{0}\}. \quad (6.21)$$

Označme řádky matic \mathbf{A}, \mathbf{B} jako $\mathbf{a}_i^T, \mathbf{b}_i^T$. Podmínka $\mathbf{B}\mathbf{X} = \mathbf{0}$ jde psát jako $\mathbf{X}^T \mathbf{b}_i = \mathbf{0}$ ($i = 1, \dots, m$), což znamená, že body $\mathbf{b}_1, \dots, \mathbf{b}_m$ leží v podprostoru $X = \text{null}(\mathbf{X}^T)$. Platí

$$\|\mathbf{A} - \mathbf{B}\|^2 = \|\mathbf{a}_1 - \mathbf{b}_1\|^2 + \dots + \|\mathbf{a}_m - \mathbf{b}_m\|^2. \quad (6.22)$$

Tedy v úlohách (6.20) a (6.21) hledáme body $\mathbf{b}_1, \dots, \mathbf{b}_m$, které leží podprostoru X dimenze nejvýše $n - k$ a jejichž součet čtverců vzdáleností k bodům \mathbf{a}_i je minimální. Viz obrázek v §6.4.1.

Ale pokud \mathbf{X} je pevné, máme

$$\min\{\|\mathbf{a} - \mathbf{b}\| \mid \mathbf{b} \in \mathbb{R}^n, \mathbf{X}^T \mathbf{b} = \mathbf{0}\} = \|\mathbf{X}^T \mathbf{a}\| \quad (6.23)$$

kde minimum se nabývá pro $\mathbf{b} = (\mathbf{I} - \mathbf{X}\mathbf{X}^T)\mathbf{a}$, což je ortogonální projekce bodu \mathbf{a} na podprostor X . Minimalizace (6.22) za podmínky $\mathbf{B}\mathbf{X} = \mathbf{0}$ se pak rozpadá na m nezávislých úloh (6.23):

$$\min\{\|\mathbf{A} - \mathbf{B}\|^2 \mid \mathbf{B} \in \mathbb{R}^{m \times n}, \mathbf{B}\mathbf{X} = \mathbf{0}\} = \|\mathbf{X}^T \mathbf{a}_1\|^2 + \dots + \|\mathbf{X}^T \mathbf{a}_m\|^2 = \|\mathbf{A}\mathbf{X}\|^2,$$

kde minimum se nabývá pro matici $\mathbf{B} = \mathbf{A}(\mathbf{I} - \mathbf{X}\mathbf{X}^T)$.

Tedy úloha (6.21) je ekvivalentní minimalizaci $\|\mathbf{A}\mathbf{X}\|$ za podmínky $\mathbf{X}^T \mathbf{X} = \mathbf{I}$, což je úloha (6.19).

6.4.3 Přeurčené homogenní lineární soustavy

Řešme homogenní lineární soustavu

$$\mathbf{A}\mathbf{x} = \mathbf{0}, \quad (6.24)$$

kde $\mathbf{A} \in \mathbb{R}^{m \times n}$. Její množinou řešení je $\text{null } \mathbf{A}$, což je lineární podprostor \mathbb{R}^n dimenze $n - \text{rank } \mathbf{A}$ viz (3.17). Může být homogenní soustava ‘přeurčená’? Přeurčenost není rozumné definovat jako u nehomogenní soustavy (viz §5.1), protože homogenní soustava má vždy triviální řešení $\mathbf{x} = \mathbf{0}$. Ovšem, kdybychom ‘přeurčenou’ soustavu zkusili přibližně řešit jako minimalizaci $\|\mathbf{A}\mathbf{x}\|$, dostali bychom triviální optimální řešení $\mathbf{x} = \mathbf{0}$.

Abychom se tomu vyhnuli, můžeme navíc požadovat $\mathbf{x}^T \mathbf{x} = 1$. To tedy vede na úlohu

$$\min\{\|\mathbf{A}\mathbf{x}\|^2 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\}. \quad (6.25)$$

Protože $\|\mathbf{A}\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x}$, dle (6.17) je řešením této úlohy vlastní vektor matice $\mathbf{A}^T \mathbf{A}$ příslušný jejímu nejmenšímu vlastnímu číslu.

Příklad 6.8. Kuželosečka (viz §6.3.1) je množina

$$K = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0\}.$$

Budiž dáno m bodů $(x_1, y_1), \dots, (x_m, y_m) \in \mathbb{R}^2$, o kterých víme, že mají ležet na kuželosečce. Body jsou ale zatíženy šumem a tedy obecně na žádné kuželosečce ležet nemusí. Hledejme tedy kuželosečku, která je ‘nejbližší’ daným bodům. Jedna možná formulace této úlohy je, že hledáme

čísla a, b, c, d, e, f , která minimalizují součet čtverců vzdáleností bodů od křivky. Vyřešit přesně tuto úlohu je ale velmi těžké.

Formulujeme proto úlohu přibližně: hledejme čísla a, b, c, d, e, f , která minimalizují součet

$$f(x_1, y_1)^2 + \dots + f(x_m, y_m)^2. \quad (6.26)$$

Tato formulace ale nevyjadřuje to, co chceme, protože minimum výrazu (6.26) se nabývá pro $a = b = c = d = e = f = 0$. V tomto případě množina K není křivka, ale celá rovina \mathbb{R}^2 . Abychom tomu zabránili, uvalíme dodatečnou podmínku

$$a^2 + b^2 + c^2 + d^2 + e^2 + f^2 = 1. \quad (6.27)$$

Minimalizace (6.26) za podmínky (6.27) se dá napsat v maticové formě jako minimalizace $\|\mathbf{M}\mathbf{p}\|^2$ za podmínky $\mathbf{p}^T \mathbf{p} = 1$, kde

$$\mathbf{M} = \begin{bmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m^2 & x_m y_m & y_m^2 & x_m & y_m & 1 \end{bmatrix}, \quad \mathbf{p} = [a \ b \ c \ d \ e \ f]^T. \quad \square$$

Obecněji, soustavu (6.24) nazveme přeurčenou tehdy, když dimenze jejího prostoru řešení je nižší než nějaká předem známá dimenze k . Přibližné řešení soustavy pak vede na úlohu (6.19), která je ovšem ekvivalentní úloze (6.20). Úloha (6.20) odpovídá tomu, že co nejméně změníme matici \mathbf{A} , aby prostor řešení soustavy (6.24) měl kýženou dimenzi k . Neboli nejprve najdeme matici \mathbf{B} s hodnotí $n - k$ nejbližší matici \mathbf{A} a potom řešíme soustavu $\mathbf{B}\mathbf{x} = \mathbf{0}$, jejíž prostor řešení již má dimenzi k .

Vztah k nehomogennímu případu. V §5.1 jsme formulovali přibližné řešení nehomogenní ($\mathbf{b} \neq \mathbf{0}$) soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ jako úlohu $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|$. Může se zdát, že tato formulace je úplně odlišná od formulace přibližného řešení homogenní soustavy, kterou jsme uvedli zde. Ale tak tomu není. Formulujeme přibližné řešení nehomogenní soustavy takto: pokud soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ nemá řešení, změníme vektor \mathbf{b} co nejméně tak, aby soustava řešení měla. Přesněji, hledáme vektor \mathbf{c} tak, aby pro nějaké \mathbf{x} platilo $\mathbf{A}\mathbf{x} = \mathbf{c}$ a přitom číslo $\|\mathbf{b} - \mathbf{c}\|$ bylo co nejmenší. Tuto úlohu lze napsat jako

$$\min\{\|\mathbf{b} - \mathbf{c}\| \mid \mathbf{A}\mathbf{x} = \mathbf{c}, \mathbf{x} \in \mathbb{R}^n, \mathbf{c} \in \mathbb{R}^m\}.$$

Zde minimalizujeme přes proměnné \mathbf{x} a \mathbf{c} (nevadí, že \mathbf{x} se nevyskytuje v účelové funkci). Ale tato úloha jde zjednodušit: dosadíme $\mathbf{c} = \mathbf{A}\mathbf{x}$ do účelové funkce $\|\mathbf{b} - \mathbf{c}\|$, čímž dostaneme $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|$. Shrňme:

- V přibližném řešení nehomogenní soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ chceme změnit vektor \mathbf{b} co nejméně tak, aby soustava měla řešení.
- V přibližném řešení homogenní soustavy $\mathbf{A}\mathbf{x} = \mathbf{0}$ chceme změnit matici \mathbf{A} co nejméně tak, aby soustava měla prostor řešení dané dimenze.

6.5 Cvičení

6.1. Spočítejte vlastní čísla a vlastní vektory matic $\begin{bmatrix} 1 & 2 \\ -1 & -3 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$.

- 6.2. Napište rovnici, jejímiž kořeny jsou vlastní čísla matice $\begin{bmatrix} 2 & 0 & 3 \\ 0 & -2 & -1 \\ 3 & -1 & 2 \end{bmatrix}$.
- 6.3. Najděte všechna vlastní čísla a vlastní vektory (a) nulové, (b) jednotkové, (c) diagonální matice. Najděte vlastní čísla trojúhelníkové matice.
- 6.4. Známe vlastní čísla a vektory matice \mathbf{A} . Jaké jsou vlastní čísla a vektory matice $\mathbf{A} + \alpha \mathbf{I}$?
- 6.5. Máme kvadratickou formu dvou proměnných $f(x, y) = 3x^2 + 2xy + 3y^2$.
- Napište ji ve tvaru $f(x, y) = \begin{bmatrix} x & y \end{bmatrix} \mathbf{A} \begin{bmatrix} x \\ y \end{bmatrix}$ se symetrickou \mathbf{A} .
 - Najděte $a, b \in \mathbb{R}$ a ortogonální \mathbf{U} tak, že $f(x, y) = au^2 + bv^2$, kde $\begin{bmatrix} u \\ v \end{bmatrix} = \mathbf{U} \begin{bmatrix} x \\ y \end{bmatrix}$.
 - Nakreslete množinu bodů (u, v) splňujících $au^2 + bv^2 = 1$.
 - Transformujte tuto množinu do souřadnic (x, y) a nakreslete.

- 6.6. Kvadratickou funkci $f(x, y) = 2x^2 + 2xy - 3y^2 - x + 2y$ napište ve tvaru (6.11) se symetrickou \mathbf{A} .
- 6.7. Je množina $\{(x, y) \in \mathbb{R}^2 \mid x^2 - 3xy + y^2 = 1\}$ elipsa nebo hyperbola? Odůvodněte.
- 6.8. Je známo, že libovolnou rotaci ve třírozměrném prostoru lze realizovat jako rotaci kolem jisté přímky (jdoucí počátkem) o jistý úhel. Geometrickou úvahou (tj. bez počítání) zjistěte co nejvíce o vlastních číslech a vektorech rotační matice rozměru 3×3 .
- 6.9. Ukažte, že je-li $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ spektrální rozklad symetrické matice \mathbf{A} , platí $\mathbf{A}^n = \mathbf{V}\mathbf{\Lambda}^n\mathbf{V}^T$.
- 6.10. (*) Ukažte, že dvě čtvercové matice komutují, právě když mají stejné vlastní vektory.
- 6.11. (*) Necht' $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{B} \in \mathbb{R}^{n \times m}$. Ukažte, že nenulová vlastní čísla matic \mathbf{AB} a \mathbf{BA} jsou stejná.
- 6.12. Pro dané matice určete, zda jsou pozitivně/negativně (semi)definitní nebo indefinitní:

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

- 6.13. Zjistěte, zda mají následující kvadratické funkce extrém a pokud ano, extrém najděte a určete jeho druh (minimum nebo maximum). Použijte doplnění na čtverec.
- $f(x, y) = x^2 + 4xy - 2y^2 + 3x - 6y + 5$
 - $f(\mathbf{x}) = \mathbf{x}^T \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \mathbf{x} + [2 \quad -1] \mathbf{x}$
- 6.14. Mějme matici $\mathbf{A} = \begin{bmatrix} 1 & -3 \\ 2 & -4 \end{bmatrix}$. Která z následujících tvrzení jsou pravdivá?
- Výraz $\mathbf{x}^T \mathbf{A} \mathbf{x}$ je nezáporný pro každé $\mathbf{x} \in \mathbb{R}^2$.
 - Výraz $\mathbf{x}^T \mathbf{A} \mathbf{x}$ je nekladný pro každé $\mathbf{x} \in \mathbb{R}^2$.
 - Funkce $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ má v bodě $\mathbf{x} = \mathbf{0}$ extrém.
- 6.15. (*) Napište v Matlabu funkci `ellipse(Q)`, která vykreslí elipsu s rovnicí $\mathbf{x}^T \mathbf{A} \mathbf{x} = 1$ pro pozitivně definitní \mathbf{A} . Zamyslete se, jak byste postupovali při návrhu funkce `conic(Q)`, která vykreslí kuželosečku $\mathbf{x}^T \mathbf{A} \mathbf{x} = 1$ pro \mathbf{A} libovolné definitnosti (nezapomeňte, že obecná kuželosečka může být neomezená, tedy je nutno ji oříznout do daného obdélníku).

- 6.16. Dokažte, že matice $\mathbf{A}^T \mathbf{A}$ je pozitivně semidefinitní pro každou matici \mathbf{A} .
- 6.17. Dokažte, že matice $\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}$ je pozitivně definitní pro každou matici \mathbf{A} a každé $\mu > 0$.
- 6.18. Dokažte, že (čtvercová symetrická) matice je pozitivně definitní právě tehdy, když její inverze je pozitivně definitní. Dokažte to nejdříve s použitím spektrálního rozkladu a pak bez použití spektrálního rozkladu.
- 6.19. Musí mít pozitivně semidefinitní matice na diagonále nezáporné prvky? Odpověď dokažte.
- 6.20. (*) Pozitivně semidefinitní symetrické matice lze vnímat jako zobecnění nezáporných čísel. Proto se někdy pozitivní semidefinitnost značí $\mathbf{A} \succeq 0$. Zápis $\mathbf{A} \succeq \mathbf{B}$ je pak zkratkou $\mathbf{A} - \mathbf{B} \succeq 0$. Na základě této analogie bychom očekávali, že:
- Pokud $\mathbf{A} \succeq \mathbf{B}$ a $\mathbf{C} \succeq \mathbf{D}$, potom $\mathbf{A} + \mathbf{C} \succeq \mathbf{B} + \mathbf{D}$.
 - Pokud $\mathbf{A} \succeq 0$ a $\alpha \geq 0$, potom $\alpha \mathbf{A} \succeq 0$.
 - Pokud $\mathbf{A} \succeq 0$, potom $\mathbf{A}^2 \succeq 0$.
 - Pokud $\mathbf{A} \succeq 0$ a $\mathbf{B} \succeq 0$, potom $\mathbf{AB} \succeq 0$.
 - Pokud $\mathbf{A} \succeq 0$ a $\mathbf{B} \succeq 0$, potom $\mathbf{ABA} \succeq 0$.

Které z těchto tvrzení platí a která neplatí? Odpovědi dokažte.

- 6.21. Dokažte, že relace \succeq z předchozího cvičení je částečné uspořádání, tj. reflexivní, tranzitivní a antisymetrická relace na množině symetrických matic $n \times n$.
- 6.22. Jsou-li λ_i vlastní čísla matice \mathbf{A} , dokažte $\lambda_1 \times \dots \times \lambda_n = \det \mathbf{A}$ pro
- libovolnou diagonalizovatelnou matici,
 - (*) pro libovolnou čtvercovou matici.
- 6.23. V §4.4 jsme definovali projekci jako matici \mathbf{P} splňující $\mathbf{P}^2 = \mathbf{P}$. Geometrickou úvahou (tedy bez počítání) najděte aspoň jedno vlastní číslo a příslušný vlastní vektor projekce.
- 6.24. (*) Geometrickou úvahou najděte aspoň dva vlastní vektory a příslušná vlastní čísla Householderovy matice ze Cvičení 4.13.
- 6.25. Dokažte, že pro libovolnou matici \mathbf{A} jsou nenulová vlastní čísla matic $\mathbf{A}^T \mathbf{A}$ a $\mathbf{A} \mathbf{A}^T$ stejná.
- 6.26. Pro čtvercovou matici \mathbf{A} je výraz $\frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$ znám jako **Rayleighův kvocient**. Dokažte, že

$$\min_{\mathbf{x} \in \mathbb{R}^n} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \min\{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\}.$$

Nápověda a řešení

- 6.4. Vlastní čísla se zmenší o α . Vlastní vektory jsou stejné.
- 6.5.a) $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$
- 6.5.b) $a = 2, b = 4, \mathbf{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$
- 6.6. $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}, \mathbf{b} = [-1 \quad 2], c = 0$
- 6.7. Hyperbola.
- 6.12. indefinitní, pozitivně definitní, indefinitní, pozitivně semidefinitní

- 6.13.a) Nemá extrém (má sedlo v bodě $(2, \frac{7}{4})$).
- 6.13.b) Má minimum v bodě $-(3, 1)/2$.
- 6.14. Žádné tvrzení neplatí. Náponěda: je matice symetrická?
- 6.19. Musí.
- 6.22.a) Použijte spektrální rozklad $\mathbf{A} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}$.
- 6.25. Vynásobte rovnici $\mathbf{A}^T\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ zleva maticí \mathbf{A} .

Kapitola 7

Singulární rozklad

Věta 7.1. Každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ hodnosti r lze rozložit jako

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \quad (7.1)$$

kde $\mathbf{S} \in \mathbb{R}^{r \times r}$ je diagonální regulární a $\mathbf{U} \in \mathbb{R}^{m \times r}$ a $\mathbf{V} \in \mathbb{R}^{n \times r}$ mají ortogonální sloupce.

Důkaz. Zkoumejme matici $\mathbf{A}^T\mathbf{A}$. Tato matice je symetrická a pozitivně semidefinitní (Cvičení 6.16), tedy má nezáporná vlastní čísla. Dle (5.5) je $\text{rank}(\mathbf{A}^T\mathbf{A}) = \text{rank} \mathbf{A} = r$. Pišme její spektrální rozklad (viz §6.1.1) jako

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\mathbf{S}^2\mathbf{V}^T, \quad (7.2)$$

kde $\mathbf{S} \in \mathbb{R}^{r \times r}$ má na diagonále nezáporná čísla (odmocniny z vlastních čísel matice $\mathbf{A}^T\mathbf{A}$) a $\mathbf{V} \in \mathbb{R}^{n \times r}$. Položme

$$\mathbf{U} = \mathbf{A}\mathbf{V}\mathbf{S}^{-1}.$$

Ukážeme, že matice $\mathbf{U}, \mathbf{S}, \mathbf{V}$ tvoří rozklad (7.1). Ověříme tedy, že

$$\mathbf{U}^T\mathbf{U} = \mathbf{S}^{-1}\mathbf{V}^T\mathbf{A}^T\mathbf{A}\mathbf{V}\mathbf{S}^{-1} = \mathbf{S}^{-1}\mathbf{V}^T\mathbf{V}\mathbf{S}^2\mathbf{V}^T\mathbf{V}\mathbf{S}^{-1} = \mathbf{S}^{-1}\mathbf{S}^2\mathbf{S}^{-1} = \mathbf{I}, \quad (7.3a)$$

$$\mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{A}\mathbf{V}\mathbf{S}^{-1}\mathbf{S}\mathbf{V}^T = \mathbf{A}\mathbf{V}\mathbf{V}^T = \mathbf{A}. \quad (7.3b)$$

Rovnost $\mathbf{V}\mathbf{V}^T = \mathbf{A}$ v (7.3b) je očividná, pokud \mathbf{V} je čtvercová a tedy $\mathbf{V}\mathbf{V}^T = \mathbf{I}$. Rovnost ale platí, i když \mathbf{V} není čtvercová (tedy je úzká). Toto je těžší část důkazu. Máme

$$\text{rng}(\mathbf{A}^T) = \text{rng}(\mathbf{A}^T\mathbf{A}) = \text{rng}(\mathbf{V}\mathbf{A}\mathbf{V}^T) = \text{rng} \mathbf{V}, \quad (7.4)$$

kde první rovnost plyne ze (5.4b), druhá rovnost ze (7.2), a třetí rovnost z Věty 3.4 protože matice $\mathbf{A}\mathbf{V}^T$ má lineárně nezávislé řádky. Dle (7.4) je každý sloupec \mathbf{A}^T lineární kombinací sloupců \mathbf{V} , což lze psát jako $\mathbf{A}^T = \mathbf{V}\mathbf{B}$. Dosazením máme $\mathbf{A}\mathbf{V}\mathbf{V}^T = \mathbf{B}^T\mathbf{V}^T\mathbf{V}\mathbf{V}^T = \mathbf{B}^T\mathbf{V}^T = \mathbf{A}$.

Lze to uvídnout i jinak. Dle (7.4) je $\mathbf{V}\mathbf{V}^T$ projektor na podprostor $\text{rng} \mathbf{V} = \text{rng}(\mathbf{A}^T)$ a proto pro každé $\mathbf{x} \in \text{rng}(\mathbf{A}^T)$ platí $\mathbf{V}\mathbf{V}^T\mathbf{x} = \mathbf{x}$ (viz §4.4). Protože každý sloupec matice \mathbf{A}^T patří do $\text{rng}(\mathbf{A}^T)$, máme $\mathbf{V}\mathbf{V}^T\mathbf{A}^T = \mathbf{A}^T$ neboli $\mathbf{A}\mathbf{V}\mathbf{V}^T = \mathbf{A}$. \square

Rozklad (7.1) se nazývá **singulární rozklad** (*singular value decomposition, SVD*) matice \mathbf{A} . Diagonální prvky matice \mathbf{S} se nazývají **singulární čísla** matice \mathbf{A} . Budeme je značit s_1, \dots, s_r a je zvykem je sestupně seřadit,

$$s_1 \geq \dots \geq s_r > 0,$$

což lze vždy udělat permutací sloupců \mathbf{U} a \mathbf{V} či vynásobením některých sloupců číslem -1 . Sloupce matice \mathbf{U} příp. \mathbf{V} se nazývají levé příp. pravé **singulární vektory** matice \mathbf{A} .

Platí

$$\text{rng } \mathbf{U} = \text{rng } \mathbf{A}, \quad (7.5a)$$

$$\text{rng } \mathbf{V} = \text{rng}(\mathbf{A}^T), \quad (7.5b)$$

kde (7.5a) plyne z Věty 3.4, neboť $\text{rng } \mathbf{A} = \text{rng}(\mathbf{USV}^T) = \text{rng}(\mathbf{US}) = \text{rng } \mathbf{U}$. Rovnost (7.5b) je rovnost (7.5a) použitá na \mathbf{A}^T .

Důkaz ukazuje, že SVD matice \mathbf{A} jde snadno spočítat ze spektrálního rozkladu matice $\mathbf{A}^T\mathbf{A}$ (stejně bychom to mohli udělat s maticí \mathbf{AA}^T). To ale není numericky nejvhodnější, protože výpočet součinu $\mathbf{A}^T\mathbf{A}$ může vést ke zbytečným zaokrouhlovacím chybám (viz §5.1.1). Na SVD se proto používají algoritmy, které se explicitnímu výpočtu $\mathbf{A}^T\mathbf{A}$ vyhýbají. Úlohy (6.19) a (6.20) se tedy kvůli zaokrouhlovacím chybám mají řešit pomocí SVD algoritmu. Na druhou stranu, pokud nám možné snížení přesnosti nevádí, počítání SVD spektrálním rozkladem může být rychlejší: např. když $m \gg n$ a potřebujeme spočítat jen matice \mathbf{V} a \mathbf{S} (nepotřebujeme \mathbf{U}), spektrální rozklad matice $\mathbf{A}^T\mathbf{A}$ bude rychlejší (tato matice je malá, $n \times n$).

Poznámka o numerické lineární algebře. Potkali jsme již tři rozklady matic: QR, spektrální rozklad, SVD. Je ještě několik jiných zajímavých rozkladů. Návrh algoritmů na operace s maticemi, řešení soustav lineárních rovnic a rozklady matic vektorů je předmětem *numerické lineární algebry*. Cílem je nalézt algoritmy, které jsou rychlé a odolné vůči zaokrouhlovacím chybám. Existují volně dostupné softwarové balíky na numerickou lineární algebru, např. LAPACK a BLAS. Matlab je postaven na balíku LAPACK.

7.1 Rozšířené formy SVD

Matice \mathbf{U} a \mathbf{V} je možno doplnit přidáním sloupců na širší matice s ortonormálními sloupci a matici \mathbf{S} doplnit přidáním nul na diagonální matici odpovídající velikosti¹:

$$\mathbf{A} = \mathbf{USV}^T = \underbrace{[\mathbf{U} \quad \bar{\mathbf{U}}]}_{\mathbf{U}'} \underbrace{\begin{bmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\mathbf{S}'} \underbrace{[\mathbf{V} \quad \bar{\mathbf{V}}]}_{\mathbf{V}'^T}, \quad (7.6)$$

kde rovnost ověříme vynásobením blokových matic (viz §2.7). Vidíme, že přidané matice $\bar{\mathbf{U}}$, $\bar{\mathbf{V}}$ mohou být libovolné až na požadavek, že matice $[\mathbf{U} \quad \bar{\mathbf{U}}]$ a $[\mathbf{V} \quad \bar{\mathbf{V}}]$ musejí mít ortonormální sloupce. Jsou dvě přirozené formy tohoto rozšířeného SVD:

- $\mathbf{S}' \in \mathbb{R}^{p \times p}$, $\mathbf{U}' \in \mathbb{R}^{m \times p}$, $\mathbf{V}' \in \mathbb{R}^{n \times p}$, kde $p = \min\{m, n\}$,
- $\mathbf{S}' \in \mathbb{R}^{m \times n}$, $\mathbf{U}' \in \mathbb{R}^{m \times m}$, $\mathbf{V}' \in \mathbb{R}^{n \times n}$.

V obou případech matice \mathbf{S}' má p diagonálních prvků, z nichž je tedy $p - r$ nulových. To jsou všechna singulární čísla matice \mathbf{A} . Budeme je značit s_1, \dots, s_p a je zvykem je zvolit kladné a sestupně je seřadit:

$$s_1 \geq \dots \geq s_r > s_{r+1} = \dots = s_p = 0.$$

To lze vždy udělat permutací sloupců \mathbf{U} a \mathbf{V} či vynásobením některých sloupců číslem -1 .

¹Připomeňme, že diagonální matice může být i obdélníková, viz §2.

Rovnost tří forem SVD je také dobře vidět, napíšeme-li SVD jako dyadický rozvoj (srov. se (6.6)). Pak

$$\mathbf{A} = \mathbf{USV}^T = \sum_{i=1}^r s_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=1}^p s_i \mathbf{u}_i \mathbf{v}_i^T = \mathbf{U}' \mathbf{S}' \mathbf{V}'^T. \quad (7.7)$$

Obě sumy se očividně rovnají, protože pro $r < i \leq p$ je $s_i = 0$.

Matlabské příkazy `[U,S,V]=svd(A)` a `[U,S,V]=svd(A,'econ')` počítají druhou a první rozšířenou formu SVD.

Příklad 7.1. Zde je příklad obyčejného a rozšířeného SVD matice 2×3 :

$$\begin{aligned} \mathbf{A} = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} &= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \end{bmatrix} \\ &= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ -2/3 & 2/3 & 1/3 \end{bmatrix} \quad \square \end{aligned}$$

7.2 Nejbližší matice nižší hodnosti

Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $q \leq r = \text{rank } \mathbf{A}$. V §6.4.2 jsme ukázali, že řešením úlohy

$$\min\{\|\mathbf{A} - \mathbf{B}\| \mid \mathbf{B} \in \mathbb{R}^{m \times n}, \text{rank } \mathbf{B} \leq q\} \quad (7.8)$$

je $\mathbf{B} = \mathbf{A}(\mathbf{I} - \mathbf{XX}^T)$, kde sloupce matice $\mathbf{X} \in \mathbb{R}^{n \times (n-q)}$ jsou vlastní vektory příslušné $n - q$ nejmenším vlastním číslům matice $\mathbf{A}^T\mathbf{A}$. Tento výsledek lze elegantně formulovat pomocí SVD.

Věta 7.2 (Eckart-Young). Necht' $\mathbf{A} = \mathbf{USV}^T$ je SVD matice \mathbf{A} . Řešení úlohy (7.8) je

$$\mathbf{B} = \mathbf{U}' \mathbf{S}' \mathbf{V}'^T = \sum_{i=1}^q s_i \mathbf{u}_i \mathbf{v}_i^T, \quad (7.9)$$

kde $\mathbf{U}' \in \mathbb{R}^{m \times q}$ je prvních q sloupců matice \mathbf{U} , $\mathbf{V}' \in \mathbb{R}^{n \times q}$ je prvních q sloupců matice \mathbf{V} , a $\mathbf{S}' \in \mathbb{R}^{q \times q}$ je prvních q sloupců+řádků matice \mathbf{S} .

Důkaz. Z §6.4.2 víme, že sloupce \mathbf{X} jsou vlastní vektory příslušné $n - q$ nejmenším vlastním číslům matice $\mathbf{A}^T\mathbf{A}$. Z důkazu Věty 7.1 víme, že sloupce \mathbf{V}' jsou vlastní vektory příslušné q největším vlastním číslům matice $\mathbf{A}^T\mathbf{A}$. Tedy $[\mathbf{V}_r \quad \mathbf{X}]$ je ortogonální matice. Dle (4.16) tedy $\mathbf{I} - \mathbf{XX}^T = \mathbf{V}'\mathbf{V}'^T$. Nyní máme

$$\mathbf{B} = \mathbf{A}(\mathbf{I} - \mathbf{XX}^T) = \mathbf{AV}'\mathbf{V}'^T = \mathbf{USV}^T\mathbf{V}'\mathbf{V}'^T = \mathbf{US} \begin{bmatrix} \mathbf{I}_q \\ \mathbf{0} \end{bmatrix} \mathbf{V}'^T = \mathbf{U} \begin{bmatrix} \mathbf{S}' \\ \mathbf{0} \end{bmatrix} \mathbf{V}'^T = \mathbf{U}' \mathbf{S}' \mathbf{V}'^T. \quad \square$$

Ze formy (7.9) jako dyadického rozvoje vidíme, že suma je vlastně prvních q členů sumy (7.7). Optimální hodnota úlohy (7.8) je

$$\|\mathbf{A} - \mathbf{B}\| = \|\mathbf{USV}^T - \mathbf{U}'\mathbf{S}'\mathbf{V}'^T\| = \|\mathbf{U}(\mathbf{S} - \mathbf{S}')\mathbf{V}^T\| = \|\mathbf{S} - \mathbf{S}'\| = (s_{q+1}^2 + \dots + s_r^2)^{1/2}. \quad (7.10)$$

V tomto smyslu singulární čísla udávají vzdálenost matice k nejbližší matice dané nižší hodnosti. Všimněte si dále, že podle (7.5) singulární vektory nejen definují ortonormální báze podprostorů $\text{rng } \mathbf{A}$ a $\text{rng}(\mathbf{A}^T)$, ale říkají také, jak by se tyto prostory změnilo, kdyby se matice nahradila nejbližší maticí dané nižší hodnosti.

7.3 Pseudoinverze obecné matice

Vraťme se nyní do kapitoly 5 k nehomogenní lineární soustavě, $\mathbf{Ax} = \mathbf{b}$ pro $\mathbf{A} \in \mathbb{R}^{m \times n}$. Tam jsme odděleně diskutovali případy, kdy soustava má žádné, jedno, nebo nekonečně mnoho řešení. Ted' všechny tyto případy spojíme do jediné formulace. Zopakujeme, že množina optimálních řešení úlohy (5.2) je rovna množině řešení soustavy normálních rovnic (5.3), a to pro libovolnou matici \mathbf{A} . Co když je ale sama soustava (5.3) nedourčená? V tom případě můžeme najít její řešení s nejmenší normou. Tedy řešíme úlohu

$$\min\{\|\mathbf{x}\|^2 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}\}. \quad (7.11)$$

Protože úloha (5.11) má pro každou matici \mathbf{A} právě jedno optimální řešení, má i úloha (7.11) pro každou matici \mathbf{A} právě jedno optimální řešení. Toto řešení \mathbf{x} má následující vlastnosti:

- Pokud soustava $\mathbf{Ax} = \mathbf{b}$ má jediné řešení, \mathbf{x} je toto řešení.
- Pokud soustava $\mathbf{Ax} = \mathbf{b}$ nemá řešení, \mathbf{x} je její přibližné řešení ve smyslu nejmenších čtverců, tj. řešení problému (5.2). Pokud ovšem problém (5.2) má více než jedno (tedy nekonečně mnoho) řešení, \mathbf{x} je řešení problému (5.2) s nejmenší normou.
- Pokud soustava $\mathbf{Ax} = \mathbf{b}$ má nekonečně mnoho řešení, \mathbf{x} je řešení této soustavy s nejmenší normou, tj. řešení problému (5.11).

Lze ukázat, že řešení úlohy (7.11) lze opět psát jako $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$ kde matici \mathbf{A}^+ nazýváme **pseudoinverze** (přesněji: Mooreova-Penroseova pseudoinverze) obecné matice \mathbf{A} . Když \mathbf{A} má lineárně nezávislé sloupce (tedy je úzká s plnou hodnotí), pseudoinverze je rovna (5.6). Když \mathbf{A} má lineárně nezávislé řádky (tedy je široká s plnou hodnotí), pseudoinverze je rovna (5.14). Když \mathbf{A} je regulární (tedy čtvercová s plnou hodnotí), je $\mathbf{A}^+ = \mathbf{A}^{-1}$. Když ovšem \mathbf{A} nemá plnou hodnot, pseudoinverzi je nutno počítat jinak. Elegantně se to udělá pomocí SVD.

Věta 7.3. *Necht' $\mathbf{A} = \mathbf{USV}^T$ je redukované SVD matice \mathbf{A} . Pak pseudoinverze matice \mathbf{A} je*

$$\mathbf{A}^+ = \mathbf{VS}^{-1} \mathbf{U}^T. \quad (7.12)$$

Důkaz. Podle §5.2, \mathbf{x} je řešení úlohy (7.11) právě tehdy, když existuje \mathbf{y} takové, že

$$\mathbf{A}^T \mathbf{Ay} = \mathbf{x}, \quad (7.13a)$$

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}. \quad (7.13b)$$

Máme $\mathbf{x} = \mathbf{A}^+ \mathbf{b} = \mathbf{VS}^{-1} \mathbf{U}^T \mathbf{b}$. Necht' $\mathbf{y} = \mathbf{VS}^{-2} \mathbf{V}^T \mathbf{x}$. Ukážeme, že (\mathbf{x}, \mathbf{y}) splňují (7.13).

Je $\mathbf{A}^T \mathbf{A} = \mathbf{VS}^T \mathbf{U}^T \mathbf{USV}^T = \mathbf{VS}^2 \mathbf{V}^T$. Máme

$$\begin{aligned} \mathbf{A}^T \mathbf{Ay} &= \mathbf{VS}^2 \mathbf{V}^T \mathbf{y} = \mathbf{VS}^2 \mathbf{V}^T \mathbf{VS}^{-2} \mathbf{V}^T \mathbf{x} = \mathbf{x}, \\ \mathbf{A}^T \mathbf{Ax} &= \mathbf{VS}^2 \mathbf{V}^T \mathbf{x} = \mathbf{VS}^2 \mathbf{V}^T \mathbf{VS}^{-1} \mathbf{U}^T \mathbf{b} = \mathbf{VSU}^T \mathbf{b} = \mathbf{A}^T \mathbf{b}. \end{aligned} \quad \square$$

7.3.1 Pseudoinverze z regularizace

Pro libovolnou matici \mathbf{A} a vektor \mathbf{b} chceme přibližně řešit soustavu $\mathbf{Ax} = \mathbf{b}$ a zároveň chceme, aby norma řešení \mathbf{x} nebyla moc velká. To lze formulovat jako

$$\min_{\mathbf{x} \in \mathbb{R}^n} (\|\mathbf{Ax} - \mathbf{b}\|^2 + \mu \|\mathbf{x}\|^2) \quad (7.14)$$

pro zvolenou váhu $\mu > 0$. Přidání členu $\mu \|\mathbf{x}\|^2$ se říká **regularizace** úlohy (5.2).

Pro každé $\mu > 0$ má úloha (7.14) právě jedno řešení $\mathbf{x} = \mathbf{A}_\mu^+ \mathbf{b}$, kde

$$\mathbf{A}_\mu^+ = (\mathbf{A}^T \mathbf{A} + \mu \mathbf{I})^{-1} \mathbf{A}^T = \mathbf{A}^T (\mathbf{AA}^T + \mu \mathbf{I})^{-1}. \quad (7.15)$$

Matice $\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}$ a $\mathbf{AA}^T + \mu \mathbf{I}$ jsou regulární pro $\mu > 0$ (viz Cvičení 6.16), tedy oba výrazy v (7.15) jsou definovány. Snadno ověříme, že oba výrazy ve (7.15) se rovnají: vynásobíme oba výrazy zleva maticí $\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}$ a zprava maticí $\mathbf{AA}^T + \mu \mathbf{I}$ a pak roznásobíme závorky.

Nyní zvolme μ velmi malé, v limitě $\mu \rightarrow 0+$. Definujme matici

$$\mathbf{A}^+ = \lim_{\mu \rightarrow 0+} \mathbf{A}_\mu^+ = \lim_{\mu \rightarrow 0+} (\mathbf{A}^T \mathbf{A} + \mu \mathbf{I})^{-1} \mathbf{A}^T = \lim_{\mu \rightarrow 0+} \mathbf{A}^T (\mathbf{AA}^T + \mu \mathbf{I})^{-1}. \quad (7.16)$$

Věta 7.4. *Obě limity v (7.16) existují a obě jsou rovny pseudoinverzi matice \mathbf{A} .*

Důkaz. Necht' $\mathbf{A} = \mathbf{USV}^T$ je redukované SVD matice \mathbf{A} . Nejdříve ukážeme, že

$$\mathbf{A}_\mu^+ = \mathbf{VS}_\mu^+ \mathbf{U}^T, \quad \text{kde } \mathbf{S}_\mu^+ = (\mathbf{S}^T \mathbf{S} + \mu \mathbf{I})^{-1} \mathbf{S}^T = (\mathbf{S}^2 + \mu \mathbf{I})^{-1} \mathbf{S}.$$

Po dosazení z (7.15) máme tedy ukázat, že $(\mathbf{A}^T \mathbf{A} + \mu \mathbf{I})^{-1} \mathbf{A}^T = \mathbf{V}(\mathbf{S}^2 + \mu \mathbf{I})^{-1} \mathbf{S} \mathbf{U}^T$, neboli

$$\mathbf{A}^T = (\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}) \mathbf{V}(\mathbf{S}^2 + \mu \mathbf{I})^{-1} \mathbf{S} \mathbf{U}^T.$$

Úpravami pravé strany této rovnosti ukážeme, že se rovná levé straně:

$$\begin{aligned} (\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}) \mathbf{V}(\mathbf{S}^2 + \mu \mathbf{I})^{-1} \mathbf{S} \mathbf{U}^T &= (\mathbf{VSU}^T \mathbf{USV}^T + \mu \mathbf{I}) \mathbf{V}(\mathbf{S}^2 + \mu \mathbf{I})^{-1} \mathbf{S} \mathbf{U}^T \\ &= (\mathbf{VS}^2 \mathbf{V}^T + \mu \mathbf{I}) \mathbf{V}(\mathbf{S}^2 + \mu \mathbf{I})^{-1} \mathbf{S} \mathbf{U}^T \\ &= (\mathbf{VS}^2 \mathbf{V}^T \mathbf{V} + \mu \mathbf{V})(\mathbf{S}^2 + \mu \mathbf{I})^{-1} \mathbf{S} \mathbf{U}^T \\ &= (\mathbf{VS}^2 + \mu \mathbf{V})(\mathbf{S}^2 + \mu \mathbf{I})^{-1} \mathbf{S} \mathbf{U}^T \\ &= \mathbf{V}(\mathbf{S}^2 + \mu \mathbf{I})(\mathbf{S}^2 + \mu \mathbf{I})^{-1} \mathbf{S} \mathbf{U}^T \\ &= \mathbf{VSU}^T = \mathbf{A}^T \end{aligned}$$

Protože \mathbf{S} je čtvercová diagonální, snadno spočítáme (ověřte!), že

$$\mathbf{S}_\mu^+ = (\mathbf{S}^2 + \mu \mathbf{I})^{-1} \mathbf{S} = \text{diag}(s_1/(s_1^2 + \mu), \dots, s_r/(s_r^2 + \mu)).$$

Limitu $\mathbf{S}^+ = \lim_{\mu \rightarrow 0+} \mathbf{S}_\mu^+$ nyní spočítáme jednoduše tak, že spočítáme limitu každého prvku $s_i/(s_i^2 + \mu)$ zvlášť. Je (spočítejte!)

$$\lim_{\mu \rightarrow 0+} \frac{s}{s^2 + \mu} = \frac{1}{s},$$

tedy $\mathbf{S}^+ = \text{diag}(1/s_1, \dots, 1/s_r) = \text{diag}(s_1, \dots, s_r)^{-1} = \mathbf{S}^{-1}$. Z toho již

$$\mathbf{A}^+ = \lim_{\mu \rightarrow 0+} \mathbf{A}_\mu^+ = \lim_{\mu \rightarrow 0+} \mathbf{VS}_\mu^+ \mathbf{U}^T = \mathbf{VS}^+ \mathbf{U}^T = \mathbf{VS}^{-1} \mathbf{U}^T.$$

Což dle Věty 7.3 je pseudoinverze matice \mathbf{A} . □

7.4 (★) Nejblížejší ortogonální matice

Hledejme ortogonální matici \mathbf{R} , která je nejblížejší (ve Frobeniově normě) dané čtvercové matici \mathbf{A} . Tedy minimalizujeme $\|\mathbf{A} - \mathbf{R}\|^2$ za podmínky $\mathbf{R}^T\mathbf{R} = \mathbf{I}$. Máme

$$\begin{aligned}\|\mathbf{A} - \mathbf{R}\|^2 &= \text{tr}((\mathbf{A} - \mathbf{R})^T(\mathbf{A} - \mathbf{R})) = \text{tr}(\mathbf{A}^T\mathbf{A} - \mathbf{A}^T\mathbf{R} - \mathbf{R}^T\mathbf{A} + \mathbf{R}^T\mathbf{R}) \\ &= \text{tr}(\mathbf{A}^T\mathbf{A}) - \text{tr}(\mathbf{A}^T\mathbf{R}) - \text{tr}(\mathbf{R}^T\mathbf{A}) + \text{tr}\mathbf{I} \\ &= \text{tr}(\mathbf{A}^T\mathbf{A}) - 2\text{tr}(\mathbf{A}^T\mathbf{R}) + \text{tr}\mathbf{I},\end{aligned}$$

kde jsme použili rovnost $\text{tr}(\mathbf{A}^T\mathbf{R}) = \text{tr}(\mathbf{R}^T\mathbf{A})$ (viz §2.3 a §4.6). Protože $\text{tr}(\mathbf{A}^T\mathbf{A}) + \text{tr}\mathbf{I}$ je konstanta, místo minimalizace $\|\mathbf{A} - \mathbf{R}\|^2$ můžeme maximalizovat $\text{tr}(\mathbf{A}^T\mathbf{R})$.

Nechť $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ je redukované SVD matice \mathbf{A} . Zavedme substituci $\mathbf{Z} = \mathbf{U}^T\mathbf{R}\mathbf{V}$. Matice \mathbf{Z} je ortogonální a proto každý její prvek splňuje $|z_{ij}| \leq 1$. Nyní

$$\text{tr}(\mathbf{A}^T\mathbf{R}) = \text{tr}(\mathbf{V}\mathbf{S}\mathbf{U}^T\mathbf{R}) = \text{tr}(\mathbf{S}\mathbf{U}^T\mathbf{R}\mathbf{V}) = \text{tr}(\mathbf{S}\mathbf{Z}) = \sum_i s_i z_{ii}.$$

Jelikož $s_i \geq 0$ a $|z_{ii}| \leq 1$, poslední výraz nabývá maxima pro $z_{ii} = 1$, což se realizuje pro $\mathbf{Z} = \mathbf{I}$. To odpovídá $\mathbf{R} = \mathbf{U}\mathbf{V}^T$.

Ortogonalní Prokrustův problém: $\min \|\mathbf{R}\mathbf{A} - \mathbf{B}\|$ za podmínky $\mathbf{R}^T\mathbf{R} = \mathbf{I}$, kde \mathbf{A}, \mathbf{B} jsou široké. Máme

$$\|\mathbf{R}\mathbf{A} - \mathbf{B}\|^2 = \text{tr}((\mathbf{R}\mathbf{A} - \mathbf{B})^T(\mathbf{R}\mathbf{A} - \mathbf{B})) = \text{tr}(\mathbf{A}^T\mathbf{A}) - \text{tr}(\mathbf{A}^T\mathbf{R}^T\mathbf{B}) - \text{tr}(\mathbf{B}^T\mathbf{R}\mathbf{A}) + \text{tr}(\mathbf{B}^T\mathbf{B}).$$

Ale $\text{tr}(\mathbf{A}^T\mathbf{R}^T\mathbf{B}) = \text{tr}(\mathbf{B}^T\mathbf{R}\mathbf{A}) = \text{tr}(\mathbf{R}\mathbf{A}\mathbf{B}^T)$. Tedy maximalizujeme $\text{tr}(\mathbf{R}\mathbf{A}\mathbf{B}^T)$, což je to samé jako nalezení nejblížejší ortogonální matice k $\mathbf{B}\mathbf{A}^T$.

7.5 Cvičení

7.1. Jsou dány matice

$$\mathbf{A} = \frac{1}{15} \begin{bmatrix} 14 & -16 & -4 \\ -2 & 13 & 22 \end{bmatrix}, \quad \mathbf{U} = \frac{1}{5} \begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix}, \quad \mathbf{V} = \frac{1}{3} \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$$

Spočítejte matici \mathbf{B} hodnosti jedna takovou, že $\|\mathbf{A} - \mathbf{B}\|$ je minimální (kde $\|\cdot\|$ značí Frobeniovu normu). Bez užití počítače spočítejte hodnotu $\|\mathbf{A} - \mathbf{B}\|$ pro tuto matici \mathbf{B} .

7.2. Jsou dány čtyři body $\mathbf{a}_1 = (3, -3, 4)$, $\mathbf{a}_2 = (-2, -3, -2)$, $\mathbf{a}_3 = (1, 0, -1)$, $\mathbf{a}_4 = (3, 1, 0)$ v \mathbb{R}^3 . Najděte množinu $X \subseteq \mathbb{R}^3$, která minimalizuje součet čtverců kolmých vzdáleností bodů k množině X , kde X je

- přímka procházející počátkem,
- rovina procházející počátkem,
- přímka která může ale nemusí procházet počátkem.

Můžete použít počítač.

7.3. S použitím SVD najděte bázi ortogonálního doplňku podprostoru $\text{span}\{(1, 1, 1, -1), (2, -1, -1, 1), (-1, 2, 2, 1)\}$. Můžete použít počítač.

7.4. (★) Vyřešte soustavu ze Cvičení 5.2 přibližně ve smyslu nejmenších čtverců pomocí SVD. Můžete použít počítač.

7.5. (★) Dokažte vlastnosti pseudoinverze ze Cvičení 5.9 pomocí vztahu (7.12) pro libovolné (čtvercové či obdélníkové) matice z libovolnou hodnotí.

7.6. Známe spektrální rozklad symetrické matice. Jak byste jednoduše našli SVD této matice?

7.7. Možná jste si již dříve položili otázku, jaký je vhodný algoritmus na nalezení báze nulového prostoru dané matice. Cvičení 3.11 naznačilo jeden způsob. V Matlabu se ovšem báze nulového prostoru najde snadno funkcí `null`. Vypíšte si implementaci této funkce matlabským příkazem `edit null` a najděte souvislost se vzorcí (7.5).

Nápověda a řešení

7.1. $\|\mathbf{A} - \mathbf{B}\| = 1$

7.2.a) $X = \text{span}\{(0.6912, -0.2181, 0.6889)\}$ (Uvědomte si ale, že báze vašeho podprostoru může být jiná než zde uvedená.)

7.2.b) $X = \text{span}\{(0.6912, -0.2181, 0.6889), (-0.3771, -0.9221, 0.0864)\}$

7.2.c) $X = (1.25, -1.25, 0.25) + \text{span}\{(0.6599, 0.0280, 0.7508)\}$

7.3. Hledaná báze je např. $(0, -1, 1, 0)$.

Kapitola 8

Nelineární funkce a zobrazení

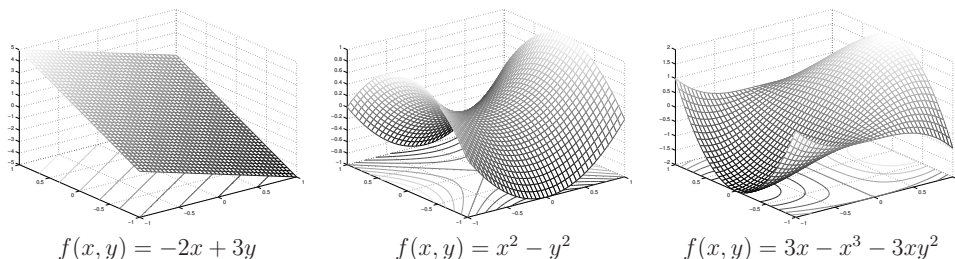
V předchozích kapitolách jsme potkali lineární a afinní zobrazení a kvadratické funkce. V této kapitole si řekneme více o nelineárních funkcích $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a zobrazeních $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ (zopakujte si značení funkcí a zobrazení z §1.1.3). Předpokládáme přitom, že student zná analýzu funkce jedné proměnné a pojem parciální derivace.

Dále budeme předpokládat, že definiční obor funkcí a zobrazení je celé \mathbb{R}^n . Tomu tak není vždy, např. definiční obor funkce $f(x) = \sqrt{1-x^2}$ je interval $[-1, 1] \subset \mathbb{R}$. Tento předpoklad ale zjednoduší výklad a pro čtenáře vždy bude snadné látku zobecnit pro jiný definiční obor.

Pro funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ užíváme tyto pojmy:

- **Graf** funkce f je množina $\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid y = f(\mathbf{x})\}$.
- **Vrstevnice** funkce f výšky y je množina $\{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) = y\}$.

Obrázek ukazuje příklady grafu a vrstevnic funkcí dvou proměnných na obdélníku $[-1, 1]^2$:



Příklad 8.1. Příklady funkcí a zobrazení více proměnných:

1. $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^2 - y^2$
2. $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(\mathbf{x}) = x_1$ (i když x_2, \dots, x_n chybí, f se rozumí jako funkce n proměnných)
3. $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} = a_1 x_1 + \dots + a_n x_n$ (lineární funkce)
4. $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b = a_1 x_1 + \dots + a_n x_n + b$ (afinní funkce)
5. $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(\mathbf{x}) = e^{-\mathbf{x}^T \mathbf{x}}$ (nenormalizované Gaussovo rozdělení)
6. $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(\mathbf{x}) = \max_{i=1}^n x_i$
7. $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^2$, $\mathbf{f}(t) = (\cos t, \sin t)$ (parametrizace kružnice, množina $\mathbf{f}([0, 2\pi])$ je kružnice)
8. $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^3$, $\mathbf{f}(t) = (\cos t, \sin t, at)$ (parametrizace šroubovice neboli helixu)

9. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\mathbf{f}(\mathbf{x}) = \mathbf{x}$ (identické zobrazení neboli identita)
10. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ (lineární zobrazení)
11. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ (afinní zobrazení)
12. $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $\mathbf{f}(u, v) = ((R + r \cos v) \cos u, (R + r \cos v) \sin u, r \sin v)$ (parametrizace toru neboli anuloidu, množina $\mathbf{f}([0, 2\pi) \times [0, 2\pi))$ je torus)
13. Při technice *image morphing* se obrázek např. obličejů zdeformuje na obrázek jiného obličejů. Morphing je realizován zobrazením $\mathbb{R}^2 \rightarrow \mathbb{R}^2$.
14. Elektrické pole přiřadí každému bodu v \mathbb{R}^3 vektor z \mathbb{R}^3 . □

8.1 Spojitost

Neformálně řečeno, zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je **spojité** v bodě $\mathbf{x} \in \mathbb{R}^n$, jestliže bodům blízkým k \mathbf{x} přiřazuje body blízké k $\mathbf{f}(\mathbf{x})$. Abychom tuto větu formalizovali, potřebovali bychom limitu funkce více proměnných, jejíž znalost u čtenáře nepředpokládáme. Uvedeme proto pouze postačující (avšak nikoliv nutnou) podmínku pro spojitost, která nám v praxi stačí. Přitom předpokládáme, že čtenář dokáže ověřit spojitost funkcí jedné proměnné.

Věta 8.1.

1. Necht' funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ je spojitá v bodě x . Necht' $k \in \{1, \dots, n\}$ a necht' funkce $g: \mathbb{R}^n \rightarrow \mathbb{R}$ je dána jako $g(x_1, \dots, x_n) = f(x_k)$ (tedy g závisí jen na proměnné x_k). Pak funkce g je spojitá v každém bodě (x_1, \dots, x_n) takovém, že $x_k = x$.
2. Necht' funkce $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$ jsou spojité v bodě \mathbf{x} . Pak funkce $f + g$, $f - g$ a f/g jsou spojité v bodě \mathbf{x} . Pokud $g(\mathbf{x}) \neq 0$, je funkce f/g spojitá v bodě \mathbf{x} .
3. Necht' $g: \mathbb{R}^n \rightarrow \mathbb{R}$ je spojitá v bodě \mathbf{x} a $f: \mathbb{R} \rightarrow \mathbb{R}$ je spojitá v bodě $y = g(\mathbf{x})$. Pak složená funkce $f \circ g: \mathbb{R}^n \rightarrow \mathbb{R}$ je spojitá v bodě \mathbf{x} .
4. Necht' funkce $f_1, \dots, f_m: \mathbb{R}^n \rightarrow \mathbb{R}$ jsou spojité v bodě \mathbf{x} . Pak zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ definované jako $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ je spojité v bodě \mathbf{x} .

Příklad 8.2. Z věty snadno ukážeme, že např. funkce dvou proměnných

$$f(x, y) = \sin(x + y^2) \tag{8.1}$$

je spojitá. Podle 1 je x spojitá funkce dvou proměnných (x, y) . Podobně, y^2 je spojitá funkce dvou proměnných (x, y) . Podle 2 je proto funkce $x + y^2$ spojitá. Protože funkce \sin je spojitá, podle 3 je funkce $\sin(x + y^2)$ spojitá. Takto jsme 'rekurzivně' dokázali spojitost celé funkce. □

8.2 Parciální derivace

Parciální derivaci funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ podle x_i značíme jedním ze symbolů

$$\frac{\partial f(\mathbf{x})}{\partial x_i} = f_{x_i}(\mathbf{x}) = \frac{\partial y}{\partial x_i},$$

kde poslední značení předpokládá, že jsme psali $y = f(\mathbf{x})$. Spočítáme ji tak, že všechny proměnné x_j , $j \neq i$, považujeme za konstanty a zderivujeme funkci podle jediné proměnné x_i .

Příklad 8.3. Parciální derivace funkce (8.1) jsou

$$\frac{\partial f(x, y)}{\partial x} = f_x(x, y) = \cos(x + y^2), \quad (8.2a)$$

$$\frac{\partial f(x, y)}{\partial y} = f_y(x, y) = 2y \cos(x + y^2). \quad (8.2b)$$

□

8.3 Totální derivace

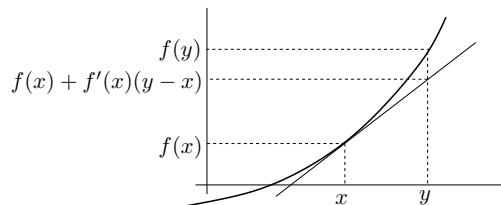
Zopakujme definici derivace funkce jedné proměnné $f: \mathbb{R} \rightarrow \mathbb{R}$ v bodě x . Existuje-li limita

$$\frac{df(x)}{dx} = f'(x) = \lim_{y \rightarrow x} \frac{f(y) - f(x)}{y - x}, \quad (8.3)$$

funkce se nazývá *diferencovatelná* v bodě x a hodnota limity se nazývá její *derivace* funkce f v bodě x . Pokud je funkce v bodě x diferencovatelná, lze ji v blízkosti bodu x ‘dobře’ aproximovat afinní funkcí

$$f(y) \approx f(x) + f'(x)(y - x). \quad (8.4)$$

Viz obrázek:



Jak se dá pojem diferencovatelnosti a derivace zobecnit na zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$? Toho nelze dosáhnout zobecněním limity (8.3), ale je lépe vycházet ze vzorce (8.4). Zkusme zobrazení v blízkosti bodu \mathbf{x} aproximovat jako

$$\mathbf{f}(\mathbf{y}) \approx \mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad (8.5)$$

kde symbol $\mathbf{f}'(\mathbf{x})$ označuje matici rozměru $m \times n$, o které ale zatím nevíme nic. Je-li \mathbf{x} pevné, pravá strana výrazu (8.5) je afinní zobrazení v proměnné \mathbf{y} (porovnejte s (3.5): pravou stranu (8.5) lze psát jako $\mathbf{A}\mathbf{y} + \mathbf{b}$, kde $\mathbf{A} = \mathbf{f}'(\mathbf{x})$ a $\mathbf{b} = \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{x}$). Zobrazení je *diferencovatelné* v bodě \mathbf{x} , jestliže je v okolí tohoto bodu ‘podobné’ afinnímu zobrazení, neboli existuje matice $\mathbf{f}'(\mathbf{x})$ taková, že chyba aproximace $\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x})$ je ‘malá’ pro ‘malé’ $\mathbf{y} - \mathbf{x}$. Abychom tuto podmínku formulovali přesně, potřebovali bychom limitu funkce více proměnných, jejíž znalost u čtenáře nepředpokládáme. Ponecháme proto pojem ‘diferencovatelné zobrazení’ nedefinovaný a definujeme pouze o něco silnější vlastnost, která nám v tomto kurzu postačí.

Definice 8.1. Zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je v bodě \mathbf{x} *spojitě diferencovatelné*, jestliže v bodě \mathbf{x} existují všechny parciální derivace $\partial f_i(\mathbf{x})/\partial x_j$ a jsou v tomto bodě *spojité*.

Věta 8.2. Je-li zobrazení v bodě *spojitě diferencovatelné*, je v tomto bodě *diferencovatelné*.

Příklad 8.4. Obě parciální derivace (8.2) funkce (8.1) jsou spojité funkce na celém \mathbb{R}^2 (neboť splňují předpoklady Věty 8.1). Proto je funkce (8.1) *spojitě diferencovatelná* (a tedy diferencovatelná) na celém \mathbb{R}^2 . □

Zdůrazněme, že pouhá existence všech parciálních derivací pro diferencovatelnost nestačí.

Příklad 8.5. Necht' je funkce $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definována jako

$$f(x, y) = \begin{cases} 1 & \text{když } x = 0 \text{ nebo } y = 0, \\ 0 & \text{když } x \neq 0 \text{ a } y \neq 0. \end{cases}$$

Je

$$\frac{\partial f(x, y)}{\partial x} = \begin{cases} 0 & \text{když } x \neq 0 \text{ nebo } y = 0, \\ \text{neexistuje} & \text{když } x = 0 \text{ a } y \neq 0 \end{cases}$$

a podobně pro $\partial f(x, y)/\partial y$. V bodě (0, 0) tedy existují obě parciální derivace (obě jsou rovny nule). Lze ukázat, že v bodě (0, 0) funkce není diferencovatelná. To nepřekvapí, neboť funkce se v okolí tohoto bodu afinní funkcí vůbec nepodobá. □

V případě, že je zobrazení \mathbf{f} je v bodě \mathbf{x} diferencovatelné, má matice $\mathbf{f}'(\mathbf{x})$ přirozený tvar: její složky jsou parciální derivace všech složek zobrazení podle všech proměnných:

$$\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}} = \mathbf{f}'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{bmatrix}. \quad (8.6)$$

Matice (8.6) se nazývá **totální derivace**¹ (nebo krátce jen **derivace**) zobrazení \mathbf{f} v bodě \mathbf{x} . Z historických důvodů se jí říká také **Jacobiho matice**. Speciální případy:

- Pro $f: \mathbb{R} \rightarrow \mathbb{R}$ je $f'(x)$ *skalár* a splývá s obyčejnou derivací (8.3).
- Pro $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^m$ je $\mathbf{f}'(x)$ *sloupcový vektor*, jehož složky jsou obyčejné derivace složek \mathbf{f} .
- Pro $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je $f'(\mathbf{x})$ *řádkový vektor*.

8.3.1 Derivace složeného zobrazení

Známe ‘řetězové pravidlo’ pro derivaci složených funkcí jedné proměnné lze přirozeným způsobem rozšířit na zobrazení. Důkaz následující věty není krátký a nebudeme ho uvádět.

Věta 8.3. Necht' $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ a $\mathbf{g}: \mathbb{R}^m \rightarrow \mathbb{R}^l$ jsou *diferencovatelná zobrazení*. Derivace složeného zobrazení $\mathbf{g} \circ \mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^l$ je

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}) = \frac{d\mathbf{g}(\mathbf{f}(\mathbf{x}))}{d\mathbf{x}} = \mathbf{g}'(\mathbf{f}(\mathbf{x})) \mathbf{f}'(\mathbf{x}). \quad (8.7)$$

Dimenze zúčastněných prostorů lze přehledně znázornit diagramem (viz §1.1.2)

$$\mathbb{R}^n \xrightarrow{\mathbf{f}} \mathbb{R}^m \xrightarrow{\mathbf{g}} \mathbb{R}^l.$$

¹Někdy se místo pojmu ‘totální derivace’ používá pojem ‘totální diferenciál’. Tyto pojmy jsou si podobné ale ne identické: totální derivace je *matice* a totální diferenciál je *lineární zobrazení* reprezentované touto maticí.

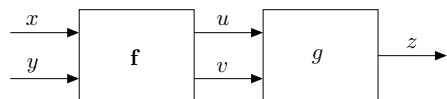
Pokud píšeme $\mathbf{u} = \mathbf{f}(\mathbf{x})$ a $\mathbf{y} = \mathbf{g}(\mathbf{u})$, pravidlo lze psát také v Leibnizově značení jako

$$\frac{d\mathbf{y}}{d\mathbf{x}} = \frac{d\mathbf{y}}{d\mathbf{u}} \frac{d\mathbf{u}}{d\mathbf{x}}, \quad (8.8)$$

což se dobře pamatuje, protože $d\mathbf{u}$ se 'jakoby vykrátí' (což ale není důkaz!). Zdůrazněme, že tato rovnost je násobením matic. Výraz na levé straně je matice $l \times n$, první výraz na pravé straně je matice $l \times m$ a druhý výraz na pravé straně je matice $m \times n$. Pro $l = m = n = 1$ dostaneme řetězové pravidlo pro derivaci složení funkcí jedné proměnné. Pravidlo se dá zjevným způsobem rozšířit na složení více než dvou zobrazení: *Jacobiho matice složeného zobrazení je součinem Jacobiho matic jednotlivých zobrazení.*

Příklad 8.6. Necht' $g(u, v)$ je diferencovatelná funkce dvou proměnných. Určeme derivaci funkce $z = h(x, y) = g(x + y, xy)$ podle vektoru (x, y) , tedy její parciální derivace podle x a y .

Máme $\mathbb{R}^2 \xrightarrow{\mathbf{f}} \mathbb{R}^2 \xrightarrow{\mathbf{g}} \mathbb{R}$, kde $\mathbf{f}(x, y) = (u, v) = (x + y, xy)$. Viz obrázek:



Derivace zobrazení g podle vektoru (u, v) je matice 1×2 (řádkový vektor)

$$g'(u, v) = \left[\frac{\partial g(u, v)}{\partial u} \quad \frac{\partial g(u, v)}{\partial v} \right] = [g_u(u, v) \quad g_v(u, v)].$$

Derivace zobrazení \mathbf{f} podle vektoru (x, y) je matice 2×2

$$\mathbf{f}'(x, y) = \frac{d\mathbf{f}(x, y)}{d(x, y)} = \begin{bmatrix} \frac{\partial(x+y)}{\partial x} & \frac{\partial(x+y)}{\partial y} \\ \frac{\partial(xy)}{\partial x} & \frac{\partial(xy)}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ y & x \end{bmatrix}.$$

Derivace zobrazení $h = g \circ \mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}$ podle vektoru (x, y) je matice 1×2 (řádkový vektor)

$$\begin{aligned} \frac{dh(x, y)}{d(x, y)} &= \frac{dg(\mathbf{f}(x, y))}{d(x, y)} = g'(u, v)\mathbf{f}'(x, y) \\ &= [g_u(u, v) \quad g_v(u, v)] \begin{bmatrix} 1 & 1 \\ y & x \end{bmatrix} \\ &= [g_u(u, v) + yg_v(u, v) \quad g_u(u, v) + xg_v(u, v)], \end{aligned}$$

kde $u = x + y$ a $v = xy$. □

Příklad 8.7. Ukažme dva způsoby, jak najít parciální derivaci f_x funkce $f(x, y) = e^{(x+y)^2+(xy)^2}$:

- Považujeme y za konstantu a derivujeme f jako funkci jedné proměnné x :

$$f_x = [2(x + y) + 2(xy)y]e^{(x+y)^2+(xy)^2} = 2(x + y + xy^2)e^{(x+y)^2+(xy)^2}.$$

- Položíme $u = x + y$, $v = xy$, $f(u, v) = e^{u^2+v^2}$. Z Příkladu 8.6 máme $f_x = f_u + yf_v$. Jelikož

$$f_u = 2ue^{u^2+v^2}, \quad f_v = 2ve^{u^2+v^2},$$

$$\text{máme } f_x = f_u + yf_v = 2ue^{u^2+v^2} + y(2v)e^{u^2+v^2} = 2(x + y + xy^2)e^{(x+y)^2+(xy)^2}. \quad \square$$

Příklad 8.8. Máme diferencovatelnou funkci $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ a chceme spočítat derivaci funkce $g(t + t^2, \sin t)$ podle t .

Máme $\mathbb{R} \xrightarrow{\mathbf{f}} \mathbb{R}^2 \xrightarrow{\mathbf{g}} \mathbb{R}$, kde $\mathbf{f}(t) = (u, v) = (t + t^2, \sin t)$. Je

$$\frac{dg(t + t^2, \sin t)}{dt} = g'(u, v)\mathbf{f}'(t) = [g_u(u, v) \quad g_v(u, v)] \begin{bmatrix} 1 + 2t \\ \cos t \end{bmatrix} = g_u(u, v)(1 + 2t) + g_v(u, v) \cos t. \quad \square$$

8.3.2 Derivace maticových výrazů

Jsou-li funkce nebo zobrazení zadány výrazem obsahujícím vektory a matice, derivaci lze vždy spočítat 'hrubou silou' tak, že výraz rozepíšeme do složek a spočítáme parciální derivace každé složky podle každé proměnné. Přísně vzato, úkol jsme tím splnili. Je ovšem výhodné tento výsledek zjednodušit tak, že pro něj najdeme maticový výraz.

Příklad 8.9. Odvod'me derivaci zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ daného vzorcem $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, kde $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{b} \in \mathbb{R}^m$. Máme

$$\begin{aligned} f_1(\mathbf{x}) &= a_{11}x_1 + \dots + a_{1n}x_n + b_1 \\ &\vdots \\ f_m(\mathbf{x}) &= a_{m1}x_1 + \dots + a_{mn}x_n + b_m. \end{aligned}$$

Ale $\partial f_i(\mathbf{x})/\partial x_j = a_{ij}$, tedy dle (8.6) je $\mathbf{f}'(\mathbf{x}) = \mathbf{A}$. Tomu se nedivíme, neboť zobrazení \mathbf{f} je afinní, tedy jeho afinní aproximace (8.5) musí být to samé afinní zobrazení. Opravdu: pravá strana výrazu (8.5) je rovna $\mathbf{A}\mathbf{y} + \mathbf{b}$ (zkontrolujte!). □

Příklad 8.10. Počítejme derivaci zobrazení $\mathbf{g}(\mathbf{A}\mathbf{x} + \mathbf{b})$ podle \mathbf{x} , kde $\mathbf{g}: \mathbb{R}^l \rightarrow \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{l \times n}$, $\mathbf{b} \in \mathbb{R}^l$. V řetězovém pravidle máme $\mathbb{R}^n \xrightarrow{\mathbf{f}} \mathbb{R}^l \xrightarrow{\mathbf{g}} \mathbb{R}^m$, kde $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$. Je $\mathbf{f}'(\mathbf{x}) = \mathbf{A}$, tedy

$$\frac{d\mathbf{g}(\mathbf{A}\mathbf{x} + \mathbf{b})}{d\mathbf{x}} = \mathbf{g}'(\mathbf{A}\mathbf{x} + \mathbf{b})\mathbf{A}. \quad \square$$

Příklad 8.11. Odvod'me derivaci kvadratické formy $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A}\mathbf{x}$, kde \mathbf{A} je libovolná (ne nutně symetrická) matice velikosti $n \times n$. Napíšeme si funkci f podrobně:

$$\begin{aligned} \mathbf{x}^T \mathbf{A}\mathbf{x} &= a_{11}x_1^2 + a_{21}x_2x_1 + \dots + a_{n1}x_nx_1 + \\ &\quad a_{12}x_1x_2 + a_{22}x_2^2 + \dots + a_{n2}x_nx_2 + \\ &\quad \vdots \\ &\quad a_{1n}x_1x_n + a_{2n}x_2x_n + \dots + a_{nn}x_n^2. \end{aligned}$$

Z tohoto výrazu po troše snahy vidíme, že

$$\frac{\partial f(\mathbf{x})}{\partial x_1} = 2a_{11}x_1 + (a_{21} + a_{12})x_2 + \dots + (a_{n1} + a_{1n})x_n$$

a podobně pro derivace podle ostatních proměnných. Ale tyto parciální derivace lze uspořádat do řádkového vektoru

$$\mathbf{f}'(\mathbf{x}) = \left[\frac{\partial f(\mathbf{x})}{\partial x_1} \quad \dots \quad \frac{\partial f(\mathbf{x})}{\partial x_n} \right] = \mathbf{x}^T (\mathbf{A} + \mathbf{A}^T). \quad \square$$

Následující tabulka uvádí derivace často potkávaných zobrazení. Odvoďte je jako cvičení!

| $\mathbf{f}(\mathbf{x})$ | $\mathbf{f}'(\mathbf{x})$ | poznámka |
|---|---|---|
| \mathbf{x} | \mathbf{I} | $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ |
| $\mathbf{A}\mathbf{x}$ | \mathbf{A} | $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ |
| $\mathbf{x}^T \mathbf{x}$ | $2\mathbf{x}^T$ | $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$ |
| $\mathbf{x}^T \mathbf{A}\mathbf{x}$ | $\mathbf{x}^T(\mathbf{A} + \mathbf{A}^T)$ | $\mathbf{A} \in \mathbb{R}^{n \times n}, \mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$ |
| $\mathbf{x}^T \mathbf{a} = \mathbf{a}^T \mathbf{x}$ | \mathbf{a}^T | $\mathbf{a} \in \mathbb{R}^n, \mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$ |
| $\ \mathbf{x}\ $ | $\mathbf{x}^T / \ \mathbf{x}\ $ | $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$ |
| $\mathbf{g}(\mathbf{x})^T \mathbf{g}(\mathbf{x})$ | $2\mathbf{g}(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$ | $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$ |
| $\mathbf{g}(\mathbf{x})^T \mathbf{h}(\mathbf{x})$ | $\mathbf{g}(\mathbf{x})^T \mathbf{h}'(\mathbf{x}) + \mathbf{h}(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$ | $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{h}: \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$ |

8.4 Směrová derivace

Řez zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ v bodě $\mathbf{x} \in \mathbb{R}^n$ ve směru $\mathbf{v} \in \mathbb{R}^n$ je zobrazení $\varphi: \mathbb{R} \rightarrow \mathbb{R}^m$ dané jako

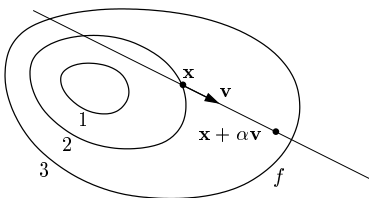
$$\varphi(\alpha) = \mathbf{f}(\mathbf{x} + \alpha \mathbf{v}). \quad (8.9)$$

Směrová derivace³ zobrazení \mathbf{f} v bodě \mathbf{x} ve směru \mathbf{v} je číslo

$$(\varphi'_1(0), \dots, \varphi'_m(0)) = \lim_{\alpha \rightarrow 0} \frac{\varphi(\alpha) - \varphi(0)}{\alpha} = \lim_{\alpha \rightarrow 0} \frac{\mathbf{f}(\mathbf{x} + \alpha \mathbf{v}) - \mathbf{f}(\mathbf{x})}{\alpha}, \quad (8.10)$$

kde $\varphi'_i(0)$ označuje derivaci i -té složky zobrazení φ v bodě $\alpha = 0$.

Pojem směrové derivace se geometricky snadněji představí pro případ $m = 1$, tedy pro funkci $f: \mathbb{R}^2 \rightarrow \mathbb{R}$. Obrázek ilustruje situaci pro funkci $f: \mathbb{R}^2 \rightarrow \mathbb{R}$:



Parciální derivace funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ není nic jiného než její směrová derivace ve směru i -tého vektoru standardní báze $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$ (jednička na i -tém místě).

Věta 8.4. Necht' zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je diferencovatelné v bodě \mathbf{x} . Pak jeho směrová derivace v bodě \mathbf{x} ve směru \mathbf{v} je rovna $\mathbf{f}'(\mathbf{x})\mathbf{v}$.

Důkaz. Zobrazení $\mathbf{y} = \varphi(\alpha) = \mathbf{f}(\mathbf{x} + \alpha \mathbf{v})$ je složením dvou zobrazení $\mathbf{y} = \mathbf{f}(\mathbf{u})$ a $\mathbf{u} = \mathbf{x} + \alpha \mathbf{v}$. Máme diagram $\mathbb{R} \xrightarrow{\mathbf{u}=\mathbf{x}+\alpha\mathbf{v}} \mathbb{R}^n \xrightarrow{\mathbf{y}=\mathbf{f}(\mathbf{x})} \mathbb{R}^m$. Máme $d\mathbf{u}/d\alpha = \mathbf{v}$. Podle řetězového pravidla je

$$\varphi'(\alpha) = \frac{d\mathbf{y}}{d\alpha} = \frac{d\mathbf{y}}{d\mathbf{u}} \frac{d\mathbf{u}}{d\alpha} = \frac{d\mathbf{f}(\mathbf{u})}{d\mathbf{u}} \mathbf{v}.$$

Pro $\alpha = 0$ je $\mathbf{u} = \mathbf{x}$, čímž je věta dokázána. \square

²Někdy se řez a směrová derivace uvažují jen pro *normalizované* směry \mathbf{v} , tj. $\|\mathbf{v}\| = 1$. My ale dovolujeme libovolný vektor \mathbf{v} .

³Přesněji jde o *oboustrannou* směrovou derivaci. Jednostrannou směrovou derivaci bychom dostali, kdybychom místo oboustranné limity (8.10) vzali jednostrannou limitu zprava.

Věta 8.4 říká, že je-li zobrazení \mathbf{f} diferencovatelné, je jeho směrová derivace (v pevném bodě \mathbf{x}) lineární zobrazení směru \mathbf{v} reprezentované Jacobiho maticí $\mathbf{f}'(\mathbf{x})$. Zdůrazněme, že nic takového neplatí, pokud zobrazení \mathbf{f} není diferencovatelné.

Příklad 8.12. Spočítejme směrovou derivaci funkce $f(x, y) = \sin(x + y^2)$ v bodě (x, y) ve směru (u, v) . Podle definice je

$$\begin{aligned} \varphi(\alpha) &= \sin(x + \alpha u + (y + \alpha v)^2), \\ \varphi'(\alpha) &= (u + 2v(y + \alpha v)) \cos(x + \alpha u + (y + \alpha v)^2), \\ \varphi'(0) &= (u + 2vy) \cos(x + y^2). \end{aligned} \quad (8.11)$$

Podle Věty 8.4 je směrová derivace rovna

$$uf_x(x, y) + vf_y(x, y) = u \cos(x + y^2) + 2vy \cos(x + y^2),$$

což je stejné jako (8.11). \square

8.5 Gradient

Transpozici totální derivace funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ se říká **gradient** a značí se

$$f'(\mathbf{x})^T = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix} = \nabla f(\mathbf{x})$$

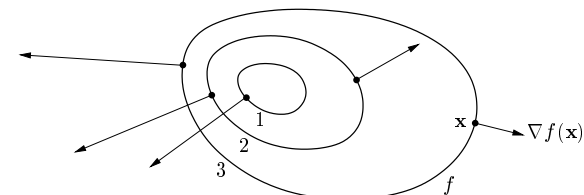
(∇ čteme 'nabla'). Jelikož $f'(\mathbf{x})$ je řádkový vektor, je gradient sloupcový vektor⁴.

Zkoumejme směrovou derivaci v pevném bodě \mathbf{x} pro různé normalizované směry \mathbf{v} (tedy $\|\mathbf{v}\| = 1$). Tato derivace je rovna $f'(\mathbf{x})\mathbf{v} = \nabla f(\mathbf{x})^T \mathbf{v}$, což je skalární součin gradientu v bodě \mathbf{x} a vektoru \mathbf{v} . Je jasné (ale promyslete!), že:

- Směrová derivace je největší ve směru $\mathbf{v} = \nabla f(\mathbf{x}) / \|\nabla f(\mathbf{x})\|$, tedy když je \mathbf{v} rovnoběžný s gradientem a stejně orientovaný. Tedy směr gradientu je *směr největšího růstu* funkce.
- Velikost gradientu $\|\nabla f(\mathbf{x})\|$ je strmost funkce ve směru největšího růstu.
- Směrová derivace ve směru kolmém na gradient je nulová.

Dále lze ukázat (viz diskuze v §9.4.1), že gradient je vždy *kolmý k vrstevnici*.

Obrázek ukazuje tři vrstevnice funkce $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ a její gradienty v několika bodech:



⁴Zavedení nového slova pro transpozici derivace se zdá zbytečné – nicméně důvod je v tom, že totální diferenciál je *lineární funkce*, kdežto gradient je *vektor*. Literatura bohužel není jednotná v rozlišování gradientu a (totální) derivace funkce. Někdy se obojí ztotožňuje a značí $\nabla f(\mathbf{x})$. To ale vede k nekonzistenci se značením používaným v lineární algebře, protože derivace funkce $\mathbb{R}^n \rightarrow \mathbb{R}$ pak není speciálním případem derivace zobrazení $\mathbb{R}^n \rightarrow \mathbb{R}^m$ (tedy Jacobiho matice) pro $m = 1$, což je řádkový vektor.

8.6 Parciální derivace druhého řádu

Zderivujeme-li funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ nejdříve podle proměnné x_i a pak podle proměnné x_j , dostaneme parciální derivaci druhého řádu, kterou značíme

$$\frac{\partial}{\partial x_j} \frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}.$$

Je-li $i = j$, píšeme zkráceně

$$\frac{\partial}{\partial x_i} \frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{\partial^2 f(\mathbf{x})}{\partial x_i^2}.$$

Věta 8.5. Pokud jsou druhé parciální derivace

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}, \quad \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i}$$

spojité v bodě \mathbf{x} , pak jsou si rovny.

Příklad 8.13. Určeme všechny druhé parciální derivace funkce $f(x, y) = \sin(x + y^2)$ z Příkladu 8.2. První derivace už tam máme. Nyní druhé derivace:

$$\begin{aligned} \frac{\partial^2 f(x, y)}{\partial x^2} &= \frac{\partial}{\partial x}(\cos(x + y^2)) = -\sin(x + y^2), \\ \frac{\partial^2 f(x, y)}{\partial x \partial y} &= \frac{\partial}{\partial y}(\cos(x + y^2)) = -2y \sin(x + y^2), \\ \frac{\partial^2 f(x, y)}{\partial y \partial x} &= \frac{\partial}{\partial x}(2y \cos(x + y^2)) = -2y \sin(x + y^2), \\ \frac{\partial^2 f(x, y)}{\partial y^2} &= \frac{\partial}{\partial y}(2y \cos(x + y^2)) = 2 \cos(x + y^2) - 4y^2 \sin(x + y^2). \end{aligned}$$

Vidíme, že vskutku nezáleží na pořadí derivování podle x a podle y . □

Pro funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ budeme značit matici všech druhých parciálních derivací

$$f''(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_n} \end{bmatrix}.$$

Je to symetrická matice velikosti $n \times n$, která se často nazývá **Hessova matice**.

Co by byla druhá derivace zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$? Nebyla by to již dvojrozměrná tabulka (tedy matice) velikosti $n \times n$, nýbrž třírozměrná tabulka velikosti $m \times n \times n$.

8.7 Taylorův polynom

Nechť funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ má v bodě x derivace až do řádu k . Její **Taylorův polynom** stupně k v bodě x je polynom $T_k: \mathbb{R} \rightarrow \mathbb{R}$ takový, že v bodě x má všechny derivace až do řádu k stejné jako funkce f . V tomto smyslu je polynom T_k aproximační funkce f v okolí bodu x .

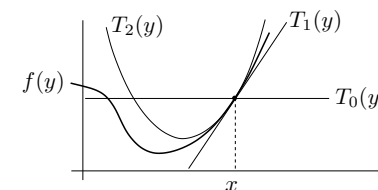
Taylorův polynom je tímto požadavkem definován jednoznačně a má tvar (odvoďte!)

$$T_k(y) = \sum_{i=0}^k \frac{1}{i!} f^{(i)}(x) (y - x)^i, \quad (8.12)$$

kde $f^{(i)}$ označuje i -tou derivaci funkce f (kde nultá derivace je funkce sama, $f^{(0)} = f$) a kde klademe $0! = 1$. Tvary polynomu až do stupně 2:

$$\begin{aligned} T_0(y) &= f(x), \\ T_1(y) &= f(x) + f'(x)(y - x), \\ T_2(y) &= f(x) + f'(x)(y - x) + \frac{1}{2} f''(x)(y - x)^2. \end{aligned}$$

Taylorův polynom nultého stupně T_0 je hodně špatná aproximace, rovná jednoduše konstantní funkci. Polynom prvního stupně $T_1(x)$ už známe ze vzorce (8.4). Polynom druhého stupně T_2 je parabola, která má s funkcí f v bodě x společnou hodnotu a první dvě derivace. Viz obrázek:



Jak zobecnit Taylorův polynom pro funkci více proměnných $f: \mathbb{R}^n \rightarrow \mathbb{R}$? Definujme Taylorův polynom k -tého stupně (funkce f v okolí bodu \mathbf{x}) jako polynom $T_k: \mathbb{R}^n \rightarrow \mathbb{R}$, který má s funkcí f v bodě \mathbf{x} společně všechny parciální derivace až do řádu k . Nebudeme uvádět vzorec pro polynom libovolného stupně, napíšeme jen polynomy do stupně dva:

$$T_0(\mathbf{y}) = f(\mathbf{x}), \quad (8.13a)$$

$$T_1(\mathbf{y}) = f(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad (8.13b)$$

$$T_2(\mathbf{y}) = f(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T \mathbf{f}''(\mathbf{x})(\mathbf{y} - \mathbf{x}). \quad (8.13c)$$

Zde $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{f}'(\mathbf{x}) \in \mathbb{R}^{1 \times n}$ je Jacobiho matice (řádkový vektor) a $\mathbf{f}''(\mathbf{x}) \in \mathbb{R}^{n \times n}$ je Hessova matice. Funkce (8.13b) je afinní a funkce (8.13c) je kvadratická.

Taylorův polynom lze zobecnit na zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ tak, že vezmeme Taylorovy polynomy všech složek f_1, \dots, f_m . Polynom prvního stupně tak vede na zobrazení

$$\mathbf{T}_1(\mathbf{y}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad (8.14)$$

což není nic jiného než pravá strana (8.5). Polynom druhého stupně vede na zobrazení \mathbf{T}_2 , jehož složky jsou funkce (8.13c). To nejde napsat v maticové formě, protože všech $m \times n \times n$ druhých parciálních derivací se 'nevejde' do matice.

Příklad 8.14. Najděte Taylorův polynom druhého stupně funkce $f(x, y) = x^{-1} + y^{-1} + xy$ v bodě $(x_0, y_0) = (2, 1)$. Máme

$$\begin{aligned} f(x_0, y_0) &= x^{-1} + y^{-1} + xy \Big|_{(x,y)=(2,1)} = \frac{7}{2}, \\ f'(x_0, y_0) &= [y - x^{-2} \quad x - y^{-2}] \Big|_{(x,y)=(2,1)} = \left[\frac{3}{4} \quad 1 \right], \\ f''(x_0, y_0) &= \begin{bmatrix} 2x^{-3} & 1 \\ 1 & 2y^{-3} \end{bmatrix} \Big|_{(x,y)=(2,1)} = \begin{bmatrix} \frac{1}{4} & 1 \\ 1 & 2 \end{bmatrix}. \end{aligned}$$

Dle (8.13c) je (pozor, naše proměnné jsou označené jinak než v (8.13c))

$$\begin{aligned} T_2(x, y) &= f(x_0, y_0) + f'(x_0, y_0) \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}^T f''(x_0, y_0) \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\ &= \frac{7}{2} + \begin{bmatrix} \frac{3}{4} & 1 \end{bmatrix} \begin{bmatrix} x - 2 \\ y - 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x - 2 \\ y - 1 \end{bmatrix}^T \begin{bmatrix} \frac{1}{4} & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x - 2 \\ y - 1 \end{bmatrix} \\ &= \frac{1}{8}x^2 + xy + y^2 - \frac{3}{4}x - 3y + \frac{9}{2}. \end{aligned} \quad \square$$

8.8 Cvičení

8.1. Máme množiny $X = [-1, 1] \times \{0\} = \{(x, 0) \mid -1 \leq x \leq 1\} \subseteq \mathbb{R}^2$ a $Y = [-1, 1] \times [-1, 1]$. Načrtněte následující množiny:

- $\{\mathbf{x} \in \mathbb{R}^2 \mid 1 \geq \min_{\mathbf{y} \in X} \|\mathbf{x} - \mathbf{y}\|\}$
- $\{\mathbf{x} \in \mathbb{R}^2 \mid 2 \geq \max_{\mathbf{y} \in X} \|\mathbf{x} - \mathbf{y}\|\}$
- vrstevnice výšky 1 funkce $f(\mathbf{x}) = \min_{\mathbf{y} \in Y} \|\mathbf{x} - \mathbf{y}\|$
- vrstevnice výšky $\sqrt{2}$ funkce $f(\mathbf{x}) = \max_{\mathbf{y} \in Y} \|\mathbf{x} - \mathbf{y}\|$

8.2. Je dána funkce dvou proměnných $f(x, y)$.

- Spočítejte derivace f podle polárních souřadnic (φ, r) , kde $x = r \cos \varphi$, $y = r \sin \varphi$.
- Bod (x, y) se v čase t pohybuje po křivce dané rovnicí $(x, y) = (t^2 + 2t, \log(t^2 + 1))$. Najděte derivaci f podle času.

8.3. Spočítejte derivaci funkce $g(\mathbf{u}) = f(\mathbf{a}^T \mathbf{u}, \|\mathbf{u}\|)$ podle vektoru \mathbf{u} .

8.4. Nadmořská výška krajiny je dána vzorcem $h(d, s) = 2s^2 + 3sd - d^2 + 5$, kde d je zeměpisná délka (zvětšuje se od západu k východu) a s je zeměpisná šířka (zvětšuje se od jihu k severu). V bodě $(d, s) = (-1, 1)$ určete

- směr nejstrmějšího stoupání terénu
- strmost terénu v jihovýchodním směru.

V této úloze je logické uvažovat směr jako normalizovaný vektor.

8.5. Spočítejte druhou derivaci $f''(x, y)$ (tj. Hessovu matici) funkcí (výsledek co nejvíce zjednodušte)

- $f(x, y) = e^{-x^2 - y^2}$
- $f(x, y) = \log(e^x + e^y)$

8.6. Hessova matice kvadratické formy $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ je $f''(\mathbf{x}) = \mathbf{A} + \mathbf{A}^T$. Odvod'te.

8.7. Je dána funkce $f(x, y) = 6xy^2 - 2x^3 - 3y^3$. V bodě $(x_0, y_0) = (1, -2)$ najděte Taylorův polynom nultého, prvního a druhého stupně.

8.8. *Metoda konečných diferencí* počítá derivaci funkce přibližně jako

$$f'(x) \approx \frac{f(x+h) - f(x)}{h},$$

kde h je malé číslo (dobrá volba je $h = \sqrt{\varepsilon}$, kde ε je strojová přesnost). Toto jde použít i na parciální derivace. Vymyslete si dvě zobrazení $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ a $\mathbf{f}: \mathbb{R}^m \rightarrow \mathbb{R}^l$ pro nějaké navzájem různé dimenze $n, m, l > 1$. Zvolte bod $\mathbf{x} \in \mathbb{R}^n$. Spočítejte přibližně totální derivace (Jacobiho matice) $\mathbf{g}'(\mathbf{x})$ a $\mathbf{f}'(\mathbf{g}(\mathbf{x}))$ v Matlabu metodou konečných diferencí. Potom spočítejte derivaci složeného zobrazení $(\mathbf{f} \circ \mathbf{g})'(\mathbf{x})$ jednak metodou konečných diferencí a jednak vynásobením matic $\mathbf{g}'(\mathbf{x})$ a $\mathbf{f}'(\mathbf{g}(\mathbf{x}))$. Porovnejte.

Nápověda a řešení

$$8.2.a) \quad f_\varphi(x, y) = -f_x(x, y)r \sin \varphi + f_y(x, y)r \cos \varphi, \quad f_r(x, y) = f_x(x, y) \cos \varphi + f_y(x, y) \sin \varphi$$

$$8.2.b) \quad f_t(x, y) = 2f_x(x, y)(t+1) + 2tf_y(x, y)/(t^2+1)$$

$$8.4.a) \quad (5, 1)/\sqrt{26}$$

$$8.4.b) \quad (5, 1)^T(1, -1)/\sqrt{2} = 2\sqrt{2}$$

$$8.5.a) \quad 2e^{-x^2-y^2} \begin{bmatrix} 2x^2 - 1 & 2xy \\ 2xy & 2x^2 - 1 \end{bmatrix}$$

$$8.5.b) \quad \frac{e^{x+y}}{(e^x+e^y)^2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$8.7. \quad T_0(x, y) = 46, \quad T_1(x, y) = 18x - 60y - 92, \quad T_2(x, y) = -6x^2 - 24xy - 18x + 24y^2 + 60y + 46$$

Kapitola 9

Analytické podmínky na lokální extrémů

9.1 Vlastnosti bodu vzhledem k podmnožině \mathbb{R}^n

Množina $U \subseteq \mathbb{R}^n$ se nazývá **okolí** bodu $\mathbf{x} \in \mathbb{R}^n$, jestliže existuje $\varepsilon > 0$ tak, že¹

$$U = \{ \mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{y}\| < \varepsilon \}. \quad (9.1)$$

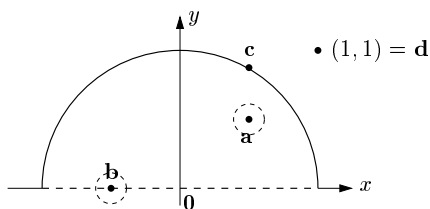
Množina U je koule (bez hranice) se středem \mathbf{x} a nenulovým poloměrem ε .

Definice 9.1. Mějme množinu $X \subseteq \mathbb{R}^n$. Bod $\mathbf{x} \in \mathbb{R}^n$ se nazývá její

- **vnitřní bod**, jestliže existuje okolí U bodu \mathbf{x} takové, že $U \subseteq X$
- **hraniční bod**, jestliže pro každé okolí U bodu \mathbf{x} je $U \cap X \neq \emptyset$ a $U \cap (\mathbb{R}^n \setminus X) \neq \emptyset$

Všimněte si, že hraniční bod množiny nemusí patřit do této množiny. Pokud leží bod v množině, je buď vnitřní nebo hraniční, ale ne obojí najednou (dokažte!). **Vnitřek** [hranice] množiny je množina všech jejích vnitřních [hraničních] bodů.

Příklad 9.1. Máme množinu $\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1, y > 0 \} \cup \{ (1, 1) \}$ na obrázku:



Bod **a** je vnitřní bod množiny, protože existuje jeho okolí U , které celé leží v množině. Bod **b** je hraniční, protože každé jeho okolí má neprázdný průnik s množinou i s jejím doplňkem. Všimněte si, že **b** nepatří do množiny. Bod **a** není hraniční a bod **b** není vnitřní. Bod **c** není vnitřní, je hraniční a patří do množiny. Bod **d** je hraniční. \square

Příklad 9.2. Bod $1/2$ je vnitřní bod intervalu $(0, 1] \subseteq \mathbb{R}$ a body 0 a 1 jsou hraniční. \square

Příklad 9.3. Množina $[0, 1] \times \{1\} = \{ (x, y) \mid 0 \leq x \leq 1, y = 1 \} \subseteq \mathbb{R}^2$ (úsečka v rovině) nemá žádné vnitřní body. Všechny její body jsou hraniční, je tedy sama svou vlastní hranicí. \square

¹ Norma v (9.1) může být eukleidovská, ale i libovolná jiná vektorová p -norma (viz §11.3.1). Vnitřek a hranice množiny na výběru normy nezávisí.

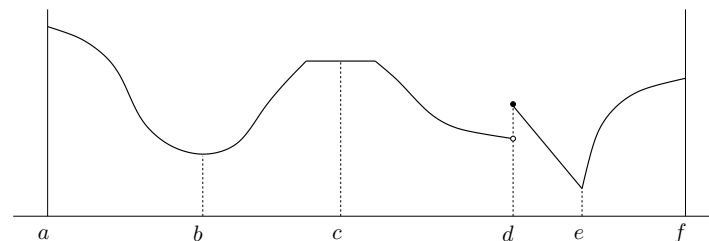
9.2 Lokální extrémů

Zopakujte si pojem extrému funkce na množině (§1.1.4): funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ nabývá na množině $X \subseteq \mathbb{R}^n$ svého minima v bodě $\mathbf{x} \in X$, jestliže $f(\mathbf{x}) \leq f(\mathbf{x}')$ pro všechna $\mathbf{x}' \in X$.

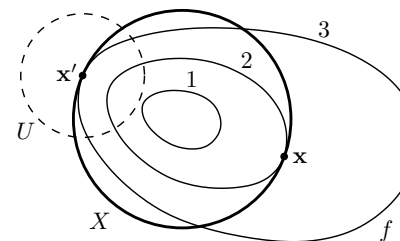
Definice 9.2. Funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ nabývá na množině $X \subseteq \mathbb{R}^n$ svého **lokálního minima** v bodě $\mathbf{x} \in X$, jestliže existuje okolí U bodu \mathbf{x} tak, že funkce f nabývá na množině $U \cap X$ svého minima v bodě \mathbf{x} .

Lokální maximum se definuje obdobně. Každé minimum funkce f na množině X je zároveň lokální minimum funkce f na množině X (naopak to ale obecně neplatí). Mluvíme-li o lokálních extrémech, pro zdůraznění někdy 'obyčejné' extrémy (ve smyslu §1.1.4) nazýváme **globální extrémy**. Pokud odkaz na množinu X chybí, myslí se jí celý definiční obor funkce f .

Příklad 9.4. Funkce jedné proměnné na obrázku nabývá na uzavřeném intervalu $[a, f] \subseteq \mathbb{R}$ v bodě a lokálního a zároveň globálního maxima, v bodě b lokálního minima, v bodě c lokálního maxima a zároveň lokálního minima, v bodě d lokálního maxima, v bodě e lokálního a zároveň globálního minima, v bodě f lokálního maxima. \square



Příklad 9.5. Necht' $X \subseteq \mathbb{R}^2$ je kružnice a funkce $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ má vrstevnice jako na obrázku:



V bodě \mathbf{x} nabývá funkce f na množině X globálního (a tedy i lokálního) minima, protože v žádném bodě na kružnici X nemá funkce menší hodnotu než $f(\mathbf{x}) = 2$. V bodě \mathbf{x}' nabývá funkce f na množině X lokálního minima, protože existuje okolí U bodu \mathbf{x}' takové, že funkce f nabývá na části kružnice $U \cap X$ svého (globálního) minima. \square

9.3 Volné lokální extrémů

Věta 9.1. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a $\mathbf{x} \in X \subseteq \mathbb{R}^n$. Necht'

- funkce f je v bodě \mathbf{x} diferencovatelná,
- \mathbf{x} je vnitřní bod množiny X ,
- \mathbf{x} je lokální extrém funkce f na množině X .

Pak $f'(\mathbf{x}) = \mathbf{0}$, neboli všechny parciální derivace funkce f v bodě \mathbf{x} jsou nulové.

Důkaz. Z Definice 9.2 plyne, že funkce f má v bodě \mathbf{x} (globální) extrém na nějakém okolí U bodu \mathbf{x} . Z toho ovšem plyne, že řez $\varphi(\alpha) = f(\mathbf{x} + \alpha\mathbf{v})$ funkce f (viz §8.4) v libovolném směru $\mathbf{v} \neq \mathbf{0}$ má (globální) extrém v bodě $\alpha = 0$ na množině $\{\alpha \in \mathbb{R} \mid \mathbf{x} + \alpha\mathbf{v} \in U\}$. Tedy funkce φ má v bodě $\alpha = 0$ lokální extrém. Tedy její derivace je v tomto bodě nulová (to víme z analýzy funkcí jedné proměnné). Ale tato derivace je směrová derivace funkce f v bodě \mathbf{x} ve směru \mathbf{v} . Parciální derivace jsou speciálním případem směrové derivace. \square

Bod, ve kterém má funkce všechny parciální derivace nulové, se nazývá její **stacionární bod**. Věta 9.1 svádí k tomu, aby se použila v situacích, kdy nejsou splněny její předpoklady. Uveďme příklady tohoto chybdého použití.

Příklad 9.6. V Příkladu 9.4 jsou předpoklady Věty 9.1 splněny pouze pro body b, c , které jsou stacionární a vnitřní. Body a, f jsou hraniční (tedy ne vnitřní) body intervalu $[a, f]$ a v bodech d, e není funkce diferencovatelná. \square

Příklad 9.7. Funkce $f(x) = x^3$ má na \mathbb{R} v bodě 0 stacionární bod, ale nemá tam lokální extrém. To není v rozporu s Větou 9.1. \square

Příklad 9.8. Funkce $f(\mathbf{x}) = \|\mathbf{x}\|$ má na hyperkrychli $X = \{\mathbf{x} \in \mathbb{R}^n \mid -1 \leq \mathbf{x} \leq 1\}$ v bodě $\mathbf{0}$ volné lokální minimum (nakreslete si množinu a vrstevnice funkce pro $n = 1$ a pro $n = 2$). Nemá tam ale stacionární bod, protože tam není diferencovatelná. Dále má funkce na množině X lokální maxima ve všech rozích hyperkrychle, např. v bodě $\mathbf{1}$. V bodě $\mathbf{1}$ ale není stacionární bod, což není v rozporu s Větou 9.1, protože $\mathbf{1}$ není vnitřní bod X . \square

Věta 9.1 říká, že stacionární body jsou body ‘podezřelé’ z volného lokálního extrému. Udává podmínku *prvního řádu* na volné extrémy, protože obsahuje první derivace. Následující podmínka *druhého řádu* pomůže zjistit, zda je stacionární bod lokálním extrémem, případně jakým.

Věta 9.2. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a $\mathbf{x} \in X \subseteq \mathbb{R}^n$. Necht'

- funkce f je v bodě \mathbf{x} dvakrát diferencovatelná,
- \mathbf{x} je vnitřní bod množiny X ,
- $f'(\mathbf{x}) = \mathbf{0}$.

Pak platí:

- Je-li \mathbf{x} lokální minimum [maximum] funkce f na množině X , pak Hessova matice $f''(\mathbf{x})$ je pozitivně [negativně] semidefinitní.
- Je-li $f''(\mathbf{x})$ pozitivně [negativně] definitní, pak \mathbf{x} je ostré lokální minimum [maximum] funkce f na množině X .
- Je-li $f''(\mathbf{x})$ indefinitní, pak \mathbf{x} není lokální extrém funkce f na množině X .

Větu 9.2 nebudeme dokazovat, uvedeme jen náznak možného důkazu. Místo funkce f vyšetřujeme v blízkosti bodu \mathbf{x} její Taylorův polynom druhého stupně (8.13c),

$$T_2(\mathbf{y}) = f(\mathbf{x}) + \underbrace{f'(\mathbf{x})(\mathbf{y} - \mathbf{x})}_0 + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T f''(\mathbf{x})(\mathbf{y} - \mathbf{x}).$$

Protože $f'(\mathbf{x}) = \mathbf{0}$, lineární člen je nulový a polynom je tedy kvadratická forma posunutá do bodu \mathbf{x} . Rozdíl je ale v tom, že pokud je kvadratická forma (pozitivně či negativně) semidefinitní, má v počátku extrém, zatímco Věta 9.2 o případu, kdy je $f''(\mathbf{x})$ semidefinitní, nic nepraví. V tom případě v bodě \mathbf{x} lokální extrém být může nebo nemusí (příkladem jsou funkce $f(x) = x^3$ a $f(x) = x^4$ v bodě $x = 0$). Bod \mathbf{x} , ve kterém je $f'(\mathbf{x}) = \mathbf{0}$ a matice $f''(\mathbf{x})$ je indefinitní, se nazývá **sedlový bod**.

Příklad 9.9. Extrémy kvadratické funkce (6.11) umíme hledat pomocí rozkladu na čtverec. Ovšem je to také možné pomocí derivací. Podmínka stacionarity je

$$\frac{d}{d\mathbf{x}}(\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c) = 2\mathbf{x}^T \mathbf{A}^T + \mathbf{b}^T = \mathbf{0}.$$

Po transpozici dostaneme rovnici (6.13a). Druh extrému určíme podle druhé derivace (Hessiánu), který je roven $2\mathbf{A}$ (předpokládáme symetrii \mathbf{A}). To souhlasí s klasifikací extrémů kvadratické formy z §6. \square

9.4 Lokální extrémy vázané rovnostmi

Hledejme minimum funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině

$$X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) = \mathbf{0}\}, \quad (9.2)$$

kde $\mathbf{g} = (g_1, \dots, g_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$. To odpovídá úloze (1.9) s omezeními typu rovnosti:

$$\begin{aligned} \min \quad & f(x_1, \dots, x_n) \\ \text{za podmíněk} \quad & g_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, m. \end{aligned} \quad (9.3)$$

Mluvíme o minimum funkce f *vázaném rovnostmi* $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

Množina X obsahuje všechna řešení soustavy $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, což je soustava m rovnic o n neznámých. Množina X obvykle nemá žádné vnitřní body, proto nelze použít Větu 9.1. V některých případech ale lze vyjádřit všechna řešení soustavy $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ parametricky a úlohu tak převést na úlohu bez omezení. Toto jsme použili v Příkladu 1.2, uveďme další příklady.

Příklad 9.10. Hledejme obdélník s jednotkovým obsahem a minimálním obvodem. Tedy minimalizujeme funkci $f(x, y) = x + y$ za podmínky $xy = 1$, neboli hledáme minima f na množině $X = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 1 - xy = 0\}$.

Množina X nemá žádné vnitřní body (dokažte!), proto nelze použít Větu 9.1. Z podmínky ale máme $y = 1/x$, což dosazeno do účelové funkce dá $f(x, 1/x) = x + 1/x$. Dle Věty 9.1 má tato funkce na svém definičním oboru dva stacionární body $x = \pm 1$. Tedy body podezřelé z lokálního extrému jsou $(x, y) = \pm(1, 1)$. \square

Příklad 9.11. Řešte úlohu

$$\begin{aligned} \min \quad & x + y \\ \text{za podmínky} \quad & x^2 + y^2 = 1, \end{aligned} \quad (9.4)$$

tedy hledáme minimum funkce $f(x, y) = x + y$ na množině $X = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 0\}$ kde $g(x, y) = 1 - x^2 - y^2$.

Množina X nemá žádné vnitřní body. Ale lze ji parametrizovat jako $X = \{(\cos t, \sin t) \mid t \in [0, 2\pi)\}$. Úlohu tak převedeme na hledání lokálních extrémů funkce jedné proměnné $f(\cos t, \sin t) = \cos t + \sin t$. Podmínka stacionarity $df(\cos t, \sin t)/dt = -\sin t + \cos t = 0$ má dvě řešení $t = \pm \frac{\pi}{2}$. Tedy body podezřelé z lokálního extrému jsou $(x, y) = \pm \frac{\sqrt{2}}{2}(1, 1)$. \square

Příklad 9.12. Chceme řešit problém $\min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}\}$. Všechna řešení soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ jdou vyjádřit jako $\mathbf{x} = \mathbf{C}\mathbf{y} + \mathbf{x}_0$, kde \mathbf{x}_0 je libovolné splňující $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$ a $\text{rng } \mathbf{C} = \text{null } \mathbf{A}$. Tedy úloha je ekvivalentní minimalizaci funkce $f(\mathbf{C}\mathbf{y} + \mathbf{x}_0)$ bez omezení. Počítat matici \mathbf{C} (např. najít bázi nulového prostoru matice \mathbf{A}) může být někdy ovšem nepříjemné. \square

Někdy ovšem množinu (9.2) parametrizovat nejde nebo je to složité. Nyní proto odvodíme obecnější postup, *metodu Lagrangeových multiplikátorů*.

9.4.1 Tečný a ortogonální prostor k povrchu

Zapomeňme nejprve na účelovou funkci f a zkoumejme jen množinu (9.2). Předpokládejme, že zobrazení \mathbf{g} je v okolí nějakého bodu $\mathbf{x} \in X$ spojitě diferencovatelné. V tom případě je množina X v okolí bodu \mathbf{x} ‘zakřivený povrch’² v \mathbb{R}^n . Pak existuje **tečný prostor** (množina všech tečných vektorů) a **ortogonální prostor** (množina všech kolmých vektorů) k povrchu X v bodě \mathbf{x} . Tyto dva prostory jsou ortogonální doplněk jeden druhého. Zde přesné definice pojmů ‘vektor tečný k povrchu’ a ‘vektor kolmý k povrchu’ neuvádíme a spoléháme na geometrickou intuici. Následující lema uvádíme bez důkazu.

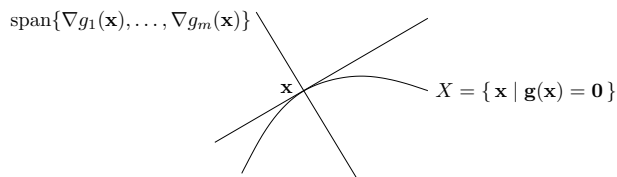
Lema 9.3. Necht' zobrazení $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je v bodě $\mathbf{x} \in X$ spojitě diferencovatelné. Necht'

$$\text{rang } \mathbf{g}'(\mathbf{x}) = m. \quad (9.5)$$

Pak ortogonální prostor k množině X v bodě \mathbf{x} je množina

$$\text{rng } \mathbf{g}'(\mathbf{x})^\perp = \text{span}\{\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})\}. \quad (9.6)$$

Viz obrázek:



² Přesněji, množina X je příkladem objektu, který se nazývá *diferencovatelný manifold*. Studium takových objektů se zabývá *diferenciální geometrií*.

Jelikož řádky Jacobiho matice $\mathbf{g}'(\mathbf{x})$ jsou gradienty $\nabla g_i(\mathbf{x})$, podmínka (9.5) vlastně říká, že gradienty $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$ musí být lineárně nezávislé. Bodu $\mathbf{x} \in X$ splňující podmínku (9.5) se někdy říká **regulární bod** povrchu.

Pro $m = 1$ podmínka (9.5) zní $\nabla g(\mathbf{x}) \neq \mathbf{0}$ a lema zobecňuje skutečnost, kterou jsme bez důkazu uvedli v §8.5, totiž že gradient funkce je v každém bodě kolmý k její vrstevnici. Lema ale navíc říká, že *každý* vektor kolmý k vrstevnici musí být násobek gradientu.

Příklad 9.13. Necht' $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ je funkce $g(x, y) = x^2 + y^2 - 1$. Množina X je jednotková kružnice v \mathbb{R}^2 . Máme $\nabla g(x, y) = (2x, 2y)$. Protože pro každé $(x, y) \in X$ je $\nabla g(x, y) \neq (0, 0)$, předpoklady Lematu 9.3 jsou splněny a ortogonální prostor k X v bodě (x, y) je množina $\text{span}\{\nabla g(x, y)\} = \{(\alpha x, \alpha y) \mid \alpha \in \mathbb{R}\}$, což je přímka kolmá ke kružnici. Tečný prostor v bodě (x, y) je ortogonální doplněk této přímky, tedy přímka tečná ke kružnici. \square

Příklad 9.14. Necht' $g: \mathbb{R}^3 \rightarrow \mathbb{R}$ je funkce $g(x, y, z) = x^2 + y^2 + z^2 - 1$. Množina X je jednotková sféra v \mathbb{R}^3 . Máme $\nabla g(x, y, z) = (2x, 2y, 2z)$. Ortogonální prostor k X v bodě (x, y, z) je množina $\text{span}\{\nabla g(x, y, z)\} = \{(\alpha x, \alpha y, \alpha z) \mid \alpha \in \mathbb{R}\}$, což je přímka kolmá ke sféře. Tečný prostor v bodě (x, y, z) je ortogonální doplněk této přímky, tedy rovina tečná ke sféře. \square

Příklad 9.15. Necht' $\mathbf{g} = (g_1, g_2): \mathbb{R}^3 \rightarrow \mathbb{R}^2$ je zobrazení

$$\mathbf{g}(x, y, z) = (x^2 + y^2 + z^2 - 1, (x - 1)^2 + y^2 + z^2 - 1).$$

Nulová vrstevnice funkce g_1 je jednotková sféra se středem v bodě $(0, 0, 0)$, nulová vrstevnice funkce g_2 je jednotková sféra se středem v bodě $(1, 0, 0)$. Množina X je průnik těchto dvou sfér, je to tedy kružnice v \mathbb{R}^3 . Máme $\nabla g_1(x, y, z) = 2(x, y, z)$ a $\nabla g_2(x, y, z) = 2(x - 1, y, z)$. Ortogonální prostor k množině X v bodě (x, y, z) je množina $\text{span}\{\nabla g_1(x, y, z), \nabla g_2(x, y, z)\} = \{\alpha_1(x, y, z) + \alpha_2(x - 1, y, z) \mid \alpha_1, \alpha_2 \in \mathbb{R}\}$, což je rovina kolmá ke kružnici v bodě (x, y, z) . Tečný prostor je ortogonální doplněk této množiny, tedy přímka tečná ke kružnici. \square

Příklad 9.16. Necht' $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ je funkce $g(x, y) = (x^2 + y^2 - 1)^2$. Množina X je stejná kružnice jako v Příkladě 9.13. Máme $\nabla g(x, y) = 4(x^2 + y^2 - 1)(x, y)$. Pro každý bod $(x, y) \in X$ je $\nabla g(x, y) = (0, 0)$, tedy předpoklady Lematu 9.3 nejsou splněny. Ortogonální prostor není množina $\text{span}\{\nabla g(x, y)\} = \{(0, 0)\}$. \square

9.4.2 Podmínky prvního řádu

Nyní přidáme do našich úvah i účelovou funkci f . Je intuitivně zřejmé (důkaz neuvádíme), že pokud \mathbf{x} má být lokální extrém funkce f na množině X , směrová derivace $f'(\mathbf{x})\mathbf{v} = \nabla f(\mathbf{x})^T \mathbf{v}$ funkce f v bodě \mathbf{x} v každém směru \mathbf{v} tečném k povrchu X musí být nulová. To znamená, že gradient $\nabla f(\mathbf{x})$ musí být kolmý k tečnému prostoru v bodě \mathbf{x} , neboli musí patřit do ortogonálního prostoru (9.6), neboli musí být lineární kombinací gradientů $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$. Tedy existují čísla $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ tak, že

$$\nabla f(\mathbf{x}) + \lambda_1 \nabla g_1(\mathbf{x}) + \dots + \lambda_m \nabla g_m(\mathbf{x}) = \mathbf{0}. \quad (9.7)$$

Výsledek těchto úvah se obvykle formuluje následujícím způsobem.

Věta 9.4. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \in X$. Necht'

- f a \mathbf{g} jsou v bodě \mathbf{x} spojitě diferencovatelné,

- $\text{rank } \mathbf{g}'(\mathbf{x}) = m$,
- bod \mathbf{x} je lokální extrém funkce f na množině X .

Pak existují čísla $(\lambda_1, \dots, \lambda_m) = \boldsymbol{\lambda} \in \mathbb{R}^m$ tak, že $L'(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$, kde funkce $L: \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ je dána jako

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x}) + \lambda_1 g_1(\mathbf{x}) + \dots + \lambda_m g_m(\mathbf{x}). \quad (9.8)$$

Zápis $L'(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$ označuje, že parciální derivace funkce L podle $x_1, \dots, x_n, \lambda_1, \dots, \lambda_m$ jsou nulové, neboli bod $(\mathbf{x}, \boldsymbol{\lambda}) \in \mathbb{R}^{m+n}$ je stacionární bod funkce L . Rovnost $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \mathbf{x} = \mathbf{0}$ je ekvivalentní rovnosti (9.7). Rovnost $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \boldsymbol{\lambda} = \mathbf{g}(\mathbf{x}) = \mathbf{0}$ je ekvivalentní omezením. Číslům λ_i se říká **Lagrangeovy multiplikátory** a funkci (9.8) **Lagrangeova funkce**.

Příklad 9.17. Řešme znovu Příklad 9.11. Lagrangeova funkce je

$$L(x, y, \lambda) = x + y + \lambda(1 - x^2 - y^2).$$

Její stacionární body (x, y, λ) jsou řešeními soustavy tří rovnic o třech neznámých

$$\begin{aligned} \partial L(x, y, \lambda)/\partial x &= 1 - 2\lambda x &= 0 \\ \partial L(x, y, \lambda)/\partial y &= 1 - 2\lambda y &= 0 \\ \partial L(x, y, \lambda)/\partial \lambda &= 1 - x^2 - y^2 &= 0. \end{aligned}$$

První dvě rovnice dávají $x = y = 1/(2\lambda)$. Dosazením do třetí máme $2/(2\lambda)^2 = 1$, což dá dva kořeny $\lambda = \pm 1/\sqrt{2}$. Stacionární body funkce L jsou dva, $(x, y, \lambda) = \pm(1, 1, 1)/\sqrt{2}$. Tedy máme dva kandidáty na lokální extrémy, $(x, y) = \pm(1, 1)/\sqrt{2}$.

Tuto jednoduchou úlohu je samozřejmě snadné vyřešit úvahou. Nakreslete si kružnici $X = \{(x, y) \mid x^2 + y^2 = 1\}$ a několik vrstevnic funkce f a najděte kýžené extrémy! \square

Příklad 9.18. Řešme Příklad 9.11, kde ale omezení změňme na $g(x, y) = (1 - x^2 - y^2)^2 = 0$. Podle Příkladu 9.16 máme $g'(x, y) = (0, 0)$ pro každé $(x, y) \in X$, čekáme tedy potíž.

Stacionární body Lagrangeovy funkce $L(x, y, \lambda) = x + y + \lambda(1 - x^2 - y^2)^2$ musí splňovat

$$\begin{aligned} \partial L(x, y, \lambda)/\partial x &= 1 - 4\lambda x(1 - x^2 - y^2) = 0 \\ \partial L(x, y, \lambda)/\partial y &= 1 - 4\lambda y(1 - x^2 - y^2) = 0 \\ \partial L(x, y, \lambda)/\partial \lambda &= (1 - x^2 - y^2)^2 = 0. \end{aligned}$$

Tyto rovnice si odporují. Jelikož $1 - x^2 - y^2 = 0$, tak např. první rovnice říká $1 - 4\lambda x \cdot 0 = 0$, což neplatí pro žádné (x, λ) . Závěr je, že lokální extrémy $(x, y) = \pm(1, 1)/\sqrt{2}$ jsme nenašli. \square

Příklad 9.19. Vraťme se k úloze (5.11), tedy k hledání řešení nehomogenní lineární soustavy s nejmenší normou. Lagrangeova funkce je

$$L(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{x}^T \mathbf{x} + 2\boldsymbol{\lambda}^T (\mathbf{b} - \mathbf{A}\mathbf{x}),$$

kde přidaná dvojka nemění situaci. Je $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \mathbf{x} = 2\mathbf{x}^T - 2\boldsymbol{\lambda}^T \mathbf{A}$ (odvoďte!). Stacionární body funkce L tedy získáme řešením soustavy (5.13), kterou jsme v 5.2 odvodili úvahou. \square

Předchozí příklad vyžaduje od studenta nejen znalost metody Lagrangeových multiplikátorů, ale i jistou zručnost v manipulaci s maticovými výrazy. Cvičte tuto zručnost ve Cvičnicích 9.22–9.25!

Věta 9.4 udává podmínky prvního řádu na extrémy vázané rovnostmi. Říká, že pokud $(\mathbf{x}, \boldsymbol{\lambda})$ je stacionární bod Lagrangeovy funkce, pak bod \mathbf{x} je 'podezřelý' z lokálního extrému funkce f na množině X . Jak poznáme, zda tento bod je lokální extrém, případně jaký? Podmínky druhého řádu pro vázané extrémy uvádíme nepovinně v §9.4.3. Zde pouze zdůrazníme, že druh lokálního extrému *nelze* zjistit podle definitnosti Hessovy matice $L''(\mathbf{x}, \boldsymbol{\lambda})$, tedy je chybou použít Větu 9.2 na funkci L .

9.4.3 Podmínky druhého řádu

Řekneme, že matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je pozitivně semidefinitní na podprostoru $X \subseteq \mathbb{R}^n$, jestliže $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$ pro každé $\mathbf{x} \in X$. Jak tuto podmínku ověříme? Najdeme-li matici \mathbf{B} , jejíž sloupce tvoří bázi podprostoru X , pak každý prvek $\mathbf{y} \in X$ lze parametrizovat jako $\mathbf{y} = \mathbf{B}\mathbf{x}$. Protože $\mathbf{y}^T \mathbf{A} \mathbf{y} = \mathbf{x}^T \mathbf{B}^T \mathbf{A} \mathbf{B} \mathbf{x}$, převedli jsme problém na ověřování pozitivní semidefinitnosti matice $\mathbf{B}^T \mathbf{A} \mathbf{B}$. Podobně definujeme pozitivní a negativní (semi)definitnost a indefinitnost matice \mathbf{A} na podprostoru X .

Věta 9.5. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \in \mathbb{R}^n$ a $\boldsymbol{\lambda} \in \mathbb{R}^m$. Necht'

- $(\mathbf{x}, \boldsymbol{\lambda})$ je stacionární bod Lagrangeovy funkce, neboli $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \mathbf{x} = \mathbf{0}$ a $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \boldsymbol{\lambda} = \mathbf{0}$,
- f a \mathbf{g} jsou dvakrát diferencovatelné v bodě \mathbf{x} .

Pak platí:

- Je-li $\partial^2 L(\mathbf{x}, \boldsymbol{\lambda})/\partial \mathbf{x}^2$ pozitivně [negativně] definitní na nulovém prostoru matice $\mathbf{g}'(\mathbf{x})$, má f v bodě \mathbf{x} ostré lokální minimum [maximum] vázané podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.
- Je-li $\partial^2 L(\mathbf{x}, \boldsymbol{\lambda})/\partial \mathbf{x}^2$ indefinitní na nulovém prostoru matice $\mathbf{g}'(\mathbf{x})$, nemá f v bodě \mathbf{x} lokální minimum ani lokální maximum vázané podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

Zde výraz

$$\frac{\partial^2 L(\mathbf{x}, \boldsymbol{\lambda})}{\partial \mathbf{x}^2} = f''(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i''(\mathbf{x})$$

značí druhou derivaci (Hessovu matici) funkce $L(\mathbf{x}, \boldsymbol{\lambda})$ podle \mathbf{x} v bodě $(\mathbf{x}, \boldsymbol{\lambda})$.

Příklad 9.20. Najdeme strany kváдру s jednotkovým objem a minimálním povrchem. Tedy minimalizujeme $xy + xz + yz$ za podmínky $xyz = 1$. Lagrangeova funkce je

$$L(x, y, z, \lambda) = xy + xz + yz + \lambda(1 - xyz).$$

Položením derivací L rovným nule máme soustavu

$$\begin{aligned} L'_x(x, y, z, \lambda) &= y + z - \lambda yz = 0 \\ L'_y(x, y, z, \lambda) &= x + z - \lambda xz = 0 \\ L'_z(x, y, z, \lambda) &= x + y - \lambda xy = 0 \\ L'_\lambda(x, y, z, \lambda) &= xyz - 1 = 0. \end{aligned}$$

Soustava je zjevně splněna pro $(x, y, z, \lambda) = (1, 1, 1, 2)$. Máme ukázat, že tento bod odpovídá lokálnímu minimu. Máme

$$\frac{\partial^2 L(x, y, z, \lambda)}{\partial(x, y, z)^2} = \begin{bmatrix} 0 & 1 - \lambda z & 1 - \lambda y \\ 1 - \lambda z & 0 & 1 - \lambda x \\ 1 - \lambda y & 1 - \lambda x & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}. \quad (9.9)$$

Ukážeme, že tato matice je pozitivně definitní na nulovém prostoru Jacobiho matice

$$g'(x, y, z) = [-yz \quad -xz \quad -xy] = [-1 \quad -1 \quad -1].$$

Nejdříve zkusme štěstí, zda matice (9.9) není pozitivně definitní již na \mathbb{R}^3 – v tom případě by zjevně byla pozitivně definitní i na nulovém prostoru $g'(x, y, z)$ (promyslete, proč to tak je!). Není tomu tak, protože její vlastní čísla jsou $\{-2, 1, 1\}$, tedy je indefinitní.

Nějakou bázi nulového prostoru matice $g'(x, y, z)$ snadno najdeme ručně, např.

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Snadno zjistíme, že matice

$$\mathbf{B}^T \frac{\partial^2 L(x, y, z, \lambda)}{\partial(x, y, z)^2} \mathbf{B} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

má vlastní čísla $\{2, 1\}$, tedy je pozitivně definitní. \square

9.5 Cvičení

9.1. Co je vnitřek a hranice těchto množin? Výsledek napište v množinovém zápisu.

- $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1, y \geq 0\}$
- $\{(x, y) \in \mathbb{R}^2 \mid y = x^2, -1 < x \leq 1\}$
- $\{(x, y) \in \mathbb{R}^2 \mid xy < 1, x > 0, y > 0\}$
- $\{\mathbf{x} \in \mathbb{R}^n \mid \max_{i=1}^n x_i \leq 1\}$
- $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}$, kde $\mathbf{a} \in \mathbb{R}^n, b \in \mathbb{R}$ (nadrovina)
- $\{\mathbf{x} \in \mathbb{R}^n \mid b < \mathbf{a}^T \mathbf{x} \leq c\}$, kde $\mathbf{a} \in \mathbb{R}^n, b, c \in \mathbb{R}$
- $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}\}$ kde $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m$

9.2. Je dána funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$, množiny $Y \subseteq X \subseteq \mathbb{R}^n$, a bod $\mathbf{x} \in Y$. Uvažujme dva výroky:

- Funkce f má v bodě \mathbf{x} lokální minimum na množině X .
- Funkce f má v bodě \mathbf{x} lokální minimum na množině Y .

Vyplývá (b) z (a)? Vyplývá (a) z (b)? Dokažte z definice lokálního extrému nebo vyvrát'te nalezením protipříkladu.

9.3. Může nastat případ, kdy funkce na množině má lokální minimum ale nemá na ní globální minimum? Odpověď dokažte.

9.4. Funkce $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ má stacionární bod $(2, 1, 5)$. Co se dá o tomto stacionárním bodě říci, když Hessova matice $f''(2, 1, 5)$ v něm má vlastní čísla

- $\{2, 3, -1\}$
- $\{2, 3, 0\}$
- $\{2, 1, 1\}$

9.5. Pro následující funkce spočítejte (na papíře) stacionární body. Pro každý stacionární bod určete, zda je to lokální minimum, lokální maximum, či sedlový bod. Pokud to určit nedokážete, odůvodněte.

- $f(x, y) = x(1 - \frac{2}{3}x^2 - y^2)$
- $f(x, y) = 1/x + 1/y + xy$
- $f(x, y) = e^y(y^2 - x^2)$
- $f(x, y) = 3x - x^3 - 3xy^2$
- $f(x, y) = 6xy^2 - 2x^3 - 3y^4$
- $f(x, y) = x^4/3 + y^4/2 - 4xy^2 + 2x^2 + 2y^2 + 3$
- $f(x, y, z) = x^3 + y^3 + 2xyz + z^2$

9.6. Dokažte, že funkce $f(x, y) = x$ nabývá za podmínky $x^3 = y^2$ minima pouze v počátku. Ukažte, že metoda Lagrangeových multiplikátorů toto minimum nenajde.

Následující úlohy se pokuste vyřešit parametrizací podmínky (analogicky k Příkladu 9.11) a pak metodou Lagrangeových multiplikátorů. Pokud jedna z těchto metod není použitelná, vnechte ji. Při použití metody Lagrangeových multiplikátorů stačí pouze najít stacionární body Lagrangeovy funkce – nemusíte určovat, jde-li o lokální extrémy a případně jaké.

9.7. Najděte lokální extrémy funkcí

- $f(x, y) = 2x - y$
- $f(x, y) = x(y - 1)$
- $f(x, y) = x^2 + 2y^2$
- $f(x, y) = x^2 y$
- $f(x, y) = x^4 + y^2$
- $f(x, y) = \sin(xy)$
- $f(x, y) = e^{xy}$

na kružnici $x^2 + y^2 = 1$. Nápodvěda: Někdy je dobré účelovou funkci zjednodušit, pokud to nezmění řešení.

9.8. Najděte extrémy funkce

- $f(x, y, z) = x + yz$ za podmínek $x^2 + y^2 + z^2 = 1$ a $z^2 = x^2 + y^2$
- $f(x, y, z) = xyz$ za podmínek $x^2 + y^2 + z^2 = 1$ a $xy + yz + zx = 1$

9.9. Najděte extrémy funkce

- $f(x, y, z) = (x + y)(y + z)$
- $f(x, y, z) = a/x + b/y + c/z$, kde $a, b, c > 0$ jsou dány
- $f(x, y, z) = x^3 + y^2 + z$
- $f(x, y, z) = x^3 + y^3 + z^3 + 2xyz$

e) (*) $f(x, y, z) = x^3 + y^3 + z^3 - 3xyz$

f) (*) $f(x, y, z) = x^3 + 2xyz - z^3$

na sféře $x^2 + y^2 + z^2 = 1$.

9.10. Rozložte dané kladné reálné číslo na součin n kladných reálných čísel tak, aby jejich součet byl co nejmenší.

9.11. Spočítejte rozměry tělesa tak, aby mělo při daném objemu nejmenší povrch:

a) kvádr

b) kvádr bez víka (má jednu dolní stěnu a čtyři boční, horní stěna chybí)

c) válec

d) püllitr (válec bez víka)

e) (*) kelímek (komolý kužel bez víka). Objem komolého kužele je $V = \frac{\pi}{3}h(R^2 + Rr + r^2)$ a povrch pláště (bez podstav) je $S = \pi(R+r)\sqrt{(R-r)^2 + h^2}$. Můžete použít vhodný numerický software na řešení vzniklé soustavy rovnic.

9.12. Najděte bod nejbližší počátku na křivce

a) $x + y = 1$

b) $x + 2y = 5$

c) $y = x^3 + 1$

d) $x^2 + 2y^2 = 1$

9.13. Necht' \mathbf{x}^* je bod nejbližší počátku na nadploše $h(\mathbf{x}) = 0$. Ukažte metodou Lagrangeových multiplikátorů, že vektor \mathbf{x}^* je kolmý k tečné nadrovině plochy v bodě \mathbf{x}^* .

9.14. Máme kouli o poloměru r a středu \mathbf{x}_0 , tj. množinu $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\| \leq r\}$. Máme nadrovinu $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}$.

9.15. Do elipsy o daných délkách os vepište obdélník s maximálním obsahem. Předpokládejte přitom, že strany obdélníku jsou rovnoběžné s osami elipsy.

9.16. *Fermatův princip* v paprskové optice říká, že cesta mezi libovolnými dvěma body na paprsku má takový tvar, aby ji světlo proběhlo za čas kratší než jí blízké dráhy. Později se zjistilo, že správným kritériem není *nejkratší* ale *extrémní* čas. Tedy skutečná dráha paprsku musí mít čas větší nebo menší než jí blízké dráhy. Z tohoto principu odvoďte:

a) Zákon odrazu od zrcadla: úhel dopadu se rovná úhlu odrazu.

b) Snellův zákon lomu: na rozhraní dvou prostředí se světlo lomí tak, že

$$\frac{c_1}{c_2} = \frac{\sin \alpha_1}{\sin \alpha_2},$$

kde α_i je úhel paprsku od normály rozhraní a c_i je rychlost světla v prostředí i .

Odvození udělejte:

(i) Pro rovinné zrcadlo a rovinné rozhraní (což vede na minimalizaci bez omezení).

(ii) Pro zrcadlo a rozhraní tvaru obecné plochy s rovnicí $g(\mathbf{x}) = 0$. Dokážete najít situaci, kdy skutečná dráha paprsku má čas *větší* než jí blízké dráhy?

9.17. Rozdělení pravděpodobnosti diskretní náhodné proměnné je funkce $p: \{1, \dots, n\} \rightarrow \mathbb{R}_+$ (tj. soubor nezáporných čísel $p(1), \dots, p(n)$) splňující $\sum_{x=1}^n p(x) = 1$.

a) *Entropie* náhodné proměnné s rozdělením p je rovna $-\sum_{x=1}^n p(x) \log p(x)$, kde \log je přirozený logaritmus. Najděte rozdělení s maximální entropií.

b) Dokažte *Gibbsovu nerovnost* (též zvanou *informační nerovnost*): pro každé dvě rozdělení p, q platí

$$\sum_{x=1}^n p(x) \log q(x) \geq \sum_{x=1}^n p(x) \log p(x),$$

přičemž rovnost nastává jen tehdy, když $p = q$.

9.18. (*) Máme trojúhelník se stranami délek a, b, c . Uvažujme bod, který má takovou polohu, že součet čtverců jeho vzdáleností od stran trojúhelníku je nejmenší možný. Jaké budou vzdálenosti x, y, z tohoto bodu od stran trojúhelníku?

9.19. (*) Máme krychli s délkou hrany 2. Do stěny krychle je vepsána kružnice (která má tedy poloměr 1) a okolo sousední stěny je opsána kružnice (která má tedy poloměr $\sqrt{2}$). Najděte nejmenší a největší vzdálenost mezi body na kružnicích.

9.20. (*) Najděte extrémy funkce

$$f(x, y, z, u, v, w) = (1 + x + u)^{-1} + (1 + y + v)^{-1} + (1 + z + w)^{-1}$$

za podmínek $xyz = a^3$, $uvw = b^3$ a $x, y, z, u, v, w > 0$.

9.21. Popište množinu řešení soustavy

$$x + 2y + z = 1$$

$$2x - y - 2z = 2.$$

Najděte takové řešení soustavy, aby výraz $\sqrt{x^2 + y^2 + z^2}$ byl co nejmenší. Najděte co nejvíce způsobů řešení.

9.22. Minimalizujte $\mathbf{x}^T \mathbf{x}$ za podmínky $\mathbf{a}^T \mathbf{x} = 1$. Jaký je geometrický význam úlohy?

9.23. Maximalizujte $\mathbf{a}^T \mathbf{x}$ za podmínky $\mathbf{x}^T \mathbf{x} = 1$. Jaký je geometrický význam úlohy?

9.24. Minimalizujte $\mathbf{x}^T \mathbf{A} \mathbf{x}$ za podmínky $\mathbf{b}^T \mathbf{x} = 1$, kde \mathbf{A} je symetrická pozitivně definitní.

9.25. Minimalizujte $\|\mathbf{C} \mathbf{x}\|$ za podmínky $\mathbf{A} \mathbf{x} = \mathbf{b}$, kde \mathbf{A} má lineárně nezávislé řádky a \mathbf{C} má lineárně nezávislé sloupce.

9.26. (*) Minimalizujte $\|\mathbf{A} \mathbf{x} - \mathbf{b}\|$ za podmínky $\mathbf{C} \mathbf{x} = \mathbf{0}$, kde \mathbf{A} má lineárně nezávislé sloupce a \mathbf{C} má lineárně nezávislé řádky.

9.27. (*) Minimalizujte $\|\mathbf{C} \mathbf{x}\|$ za podmínek $\mathbf{A} \mathbf{x} = \mathbf{0}$ a $\mathbf{x}^T \mathbf{x} = 1$.

9.28. (*) Minimalizujte $\|\mathbf{A} \mathbf{x}\|$ za podmínky $\mathbf{x}^T \mathbf{C} \mathbf{x} = 1$, kde \mathbf{C} je pozitivně definitní.

9.29. (*) Minimalizujte $\mathbf{a}^T \mathbf{x}$ za podmínky $\mathbf{x}^T \mathbf{C} \mathbf{x} = 1$, kde \mathbf{C} je pozitivně definitní.

9.30. (*) Jaké musí být vlastnosti matice \mathbf{A} a vektoru \mathbf{b} , aby $\max\{\|\mathbf{A} \mathbf{x}\| \mid \mathbf{b}^T \mathbf{x} = 0\} = 0$?

9.31. Necht' \mathbf{x} je lokální minimum funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině $X \subseteq \mathbb{R}^n$. Necht' $X' \subseteq X$. Je pravda, že \mathbf{x} je lokální minimum funkce f na množině X' ?

Nápověda a řešení

- 9.1.a) vnitřek \emptyset , hranice původní množina
9.1.b) vnitřek \emptyset , hranice $\{(x, y) \in \mathbb{R}^2 \mid y = x^2, -1 \leq x \leq 1\}$
9.1.c) vnitřek původní množina, hranice $\{(x, 0) \mid x \geq 0\} \cup \{(0, y) \mid y \geq 0\} \cup \{(x, y) \mid xy = 1\}$
9.1.d) $\max_{i=1}^n x_i \leq 1$ je totéž co $x_i \leq 1$ pro všechna i , tedy množina jde napsat také jako $(-\infty, 1]^n$ (kartézský součin n stejných polootevřených intervalů). Vnitřek je $(-\infty, 1)^n$, hranice (těžko se popíše krátkěji) je $(-\infty, 1]^n \setminus (-\infty, 1)^n$
9.1.e) vnitřek \emptyset , hranice původní množina
9.1.f) vnitřek $\{\mathbf{x} \mid b < \mathbf{a}^T \mathbf{x} < c\}$, hranice $\{\mathbf{x} \mid \mathbf{a}^T \mathbf{x} = b\} \cup \{\mathbf{x} \mid \mathbf{a}^T \mathbf{x} = c\}$
9.1.g) Vnitřek nemá, hranice je původní množina.
9.3. může
9.4.a) funkce nemá v tomto bodě lokální extrém
9.4.b) nemůžeme rozhodnout, zda má funkce má v tomto bodě lokální extrém
9.4.c) funkce má v tomto bodě lokální minimum
9.5.d) Stacionární body jsou 4.
9.5.e) Stacionární body jsou 3.
9.5.f) Stacionárních bodů je 5.
9.5.g) Stacionární body jsou 3, a to $(0, 0, 0)$, $(3/2, 3/2, -9/4)$, $(3/2, 3/2, -9/4)$.
9.16.a) Uděláme jen pro obecný případ (ii). Máme dva body \mathbf{a}, \mathbf{b} a hledáme bod \mathbf{x} splňující $g(\mathbf{x}) = 0$ pro který je celková dráha $\|\mathbf{x} - \mathbf{a}\| + \|\mathbf{x} - \mathbf{b}\|$ extrémální. Stacionární body Lagrangeovy funkce splňují $(\mathbf{x} - \mathbf{a})^0 + (\mathbf{x} - \mathbf{b})^0 = \lambda \nabla g(\mathbf{x})$, kde $\mathbf{y}^0 = \mathbf{y}/\|\mathbf{y}\|$. Ale to říká, že vektor $\nabla g(\mathbf{x})$ leží v jedné rovině s jednotkovými vektory $(\mathbf{x} - \mathbf{a})^0$ a $(\mathbf{x} - \mathbf{b})^0$ a půl úhel mezi nimi. Pro důkaz druhého tvrzení násobte rovnici vektory $(\mathbf{x} - \mathbf{a})^0$ a $(\mathbf{x} - \mathbf{b})^0$ a porovnejte.
9.24. $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}/(\mathbf{b}^T \mathbf{A}^{-1}\mathbf{b})$
9.25. $\mathbf{x} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{A}^T (\mathbf{A} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{A}^T)^{-1} \mathbf{b}$.
9.26. $\mathbf{x} = [\mathbf{I} - (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{C}^T (\mathbf{C} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{C}^T)^{-1} \mathbf{C}^T] (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$
9.31. Ano, z definice lokálního minima funkce na množině.

Kapitola 10

Iterační algoritmy na volné lokální extrémy

Zde se budeme věnovat numerickým iteračním algoritmům na nalezení volného lokálního minima diferencovatelných funkcí na množině \mathbb{R}^n .

10.1 Sestupné metody

Iterační algoritmy na hledání lokálního minima spojitě funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ mají tvar

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{v}_k, \quad (10.1)$$

kde vektor $\mathbf{v}_k \in \mathbb{R}^n$ je **směr hledání** a skalár $\alpha_k > 0$ je **délka kroku**. Ve třídě algoritmů zvaných **sestupné metody** (*descent methods*) hodnota účelové funkce monotonně klesá¹, $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$.

Nechť je funkce f diferencovatelná. Směr \mathbf{v}_k se nazývá **sestupný** v bodě \mathbf{x}_k , jestliže

$$f'(\mathbf{x}_k) \mathbf{v}_k < 0, \quad (10.2)$$

tedy směrová derivace ve směru \mathbf{v}_k je záporná. Pokud v bodě \mathbf{x}_k existuje sestupný směr, existuje $\alpha_k > 0$ tak, že $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$. Pokud v bodě \mathbf{x}_k sestupný směr neexistuje, vektor $f'(\mathbf{x}_k)$ je nutně nulový (proč?) a tedy \mathbf{x}_k je stacionární bod.

Máme-li sestupný směr, optimální délku kroku α_k najdeme minimalizací funkce f na polo-přímce z bodu \mathbf{x}_k ve směru \mathbf{v}_k . Tedy minimalizujeme funkci jedné proměnné

$$\varphi(\alpha_k) = f(\mathbf{x}_k + \alpha_k \mathbf{v}_k) \quad (10.3)$$

přes všechny $\alpha_k \geq 0$. Tato úloha je v kontextu vícerozměrné optimalizace nazývána *line search*. Úlohu stačí řešit přibližně. Takovou přibližnou metodu není obtížné vymyslet a proto se jí dále nebudeme zabývat.

Dále uvedeme nejznámější zástupce sestupných metod.

¹ Existují totiž i algoritmy, ve kterých hodnota $f(\mathbf{x}_k)$ neklesá monotonně (tj. někdy stoupne a někdy klesne) a přesto konvergují k optimu (např. *subgradientní metody*).

10.2 Gradientní metoda

Tato nejjednodušší metoda volí směr sestupu jako záporný gradient funkce f v bodě \mathbf{x}_k :

$$\mathbf{v}_k = -f'(\mathbf{x}_k)^T = -\nabla f(\mathbf{x}_k). \quad (10.4)$$

Tento směr je sestupný, což je okamžitě vidět dosazením do (10.2).

Nevýhodou gradientní metody je to, že konvergence může být pomalá kvůli 'cik-cak' chování. To se může stát tehdy, když funkce v okolí lokálního optima je v některých směrech mnohem protaženější než v jiných (přesněji, když vlastní čísla Hessiánu $f''(\mathbf{x})$ mají velmi různé velikosti). Výhodou metody je spolehlivost, protože směr je vždy sestupný.

10.2.1 (*) Závislost na lineární transformaci souřadnic

Transformujme vektor proměnných \mathbf{x} lineární transformací $\tilde{\mathbf{x}} = \mathbf{A}\mathbf{x}$, kde \mathbf{A} je regulární matice. Je jasné, že funkce f původních proměnných \mathbf{x} bude mít stejné extrémy jako funkce

$$\tilde{f}(\tilde{\mathbf{x}}) = \tilde{f}(\mathbf{A}\mathbf{x}) = f(\mathbf{x}) = f(\mathbf{A}^{-1}\tilde{\mathbf{x}}).$$

Iterace gradientní metody v nových proměnných je

$$\tilde{\mathbf{x}}_{k+1} = \tilde{\mathbf{x}}_k - \alpha_k \tilde{f}'(\tilde{\mathbf{x}}_k)^T. \quad (10.5)$$

Zkoumejme, jaké iteraci to odpovídá v původních proměnných. K tomu potřebujeme vyjádřit (10.5) v proměnných \mathbf{x} . Použitím řetězového pravidla odvodíme

$$\tilde{f}'(\tilde{\mathbf{x}}) = \frac{d\tilde{f}(\tilde{\mathbf{x}})}{d\tilde{\mathbf{x}}} = \frac{d\tilde{f}(\tilde{\mathbf{x}})}{d\mathbf{x}} \frac{d\mathbf{x}}{d\tilde{\mathbf{x}}} = \frac{df(\mathbf{x})}{d\mathbf{x}} \frac{d\mathbf{x}}{d\tilde{\mathbf{x}}} = f'(\mathbf{x})\mathbf{A}^{-1}.$$

Dosazením za $\tilde{\mathbf{x}}$ a $\tilde{f}'(\tilde{\mathbf{x}})$ do (10.5) a úpravou dostaneme

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k (\mathbf{A}^T \mathbf{A})^{-1} f'(\mathbf{x}_k)^T. \quad (10.6)$$

To lze napsat ve tvaru (10.1) se směrem hledání

$$\mathbf{v}_k = -(\mathbf{A}^T \mathbf{A})^{-1} f'(\mathbf{x}_k)^T. \quad (10.7)$$

Tento směr se liší od původního směru (10.4) vynásobením maticí $(\mathbf{A}^T \mathbf{A})^{-1}$. Vidíme, že gradientní metoda *není invariantní* vůči lineární transformaci souřadnic.

Ovšem lze ukázat, že nový směr (10.7) je také sestupný. Dosazením (10.4) do (10.2) to znamená, že $-f'(\mathbf{x}_k)(\mathbf{A}^T \mathbf{A})^{-1} f'(\mathbf{x}_k)^T < 0$. To je ale pravda, neboť matice $\mathbf{A}^T \mathbf{A}$ a tedy i její inverze je pozitivně definitní, viz Cvičení 6.16.

Na vzorec (10.7) se lze dívat ještě obecněji. Je jasné, že směr $\mathbf{v}_k = -\mathbf{C}_k^{-1} f'(\mathbf{x}_k)^T$ je sestupný, je-li matice \mathbf{C}_k je pozitivně definitní. Opačně, každý sestupný směr lze napsat takto. Uvidíme, že algoritmy uvedené dále budou mít vždy tento tvar sestupného směru, ovšem matice \mathbf{C}_k bude jiná v každém kroku.

10.3 Newtonova metoda

Newtonova metoda (přesněji Newton-Raphsonova) je slavný iterační algoritmus na řešení soustav nelineárních rovnic. Lze ho použít i na minimalizaci funkce tak, že hledáme její bod s nulovým gradientem. Oba způsoby použití popíšeme.

10.3.1 Použití na soustavy nelineárních rovnic

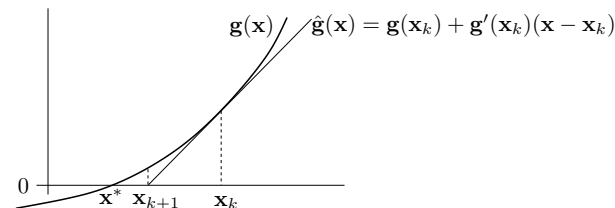
Řešme rovnici $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, kde $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ je diferencovatelné zobrazení. Jedná se tedy o soustavu n rovnic s n neznámými. Zobrazení \mathbf{g} aproximujeme v okolí bodu \mathbf{x}_k Taylorovým polynomem prvního stupně

$$\mathbf{g}(\mathbf{x}) \approx \hat{\mathbf{g}}(\mathbf{x}) = \mathbf{g}(\mathbf{x}_k) + \mathbf{g}'(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k), \quad (10.8)$$

kde Jacobiho matice $\mathbf{g}'(\mathbf{x}_k) \in \mathbb{R}^{n \times n}$ je derivace zobrazení v bodě \mathbf{x}_k . Další iteraci \mathbf{x}_{k+1} najdeme řešením nehomogenní lineární soustavy $\hat{\mathbf{g}}(\mathbf{x}_{k+1}) = \mathbf{0}$. Pokud je Jacobiho matice regulární, řešením je

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{g}'(\mathbf{x}_k)^{-1} \mathbf{g}(\mathbf{x}_k). \quad (10.9)$$

Viz obrázek:



Hlavní výhodou Newtonovy metody je, že v blízkém okolí řešení obvykle konverguje velmi rychle (mnohem rychleji než gradientní metoda). Nevýhodou je, že je nutno začít s poměrně přesnou aproximací \mathbf{x}_0 skutečného řešení, jinak algoritmus snadno diverguje.

Příklad 10.1. *Babylónská metoda* na výpočet druhé odmocniny čísla $a \geq 0$ je dána iterací

$$x_{k+1} = \frac{1}{2} \left(x_k + \frac{a}{x_k} \right).$$

To není nic jiného než Newtonova metoda pro řešení rovnice $0 = g(x) = x^2 - a$. Opravdu,

$$x_{k+1} = x_k - \frac{g(x_k)}{g'(x_k)} = x_k - \frac{x_k^2 - a}{2x_k} = x_k - \frac{1}{2} \left(x_k - \frac{a}{x_k} \right) = \frac{1}{2} \left(x_k + \frac{a}{x_k} \right). \quad \square$$

Příklad 10.2. Hledáme průsečík $(x, y) \in \mathbb{R}^2$ dvou křivek $(x-1)^2 + y^2 = 1$ a $x^4 + y^4 = 1$. Máme

$$\mathbf{g}(x, y) = \begin{bmatrix} (x-1)^2 + y^2 - 1 \\ x^4 + y^4 - 1 \end{bmatrix}, \quad \mathbf{g}'(x, y) = \begin{bmatrix} 2(x-1) & 2y \\ 4x^3 & 4y^3 \end{bmatrix}.$$

Iterace (10.9) je

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} - \begin{bmatrix} 2(x_k-1) & 2y_k \\ 4x_k^3 & 4y_k^3 \end{bmatrix}^{-1} \begin{bmatrix} (x_k-1)^2 + y_k^2 - 1 \\ x_k^4 + y_k^4 - 1 \end{bmatrix}.$$

Načrtneme-li si obě křivky, vidíme, že mají dva průsečíky, lišící se znaménkem druhé souřadnice. Zvolme počáteční odhad pro horní průsečík $(x_0, y_0) = (1, 1)$. První iterace bude

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 2 \\ 4 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 1 \end{bmatrix}.$$

Šestá iterace $(x_6, y_6) = (0.671859751039018, 0.944629015546222)$ je taková, že rovnice jsou splněny se strojovou přesností. \square

Příklad 10.3. Funkce $g(x) = x^2 - 1$ má dva nulové body $x = \pm 1$. Pokud v nějaké iteraci bude $x_k = 0$, nastane dělení nulou. Pokud bude x_k velmi malé, dělení nulou nenastane, ale iterace x_{k+1} se ocitne velmi daleko od kořene. \square

Příklad 10.4. Pro funkci $g(x) = x^3 - 2x + 2$ zvolme $x_0 = 0$. Další iterace bude $x_1 = 1$ a další $x_2 = 0$. Algoritmus bude oscilovat mezi hodnotami 0 a 1, tedy bude divergovat. \square

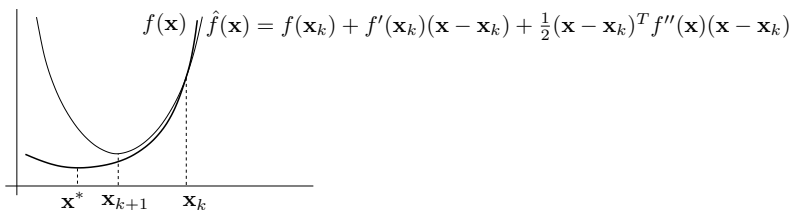
10.3.2 Použití na minimalizaci funkce

Newtonovu metodu lze použít pro hledání lokálního extrému dvakrát diferencovatelné funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ tak, že v algoritmu (10.9) položíme $\mathbf{g}(\mathbf{x}) = f'(\mathbf{x})^T$. Tím dostaneme iteraci

$$\mathbf{x}_{k+1} = \mathbf{x}_k - f''(\mathbf{x}_k)^{-1} f'(\mathbf{x}_k)^T, \quad (10.10)$$

kde $f''(\mathbf{x}_k)$ je Hessova matice funkce f v bodě \mathbf{x}_k .

Význam iterace (10.9) byl takový, že se zobrazení \mathbf{g} v bodě \mathbf{x}_k aproximovalo Taylorovým polynomem prvního stupně (tedy afinním zobrazením) a pak se našel kořen \mathbf{x}_{k+1} tohoto polynomu. Význam iterace (10.10) je takový, že se funkce f aproximuje Taylorovým polynomem druhého stupně (tedy kvadratickou funkcí) a pak se najde minimum této kvadratické funkce. Odvod'te podrobně, že tomu tak je!



Iteraci (10.10) lze napsat v obecnějším tvaru (10.1), kde

$$\mathbf{v}_k = -f''(\mathbf{x}_k)^{-1} f'(\mathbf{x}_k)^T. \quad (10.11)$$

Výhodou tohoto zobecnění je možnost zvolit optimální (ne nutně jednotkovou) délku kroku pomocí jednorozměrné minimalizace (10.3). Algoritmu (10.10) s jednotkovou délkou kroku se pak říká **čistá** Newtonova metoda.

Vektoru (10.11) říkáme **Newtonův směr**. Vidíme, že se od gradientního směru (10.4) liší násobením Hessovou maticí $f''(\mathbf{x}_k)$. Aby to byl sestupný směr, musí být

$$f'(\mathbf{x}_k) \mathbf{v}_k = -f'(\mathbf{x}_k) f''(\mathbf{x}_k)^{-1} f'(\mathbf{x}_k)^T < 0.$$

Toto platí, když $f'(\mathbf{x}_k) \neq \mathbf{0}$ (tj. \mathbf{x}_k není stacionární bod) a matice $f''(\mathbf{x}_k)$ je pozitivně definitní (neboť pak bude pozitivně definitní i její inverze, viz Cvičení 6.18).

V porovnání s gradientní metodou má Newtonova metoda (použitá na minimalizaci funkce) nevýhodu v tom, že musíme počítat Hessián $f''(\mathbf{x}_k)$ a řešit soustavu $f''(\mathbf{x}_k) \mathbf{v}_k = -f'(\mathbf{x}_k)^T$, což pro velký počet proměnných je pomalé či nemožné. Všimněte si ale, že na rozdíl od §10.3.1 je zde matice $\mathbf{g}'(\mathbf{x}_k) = f''(\mathbf{x}_k)$ symetrická, což může řešení soustavy ulehčit.

10.4 Nelineární metoda nejmenších čtverců

Řešme přeurčenou soustavu rovnic $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ pro $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ (tedy soustavu m rovnic s n neznámými) ve smyslu nejmenších čtverců. To vede na minimalizaci funkce

$$f(\mathbf{x}) = \|\mathbf{g}(\mathbf{x})\|^2 = \mathbf{g}(\mathbf{x})^T \mathbf{g}(\mathbf{x}) = \sum_{i=1}^m g_i(\mathbf{x})^2, \quad (10.12)$$

kde g_i jsou složky zobrazení \mathbf{g} . Speciálním případem je přibližné řešení lineární nehomogenní soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$, kde $\mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}$ (viz §5.1). Zde ovšem předpokládáme obecně nelineární zobrazení \mathbf{g} .

Zatímco v §10.2 a §10.3.2 bylo cílem minimalizovat *obecnou* funkci, zde chceme minimalizovat funkci ve speciálním tvaru (10.12). Nyní máme dvě možnosti. Bud' můžeme nasadit na funkci (10.12) jednu z metod pro minimalizaci obecné funkce, k čemuž se vrátíme v §10.4.2. Nebo můžeme být chytřejší a využít speciálního tvaru funkce (10.12), což popíšeme v §10.4.1.

10.4.1 Gauss-Newtonova metoda

Aproximujme opět zobrazení \mathbf{g} Taylorovým polynomem prvního stupně (10.8). Úloha (10.12) pak vyžaduje minimalizovat $\|\hat{\mathbf{g}}(\mathbf{x})\|^2$. To je úloha lineárních nejmenších čtverců, kterou již známe z §5.1. Normální rovnice (5.3) mají tvar

$$\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k) (\mathbf{x} - \mathbf{x}_k) = -\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k).$$

Pokud má Jacobiho matice $\mathbf{g}'(\mathbf{x}_k)$ lineárně nezávislé sloupce (tedy hodnost n , viz §5.1), tuto rovnici můžeme vyřešit pseudoinverzí:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{g}'(\mathbf{x}_k)^+ \mathbf{g}(\mathbf{x}_k). \quad (10.13)$$

Algoritmus (10.13) je znám jako **Gauss-Newtonova metoda**. Můžeme jej opět napsat obecněji ve tvaru (10.1) se směrem hledání

$$\mathbf{v}_k = -\mathbf{g}'(\mathbf{x}_k)^+ \mathbf{g}(\mathbf{x}_k) \quad (10.14a)$$

$$= -(\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} \mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k) \quad (10.14b)$$

$$= -\frac{1}{2} (\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} f'(\mathbf{x}_k)^T. \quad (10.14c)$$

Pro $m = n$ máme $\mathbf{g}'(\mathbf{x}_k)^+ = \mathbf{g}'(\mathbf{x}_k)^{-1}$, tedy Gauss-Newtonova metoda se redukuje na Newtonovu metodu (10.9) na řešení soustavy n rovnic s n neznámými.

Tvar (10.14c) dostaneme z (10.14b) dosazením derivace účelové funkce (10.12)

$$f'(\mathbf{x}) = 2\mathbf{g}(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$$

(viz §8.3.2). Vidíme, že Gauss-Newtonův směr (10.14c) se liší od gradientního směru (10.4) pouze násobením maticí $\frac{1}{2} (\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1}$. Aby byl tento směr sestupný, musí být

$$f'(\mathbf{x}_k) \mathbf{v}_k = -\frac{1}{2} f'(\mathbf{x}_k) (\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} f'(\mathbf{x}_k)^T < 0.$$

Toto platí, když $f'(\mathbf{x}_k) \neq \mathbf{0}$ a matice $\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k)$ je pozitivně definitní (viz Cvičení 6.18). Matice $\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k)$ je pozitivně definitní právě tehdy, když $\mathbf{g}'(\mathbf{x}_k)$ má lineárně nezávislé sloupce

(dokažte!), což ovšem již předpokládáme kvůli existenci inverze. Tedy vidíme, že za přirozených podmínek je Gauss-Newtonův směr vždy sestupný.

Čistá Gauss-Newtonova metoda (tj. s jednotkovou délkou kroku) může divergovat, a to i když je počáteční odhad \mathbf{x}_0 libovolně blízko lokálnímu minimu funkce (10.12). Protože ale Gauss-Newtonův směr je vždy sestupný, vhodnou volbou délky kroku α_k lze vždy zajistit konvergenci.

Příklad 10.5. Hledáme přibližné řešení soustavy tří rovnic $(x-1)^2 + y^2 = 1$, $x^4 + y^4 = 1$, $x^2 + (y-1)^2 = 1/2$ o dvou neznámých. Oba průsečíky křivek daných prvními dvěma rovnicemi již známe z Příkladu 10.2. Ani jeden z těchto průsečíků neleží na třetí křivce (i když je jí blízko), tedy soustava je přeřčená. Nezbyvá nám tedy, než ji řešit přibližně. Hledáme bod $(x, y) \in \mathbb{R}^2$, který minimalizuje číslo

$$f(x, y) = \mathbf{g}(x, y)^T \mathbf{g}(x, y) = ((x-1)^2 + y^2 - 1)^2 + (x^4 + y^4 - 1)^2 + (x^2 + (y-1)^2 - 1/2)^2$$

kde

$$\mathbf{g}(x, y) = \begin{bmatrix} (x-1)^2 + y^2 - 1 \\ x^4 + y^4 - 1 \\ x^2 + (y-1)^2 - 1/2 \end{bmatrix}, \quad \mathbf{g}'(x, y) = \begin{bmatrix} 2(x-1) & 2y \\ 4x^3 & 4y^3 \\ 2x & 2(y-1) \end{bmatrix}.$$

Rozumný počáteční odhad je $(x_0, y_0) = (1, 1)$. První Gauss-Newtonova iterace (10.13) je

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 2 \\ 4 & 4 \\ 2 & 0 \end{bmatrix}^+ \begin{bmatrix} 0 \\ 1 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 1 \end{bmatrix}.$$

Po osmé iteraci $(x_8, y_8) = (0.691002152515578, 0.940548357857245)$ se již hodnota $f(x_8, y_8) = 0.0008674592922855055$ v rámci strojové přesnosti nemění. \square

Příklad 10.6. V systému GPS máme m satelitů se známými souřadnicemi $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ a chceme spočítat souřadnice pozorovatele $\mathbf{x} \in \mathbb{R}^n$ z naměřených vzdáleností $y_i = \|\mathbf{a}_i - \mathbf{x}\|$ pozorovatele od satelitů. Měření jsou zatížena chybou, proto obecně tato soustava rovnic nebude mít žádné řešení. Řešme tuto přeřčenou nelineární soustavu ve smyslu nejmenších čtverců, tedy minimalizujeme funkci

$$f(\mathbf{x}) = \sum_{i=1}^m (\|\mathbf{x} - \mathbf{a}_i\| - y_i)^2.$$

Máme tedy $\mathbf{g} = (g_1, \dots, g_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$, kde $g_i(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}_i\| - y_i$. Derivace složek \mathbf{g} je (pomůžte nám §8.3.2, ale udělejte sami!) $g'_i(\mathbf{x}) = (\mathbf{x} - \mathbf{a}_i)^T / \|\mathbf{x} - \mathbf{a}_i\|$. Tedy

$$\mathbf{g}'(\mathbf{x}) = \begin{bmatrix} (\mathbf{x} - \mathbf{a}_1)^T / \|\mathbf{x} - \mathbf{a}_1\| \\ \vdots \\ (\mathbf{x} - \mathbf{a}_m)^T / \|\mathbf{x} - \mathbf{a}_m\| \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

Pak dosadíme do vzorečku (10.13). \square

10.4.2 Rozdíl proti Newtonově metodě

Předpokládejme, že bychom optimalizovali naši účelovou funkci (10.12) přímo Newtonovou metodou z §10.3.2. Spočítejme (proved'te sami!) Hessián funkce (10.12):

$$f''(\mathbf{x}) = 2\mathbf{g}'(\mathbf{x})^T \mathbf{g}'(\mathbf{x}) + 2 \sum_{i=1}^m g_i(\mathbf{x}) g''_i(\mathbf{x}). \quad (10.15)$$

Hessián je součtem členu obsahujícího derivace prvního řádu a členu obsahujícího derivace druhého řádu. Vidíme, že Gauss-Newtonův směr (10.14c) se liší od Newtonova směru (10.11) zanedbáním členu druhého řádu v Hessiánu (10.15). Jinými slovy, Gauss-Newtonovu metodu je možno vnímat jako aproximaci Newtonovy metody na minimalizaci funkce (10.12) spočívající v tom, že skutečný Hessián (10.15) se aproximuje výrazem $2\mathbf{g}'(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$.

To se projevuje tím, že Gauss-Newtonova metoda obvykle konverguje pomaleji než plná Newtonova metoda použitá na funkci (10.12). Ovšem vyhnuli jsme se počítání druhých derivací funkce \mathbf{g} , což je hlavní výhoda Gauss-Newtonovy metody.

10.4.3 Levenberg-Marquardtova metoda

Levenberg-Marquardtova metoda je široce používané vylepšení Gauss-Newtonovy metody, které její iteraci

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} \mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k)$$

nahrazuje iterací

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k) + \mu_k \mathbf{I})^{-1} \mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k) \quad (10.16)$$

kde $\mu_k > 0$. Přidání členu $\mu_k \mathbf{I}$ je vlastně regularizace, viz §7.3.1. Potom:

- Pro malé μ_k se (10.16) blíží Gauss-Newtonově iteraci.
- Pro velké μ_k je $(\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k) + \mu_k \mathbf{I})^{-1} \approx \mu_k^{-1} \mathbf{I}$, tedy (10.16) je blízká iteraci $\mathbf{x}_{k+1} = \mathbf{x}_k - \mu_k^{-1} \mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k)$ gradientní metody s délkou kroku μ_k^{-1} .

Tím jsou spojeny výhody Gauss-Newtonovy metody (typicky rychlá konvergence v okolí optima) a gradientní metody (spolehlivost i daleko od optima). Volbou parametru μ_k spojitě přecházíme mezi oběma metodami.

Parametr μ_k měníme během algoritmu. Začneme např. s $\mu_0 = 10^3$ a pak v každé iteraci:

- Pokud iterace snížila účelovou funkci, iteraci přijmeme a μ_k zmenšíme.
- Pokud iterace nesnížila účelovou funkci, iteraci odmítneme a μ_k zvětšíme.

Zvětšování a zmenšování μ_k děláme násobením a dělením konstantou, např. 10. Všimněte si, toto nahrazuje optimalizaci délky kroku α_k (*line search*).

Na algoritmus lze pohlížet i jinak. V iteraci (10.13) se počítá inverze matice $\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k)$. Tato matice je sice vždy pozitivně semidefinitní, ale může být blízká singulární (kdy se to stane?). To neblaze ovlivní stabilitu algoritmu. Matice (10.16) je ale vždy pozitivně definitní (viz Cvičení 6.17), a tedy regulární.

10.4.4 Statistické odůvodnění kritéria nejmenších čtverců

Možná se ptáte, proč se má nalezení přibližného řešení přeřčené soustavy formulovat zrovna jako (5.2). Nyní podáme statistický důvod, odkud se kritérium nejmenších čtverců vzalo.

Odhadujeme skryté parametry \mathbf{x} nějakého systému z měření \mathbf{y} na systému. Budiž vázány známou lineární závislostí $\mathbf{y} = \mathbf{Ax}$. Měření jsou zatížena chybami, které jsou způsobeny šumem senzorů, nepřesnostmi měření, nedokonalou znalostí modelu, apod. Tedy

$$\mathbf{y} = \mathbf{Ax} + \mathbf{r}, \quad (10.17)$$

kde $\mathbf{r} = (r_1, \dots, r_m)$ jsou náhodné proměnné modelující chyby měření $\mathbf{y} = (y_1, \dots, y_m)$. Metoda nejmenších čtverců říká, že máme minimalizovat $\|\mathbf{r}\|^2 = \sum_{i=1}^m r_i^2$, ale neříká proč.

Důvod odvodíme statistickou úvahou. Metoda činí dva předpoklady:

- Náhodné proměnné r_i mají normální (neboli Gaussovo) rozdělení s nulovou střední hodnotou a směrodatnou odchylkou σ , s hustotou pravděpodobnosti

$$p(r_i) = c e^{-r_i^2/(2\sigma^2)},$$

kde $c = (\sigma\sqrt{2\pi})^{-1}$ je normalizační konstanta.

- Náhodné proměnné r_1, \dots, r_m jsou na sobě nezávislé. Tedy sdružená hustota pravděpodobnosti je rovna součinu

$$p(\mathbf{r}) = p(r_1, \dots, r_m) = \prod_{i=1}^m p(r_i) = \prod_{i=1}^m c e^{-r_i^2/(2\sigma^2)}. \quad (10.18)$$

Dále použijeme *princip maxima věrohodnosti*. Ten říká, že parametry \mathbf{x} se mají najít tak, aby $p(\mathbf{r}) = p(\mathbf{y} - \mathbf{Ax})$ bylo maximální. Je pohodlnější minimalizovat záporný logaritmus

$$-\log p(r_1, \dots, r_m) = -\sum_{i=1}^m \log p(r_i) = \sum_{i=1}^m \left(\frac{r_i^2}{2\sigma^2} - \log c \right).$$

Jelikož σ je konstanta, je to totéž jako minimalizovat $\sum_i r_i^2$.

10.5 Cvičení

- 10.1. Najděte všechna řešení rovnice $\sin x = \frac{1}{2}x$ (sinus je v radiánech) na kalkulačce s největší přesností, jakou dokážete. Na kalkulačce smíte používat jen operace $+$, $-$, \times , $/$, \sin , \cos .
- 10.2. Najděte lokální extrém funkce $f(x, y) = x^2 - y + \sin(y^2 - 2x)$ čistou Newtonovou metodou. Počáteční odhad zvolte $(x_0, y_0) = (1, 1)$.
- 10.3. Máme m bodů v rovině o souřadnicích (x_i, y_i) , $i = 1, \dots, m$. Tyto body chceme proložit kružnicí ve smyslu nejmenších čtverců – tj. hledáme kružnici se středem (u, v) a poloměrem r takovou, aby součet čtverců kolmých vzdáleností bodů ke kružnici byl minimální. Zformulujte příslušnou optimalizační úlohu. Odvoďte iteraci Gauss-Newtonovy a Levenberg-Marquardtovy metody.
- 10.4. Máme soustavu rovnic

$$\begin{aligned} x + y - 2xy &= 1 \\ -x + y + xy &= -3 \\ x - y + xy &= 1 \end{aligned}$$

Je soustava lineární? Kolik má řešení a proč? Chceme soustavu řešit přibližně ve smyslu nejmenších čtverců, tj. minimalizovat funkci $f(x, y)$ ve tvaru (10.12). Napište iteraci

- a) gradientní metody,
- b) Newtonovy metody,
- c) Gauss-Newtonovy metody,
- d) Levenberg-Marquardtovy metody.

- 10.5. Rovnici $x^2 + 1 = 0$ lze vnímat jako soustavu s $n = 1$ neznámou a $m = 1$ rovnicí. Soustava nemá řešení, chceme ji tedy řešit přibližně Gauss-Newtonovou metodou. Ovšem protože je $m = n$, redukuje se Gauss-Newtonova iterace (10.13) na Newtonovu iteraci (10.9). Má tato formulace tedy smysl? K čemu bude metoda konvergovat?

Nápověda a řešení

- 10.1. Jeden kořen je $x = 0$ a pak dva další liší se znaménkem. Jeden z nich získáme Newtonovou metodou: $x \leftarrow x - (2 \sin x - x)/(2 \cos x - 1)$. Počáteční odhad zvolíme pomocí náčrtku $x = 2$. Po několika iteracích máme $x = 1.895494267033981$.
- 10.4. Soustava je nelineární. Nemá řešení, protože po zavedení proměnné $xy = z$ dostaneme lineární soustavu s řešením $(x, y, z) = (0.5, -1.5, -1)$, což je spor.

Kapitola 11

Lineární programování

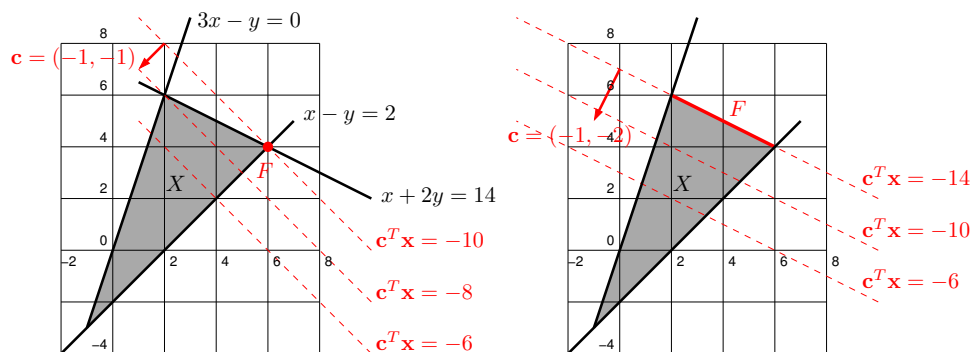
Lineární rovnici rozumíme výrok $a_1x_1 + \dots + a_nx_n = b$, neboli $h(\mathbf{x}) = 0$ kde h je afinní funkce. **Lineární nerovnici** rozumíme výrok $a_1x_1 + \dots + a_nx_n \leq b$ či $a_1x_1 + \dots + a_nx_n \geq b$, neboli $g(\mathbf{x}) \leq 0$ či $g(\mathbf{x}) \geq 0$ kde g je afinní funkce. Úloha **lineárního programování** (LP, také zvané lineární optimalizace) znamená minimalizaci lineární funkce za podmínek ve tvaru lineárních rovnic a nerovnic. Neboli v obecné formulaci (1.9) je funkce f lineární (tj. tvaru (3.6)) a funkce g_i, h_i jsou afinní (tj. tvaru (3.21)).

Jednoduché úlohy lineárního programování lze řešit graficky.

Příklad 11.1. Mějme lineární program

$$\begin{aligned} \min \quad & -x - y \\ \text{za podmíněk} \quad & x + 2y \leq 14 \\ & 3x - y \geq 0 \\ & x - y \leq 2 \end{aligned} \quad (11.1)$$

Množina $X = \{(x, y) \in \mathbb{R}^2 \mid x + 2y \leq 14, 3x - y \geq 0, x - y \leq 2\}$ přípustných řešení této úlohy je průnik tří polorovin. Tuto množinu snadno nakreslíme:



Účelová funkce $-x - y$, neboli $\mathbf{c}^T \mathbf{x}$ pro $\mathbf{x} = (x, y)$ a $\mathbf{c} = (-1, -1)$, má vrstevnice kolmé k vektoru \mathbf{c} a roste ve směru \mathbf{c} . Proto (viz levý obrázek) účelová funkce na množině X nabývá (globálního) minima v bodě $(x, y) = (6, 4)$. Úloha má tedy jediné optimální řešení.

Pokud bychom účelovou funkci úlohy (11.1) změnili na $-x - 2y$, bude tato funkce na množině X nabývat minima ve všech bodech úsečky spojující body $(2, 6)$ a $(6, 4)$ (viz pravý obrázek). Úloha má tedy nekonečně mnoho optimálních řešení. \square

Z našich úvah je patrné (přesně dokážeme později), že pro úlohu lineárního programování mohou nastat tři případy:

- úloha má (alespoň jedno) optimální řešení,
- úloha je *nepřípustná* (množina přípustných řešení je prázdná, omezení si odporují),
- úloha je *neomezená* (účelovou funkci lze za daných omezení libovolně zlepšovat).

11.1 Speciální tvary úloh LP

Algoritmy na řešení LP často předpokládají úlohu v nějakém speciálním tvaru, kdy jsou dovoleny pouze jisté typy omezení. Nejčastěji užívané speciální tvary jsou

- $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, kde $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$. Neboli

$$\begin{aligned} \min \quad & c_1x_1 + \dots + c_nx_n \\ \text{za podmíněk} \quad & a_{i1}x_1 + \dots + a_{in}x_n = b_i, \quad i = 1, \dots, m \\ & x_j \geq 0, \quad j = 1, \dots, n \end{aligned}$$

Tedy dovolujeme pouze omezení typu rovnosti a nezáporné proměnné¹.

- $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$
- $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$

Tyto speciální tvary LP nemají menší vyjadřovací schopnost než obecný tvar (ve kterém omezení mohou být libovolné lineární rovnice či nerovnice), neboť obecný tvar se dá efektivně² převést na libovolný speciální tvar následujícími úpravami:

- Maximalizaci funkce $\mathbf{c}^T \mathbf{x}$ nahradíme minimalizací funkce $-\mathbf{c}^T \mathbf{x}$.
- Nerovnost $\mathbf{a}^T \mathbf{x} \leq 0$ nahradíme nerovností $-\mathbf{a}^T \mathbf{x} \geq 0$.
- Rovnost $\mathbf{a}^T \mathbf{x} = b$ nahradíme dvěma nerovnostmi $\mathbf{a}^T \mathbf{x} \geq b$, $-\mathbf{a}^T \mathbf{x} \geq -b$.
- Nerovnost $\mathbf{a}^T \mathbf{x} \leq b$ převedeme na rovnost přidáním pomocné **slackové proměnné**³ $u_i \geq 0$ jako $\mathbf{a}^T \mathbf{x} + u_i = b_i$. Podobně převedeme nerovnost $\mathbf{a}^T \mathbf{x} \geq b$ na rovnost.
- Neomezenou proměnnou $x_i \in \mathbb{R}$ rozdělíme na dvě nezáporné proměnné $x_i^+ \geq 0$, $x_i^- \geq 0$ přidáním podmínky $x_i = x_i^+ - x_i^-$.

Úloha získaná z původní úlohy pomocí těchto úprav je ekvivalentní původní úloze v tom smyslu, že hodnota jejich optima je stejná a argument optima původní úlohy lze 'snadno' získat z argumentu optima nové úlohy.

Příklad 11.2. V úloze (11.1) chceme první podmínku převést na rovnost. To uděláme zavedením slackové proměnné $u \geq 0$. Transformovaná úloha je

$$\begin{aligned} \min \quad & -x - y \\ \text{za podmíněk} \quad & x + 2y + u = 14 \\ & 3x - y \geq 0 \\ & x - y \leq 2 \\ & u \geq 0 \end{aligned}$$

¹Tomuto tvaru se někdy říká *standardní*. Bohužel názvosloví různých tvarů LP není jednotné, názvy jako 'standardní tvar', 'základní tvar' či 'kanonický tvar' tedy mohou znamenat v různých knihách něco jiného.

²Slovem 'efektivně' zde přesně myslíme 'v lineárním čase', ve smyslu teorie algoritmů.

³Slack znamená anglicky např. mezeru mezi zdi a skříní, která není zcela přiřazená ke zdi. Termín *slack variable* nemá ustálený český ekvivalent, někdy se překládá jako *skluzová proměnná*.

Je-li (x, y, u) optimum této úlohy, optimum úlohy (11.1) je (x, y) . □

Příklad 11.3. V úloze (11.1) obě proměnné mohou mít libovolné znaménko. Chceme převést úlohu na tvar, kde všechny proměnné jsou nezáporné. Dosadíme $x = x_+ - x_-$ a $y = y_+ - y_-$, kde $x_+, x_-, y_+, y_- \geq 0$. Výsledná úloha je

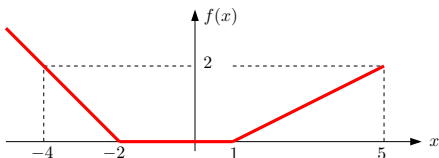
$$\begin{aligned} \min \quad & -x_+ + x_- - y_+ + y_- \\ \text{za podmínek} \quad & x_+ - x_- + 2y_+ - 2y_- \leq 14 \\ & 3x_+ - 3x_- - y_+ + y_- \geq 0 \\ & x_+ - x_- - y_+ + y_- \leq 2 \\ & x_+, x_-, y_+, y_- \geq 0 \end{aligned} \quad \square$$

11.1.1 Po částech afinní funkce

Mějme funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ danou vzorcem

$$f(\mathbf{x}) = \max_{i=1}^k (\mathbf{c}_i^T \mathbf{x} + d_i), \quad (11.2)$$

kde $\mathbf{c}_i \in \mathbb{R}^n$ a $d_i \in \mathbb{R}$ jsou dány. Tato funkce není lineární ani afinní, je po částech afinní (viz Cvičení 11.6). Příkladem pro $n = 1$ a $k = 3$ je funkce $f(x) = \max\{-x - 2, 0, \frac{1}{2}(x - 1)\}$, jejíž graf je na obrázku:



Řešme úlohu

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} \geq \mathbf{b}\}. \quad (11.3)$$

To není úloha LP, neboť její účelová funkce není lineární. Ovšem lze ji převést na LP zavedením pomocné proměnné:

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} \geq \mathbf{b}\} = \min\{z \mid (\mathbf{x}, z) \in \mathbb{R}^{n+1}, f(\mathbf{x}) \leq z, \mathbf{Ax} \geq \mathbf{b}\} \quad (11.4a)$$

$$= \min\{z \mid (\mathbf{x}, z) \in \mathbb{R}^{n+1}, \max_i (\mathbf{c}_i^T \mathbf{x} + d_i) \leq z, \mathbf{Ax} \geq \mathbf{b}\} \quad (11.4b)$$

$$= \min\{z \mid (\mathbf{x}, z) \in \mathbb{R}^{n+1}, (\forall i)(\mathbf{c}_i^T \mathbf{x} + d_i \leq z), \mathbf{Ax} \geq \mathbf{b}\} \quad (11.4c)$$

Rovnost (11.4a) platí proto, že v optimu pravé úlohy je $f(\mathbf{x}) = z$. Kdyby bylo $f(\mathbf{x}) < z$, mohli bychom totiž z zmenšit bez porušení omezení a tedy (\mathbf{x}, z) by nebylo optimum. Rovnost (11.4c) platí proto, že pro libovolná čísla a_i, b platí

$$\max_i a_i \leq b \iff (\forall i)(a_i \leq b). \quad (11.5)$$

Řádek (11.4c) je již úloha LP.

Příklad 11.4. Úloha

$$\begin{aligned} \min \quad & \max\{3x_1 + 4x_2, 2x_1 - 3x_2\} \\ \text{za podm.} \quad & x_1 + 2x_2 \leq 14 \\ & 3x_1 - x_2 \geq 0 \\ & x_1 - x_2 \leq 2 \end{aligned}$$

není LP, protože účelová funkce $f(x_1, x_2) = \max\{3x_1 + 4x_2, 2x_1 - 3x_2\}$ není lineární ani afinní (nakreslete si na papír její vrstevnice!). Úlohu lze ale přeformulovat na LP

$$\begin{aligned} \min \quad & z \\ \text{za podm.} \quad & 3x_1 + 4x_2 \leq z \\ & 2x_1 - 3x_2 \leq z \\ & x_1 + 2x_2 \leq 14 \\ & 3x_1 - x_2 \geq 0 \\ & x_1 - x_2 \leq 2 \end{aligned} \quad \square$$

Tento převod lze užít i pro funkce obsahující absolutní hodnoty, neboť $|x| = \max\{-x, x\}$. Je ale nutná opatrnost: neplatí nic takového jako $(\min_i a_i \leq b) \iff (\forall i)(a_i \leq b)$, tedy máme-li špatnou kombinaci minim/maxim a nerovností, převod na LP není možný.

11.2 Některé aplikace LP

11.2.1 Optimální výrobní program

Z m druhů surovin vyrábíme n druhů výrobků.

- a_{ij} = množství suroviny druhu i potřebné na výrobu výrobku druhu j
- b_i = množství suroviny druhu i , které máme k dispozici
- c_j = zisk z vyrobení jednoho výrobku druhu j
- x_j = počet vyrobených výrobků druhu j

Úkolem je zjistit, kolik jakých výrobků máme vyrobit, abychom dosáhli největšího zisku. Řešení:

$$\max \left\{ \sum_{j=1}^n c_j x_j \mid \sum_{j=1}^n a_{ij} x_j \leq b_i, x_j \geq 0 \right\}. \quad (11.6)$$

Příklad 11.5. Pán u stánku prodává lupínky za 120 Kč/kg a hranolky za 76 Kč/kg. Na výrobu 1 kg lupínků se spotřebuje 2 kg brambor a 0.4 kg oleje. Na výrobu 1 kg hranolky se spotřebuje 1.5 kg brambor a 0.2 kg oleje. Je nakoupeno 100 kg brambor a 16 kg oleje. Brambory stály 12 Kč/kg, olej 40 Kč/kg. Kolik má pán vyrobit lupínků a kolik hranolků, aby co nejvíce vydělal? To lze vyjádřit jako LP

$$\begin{aligned} \max \quad & 120l + 76h \\ \text{za podmínek} \quad & 2l + 1.5h \leq 100 \\ & 0.4l + 0.2h \leq 16 \\ & l, h \geq 0 \end{aligned}$$

Přitom předpokládáme, že zbytky surovin se po pracovní době vyhodí. Pokud se zbytky využijí, tak maximalizujeme $(120 - 24 - 16)l + (76 - 18 - 8)h = 80l + 50h$.

V obou případech je optimální řešení $l = 20$ kg lupínků a $h = 40$ kg hranolků. □

11.2.2 Směšovací (dietní) problém

Z n druhů surovin, z nichž každá je směsí m druhů látek, máme namíchat konečný produkt o požadovaném složení tak, aby cena surovin byla minimální.

- a_{ij} = množství látky druhu i obsažené v jednotkovém množství suroviny druhu j
- b_i = nejmenší požadované množství látky druhu i v konečném produktu
- c_j = jednotková cena suroviny druhu j
- x_j = množství suroviny druhu j

Řešení:

$$\min \left\{ \sum_{j=1}^n c_j x_j \mid \sum_{j=1}^n a_{ij} x_j \geq b_i, x_j \geq 0 \right\}. \quad (11.7)$$

Příklad 11.6. Jste kuchařka v menze a chcete uvařit pro studenty co nejlevnější oběd, ve kterém ovšem kvůli předpisům musí být dané minimální množství živin (cukrů, bílkovin a vitamínů). Oběd vaříte ze tří surovin: brambor, masa a zeleniny. Jsou dány hodnoty v tabulce:

| | na jednotku brambor | na jednotku masa | na jednotku zeleniny | min. požadavek na jeden oběd |
|----------------|------------------------|---------------------|-------------------------|---------------------------------|
| obsah cukrů | 2 | 1 | 1 | 8 |
| obsah bílkovin | 2 | 6 | 1 | 16 |
| obsah vitamínů | 1 | 3 | 6 | 8 |
| cena | 25 | 50 | 80 | |

Kolik je třeba každé suroviny na jeden oběd?

Minimalizujeme $25b+50m+80z$ za podmínek $2b+m+z \geq 8$, $2b+6m+z \geq 16$, $b+3m+6z \geq 8$ a $b, m, z \geq 0$. Optimální řešení je $b = 3.2$, $m = 1.6$, $z = 0$ s hodnotou 160. \square

11.2.3 Dopravní problém

Máme m výrobců a n spotřebitelů.

- a_i = množství zboží vyráběné výrobcem i
- b_j = množství zboží požadované spotřebitelem j
- c_{ij} = cena dopravy jednotky zboží od výrobce i ke spotřebiteli j
- x_{ij} = množství zboží vezené od výrobce i ke spotřebiteli j

Chceme co nejlevněji rozvézt zboží od výrobců ke spotřebitelům. Řešení:

$$\min \left\{ \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \mid \sum_{j=1}^n x_{ij} = a_i, \sum_{i=1}^m x_{ij} = b_j, x_{ij} \geq 0 \right\}. \quad (11.8)$$

Zadání musí splňovat $\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$ (nabídka musí být rovna poptávce), jinak bude úloha nepřipustná. Úloha jde modifikovat tak, že dovolíme $\sum_{i=1}^m a_i \geq \sum_{j=1}^n b_j$ (proved'te!).

11.3 Použití na nehomogenní lineární soustavy

11.3.1 Vektorové normy

Norma formalizuje pojem 'délky' vektoru \mathbf{x} . Známe již eukleidovskou normu, ale existují i jiné.

Definice 11.1. Funkce $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}$ se nazývá vektorová norma, jestliže splňuje tyto axiomy:

1. Jestliže $\|\mathbf{x}\| = 0$ pak $\mathbf{x} = \mathbf{0}$.
2. $\|\alpha\mathbf{x}\| = |\alpha| \|\mathbf{x}\|$ pro každé $\alpha \in \mathbb{R}$ a $\mathbf{x} \in \mathbb{R}^n$ (norma je kladně homogenní).
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ pro každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (trojúhelníková nerovnost).

Z axiomů plynou tyto další vlastnosti normy:

- $\|\mathbf{0}\| = 0$, což plyne z homogenity pro $\alpha = 0$
- $\|\mathbf{x}\| \geq 0$ pro každé $\mathbf{x} \in \mathbb{R}^n$. To jde odvodit tak, že v trojúhelníkové nerovnosti položíme $\mathbf{y} = -\mathbf{x}$, což dá

$$\|\mathbf{x} - \mathbf{x}\| = \|\mathbf{0}\| = 0 \leq \|\mathbf{x}\| + \|-\mathbf{x}\| = 2\|\mathbf{x}\|,$$

kde na pravé straně jsme použili homogenitu.

Jednotková sféra normy je množina $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$, tedy vrstevnice normy jednotkové výšky. Díky homogenitě je jednotková sféra středově symetrická a její tvar zcela určuje normu.

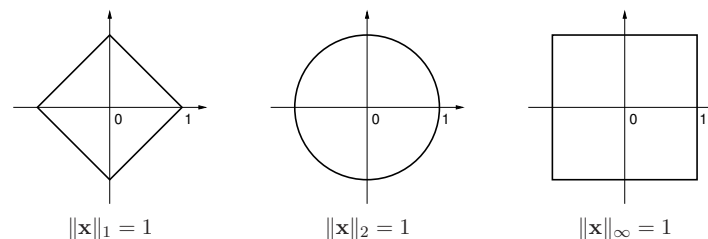
Uveďme příklady norem. Základním příkladem je p -norma

$$\|\mathbf{x}\|_p = (|x_1|^p + \dots + |x_n|^p)^{1/p}.$$

Musí být $p \geq 1$, jinak neplatí trojúhelníková nerovnost. Nejčastěji narazíte na:

- $\|\mathbf{x}\|_1 = |x_1| + \dots + |x_n|$. Někdy se jí říká *manhattanská norma*, protože v systému pravoúhlých ulic je vzdálenost mezi body \mathbf{x} a \mathbf{y} rovna $\|\mathbf{x} - \mathbf{y}\|_1$.
- $\|\mathbf{x}\|_2 = \sqrt{x_1^2 + \dots + x_n^2} = \sqrt{\mathbf{x}^T \mathbf{x}}$. Je to *eukleidovská norma*.
- $\|\mathbf{x}\|_\infty = \lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \max\{|x_1|, \dots, |x_n|\}$ (dokažte rovnost výpočtem limity!). Někdy se jí říká *Čebyševova norma* nebo *max-norma*.

Jednotkové sféry těchto norem v \mathbb{R}^2 vypadají takto:



Existují ale i normy, které nejsou p -normy, např.

- $\|\mathbf{x}\| = 2|x_1| + \sqrt{x_2^2 + x_3^2} + \max\{|x_4|, |x_5|\}$ je norma na \mathbb{R}^5 .
- Je-li $\|\mathbf{x}\|$ norma a \mathbf{A} je čtvercová nebo úzká matice s plnou hodnotostí, je také $\|\mathbf{Ax}\|$ norma.

11.3.2 Přibližné řešení přeurených soustav

Mějme přeurenou lineární soustavu $\mathbf{Ax} = \mathbf{b}$, kde $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{0} \neq \mathbf{b} \in \mathbb{R}^m$. Nalezení jejího přibližného řešení formulujeme jako úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_p. \quad (11.9)$$

Uvažujme tři případy:

- Pro $p = \infty$ hledáme takové \mathbf{x} , které minimalizuje výraz

$$\|\mathbf{Ax} - \mathbf{b}\|_\infty = \max_{i=1}^m |\mathbf{a}_i^T \mathbf{x} - b_i|, \quad (11.10)$$

tedy minimalizuje maximální residuum. Toto řešení je známé pod názvem *minimaxní* nebo *Čebyševovo*. Úloha je ekvivalentní lineárnímu programu

$$\begin{aligned} & \min z \\ & \text{za podm.} \quad \mathbf{a}_i^T \mathbf{x} - b_i \leq z, \quad i = 1, \dots, m \\ & \quad \quad \quad -\mathbf{a}_i^T \mathbf{x} + b_i \leq z, \quad i = 1, \dots, m \end{aligned}$$

neboli

$$\min \{ z \in \mathbb{R} \mid \mathbf{x} \in \mathbb{R}^n, -z\mathbf{1} \leq \mathbf{Ax} - \mathbf{b} \leq z\mathbf{1} \}. \quad (11.11)$$

- Pro $p = 2$ dostaneme řešení ve smyslu nejmenších čtverců, které jsme odvodili v §5.1.
- Pro $p = 1$ hledáme takové \mathbf{x} , které minimalizuje výraz

$$\|\mathbf{Ax} - \mathbf{b}\|_1 = \sum_{i=1}^m |\mathbf{a}_i^T \mathbf{x} - b_i|, \quad (11.12)$$

kde $\mathbf{a}_1, \dots, \mathbf{a}_m$ jsou řádky matice \mathbf{A} . Úloha je ekvivalentní lineárnímu programu

$$\begin{aligned} & \min \sum_{i=1}^m z_i \\ & \text{za podm.} \quad \mathbf{a}_i^T \mathbf{x} - b_i \leq z_i, \quad i = 1, \dots, m \\ & \quad \quad \quad -\mathbf{a}_i^T \mathbf{x} + b_i \leq z_i, \quad i = 1, \dots, m \end{aligned}$$

neboli

$$\min \{ \mathbf{1}^T \mathbf{z} \mid \mathbf{z} \in \mathbb{R}^m, \mathbf{x} \in \mathbb{R}^n, -\mathbf{z} \leq \mathbf{Ax} - \mathbf{b} \leq \mathbf{z} \}. \quad (11.13)$$

11.3.3 Lineární regrese

Vraťme se k lineární regresi z §5.1.2 (znovu přečtete!). Funkční závislost přibližně popsanou naměřenými dvojicemi (t_i, y_i) , $i = 1, \dots, m$, jsme aproximovali regresní funkcí

$$f(t, \mathbf{x}) = x_1 \varphi_1(t) + \dots + x_n \varphi_n(t) = \boldsymbol{\varphi}(t)^T \mathbf{x},$$

kde parametry \mathbf{x} jsou takové, aby $y_i \approx f(t_i, \mathbf{x})$ pro všechna i . Přibližné rovnosti \approx jsme chápali ve smyslu nejmenších čtverců, tedy hledali jsme takové \mathbf{x} které minimalizovalo funkci

$$\sum_{i=1}^m (y_i - f(t_i, \mathbf{x}))^2 = \|\mathbf{y} - \mathbf{Ax}\|_2, \quad (11.14)$$

kde $\mathbf{y} = (y_1, \dots, y_m)$ a prvky matice \mathbf{A} jsou $a_{ij} = \varphi_j(t_i)$. Tedy řešíme úlohu (11.9) pro $p = 2$. Můžeme ale použít i jiné normy než eukleidovskou. Pro $p = 1$ minimalizujeme

$$\sum_{i=1}^m |y_i - f(t_i, \mathbf{x})| = \|\mathbf{y} - \mathbf{Ax}\|_1 \quad (11.15)$$

a pro $p = \infty$ minimalizujeme

$$\max_{i=1}^m |y_i - f(t_i, \mathbf{x})| = \|\mathbf{y} - \mathbf{Ax}\|_\infty. \quad (11.16)$$

Dále ukážeme, k čemu to může být dobré.

Regrese ve smyslu ∞ -normy je vhodná např. při aproximaci funkcí.

Příklad 11.7. Na počítači bez matematického koprocesoru potřebujeme mnohokrát vyhodnocovat funkci sinus na intervalu $[0, \frac{\pi}{2}]$. Protože výpočet hodnot této funkce by trvalo příliš dlouho, chceme ji aproximovat polynomem třetího stupně $x_1 + x_2 t + x_3 t^2 + x_4 t^3$, jehož hodnoty se spočítají rychleji. Spočítejme hodnoty $y_i = \sin t_i$ funkce v dostatečném počtu bodů $t_i = \frac{\pi i}{2n}$ pro $i = 1, \dots, m$. Koeficienty polynomu hledáme minimalizací Čebyševova kritéria (11.16), neboť to nám dá záruku, že chyba aproximace nikde nepřesáhne hodnotu, která je nejmenší možná pro daný stupeň polynomu. \square

Regrese ve smyslu 1-normy je užitečná tehdy, když je malá část hodnot y_i naměřená úplně špatně (např. se někdo při zapisování čísel spletl v desetinné čárce). Takovým hodnotám se říká **vychýlené hodnoty** (*outliers*). Disciplína zabývající se modelováním funkčních závislostí za přítomnosti vychýlených hodnot se nazývá **robustní regrese**. V tomto případě řešení ve smyslu nejmenších čtverců není vhodné (není 'robustní'), protože i jediný vychýlený bod velmi ovlivní řešení. Regrese ve smyslu 1-normy je vůči vychýleným bodům odolnější.

Ukážeme to na nejjednodušším možném případě regrese: odhad hodnoty jediného čísla ze souboru jeho nepřesných měření. Pro daná čísla $\mathbf{y} = (y_1, \dots, y_m) \in \mathbb{R}^m$ hledáme $x \in \mathbb{R}$ minimalizující funkci

$$f(x) = \|(x - y_1, \dots, x - y_m)\|_1 = \|\mathbf{1}x - \mathbf{y}\|_1. \quad (11.17)$$

- Pro $p = \infty$ je $f(x) = \max_{i=1}^m |x - y_i|$. Řešením je $x = \frac{1}{2} (\min_{i=1}^m y_i + \max_{i=1}^m y_i)$, tedy bod v polovině mezi krajními body.
- Pro $p = 2$ je $f(x) = \sqrt{\sum_{i=1}^m (x - y_i)^2}$. Řešením je aritmetický průměr, $x = \frac{1}{m} \sum_{i=1}^m y_i$ (viz Příklad 5.3).
- Pro $p = 1$ je $f(x) = \sum_{i=1}^m |x - y_i|$. Řešením je *medián* z čísel y_i (dokažte jako cvičení!). Medián se vypočte tak, že seřadíme čísla y_i podle velikosti a vezmeme prostřední z nich. Pokud je m sudé, máme dva 'prostřední prvky' a v tom případě funkce f nabývá minima v jejich libovolné konvexní kombinaci. Je pak úzus definovat medián jako aritmetický průměr prostředních prvků.

Předpokládejme nyní, že jedno z čísel, např. y_1 , se zvětšuje. V tom případě se řešení x pro různá p budou chovat různě. Např. aritmetický průměr se bude zvětšovat, a to tak, že zvětšováním hodnoty y_1 dosáhneme *libovolně velké* hodnoty x . Pro medián to ovšem neplatí – zvětšováním jediného bodu y_1 ovlivníme x jen natolik, nakolik to změni pořadí bodů. Jeho libovolným zvětšováním nedosáhneme libovolně velké hodnoty x .

Příklad 11.8. Šuplérům změříme průměr ocelové kuličky v několika místech, dostaneme hodnoty $\mathbf{y} = (1.02, 1.04, 0.99, 2.03)$ (cm). Při posledním měření jsme se na stupnici přehlédli, proto je poslední hodnota úplně špatně. Z těchto měření chceme odhadnout skutečný průměr. Máme

$$\frac{1}{2} \left(\min_{i=1}^m y_i + \max_{i=1}^m y_i \right) = 1.51, \quad \frac{1}{m} \sum_{i=1}^m y_i = 1.27, \quad \text{median } y_i = 1.03.$$

Je zřejmé, že medián je neovlivněn vychýleným bodem, zatímco ostatní odhady ano. \square

Ve složitějším případě, např. prokládání dat polynomem jako v Příkladu 5.3, se nedá robustnost řešení ve smyslu 1-normy takto jednoduše formálně ukázat a analýza je obtížnější. Výsledek ale bude obdobný: řešení ve smyslu 1-normy je méně citlivé na vychýlené body než řešení ve smyslu 2-normy.

11.4 Cvičení

11.1. Najděte graficky množinu optimálních řešení úlohy

$$\begin{aligned} \min \quad & c_1 x_1 + c_2 x_2 + c_3 x_3 \\ \text{za podm.} \quad & x_1 + x_2 \geq 1 \\ & x_1 + 2x_2 \leq 3 \\ & x_1 + x_2 \leq 10 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

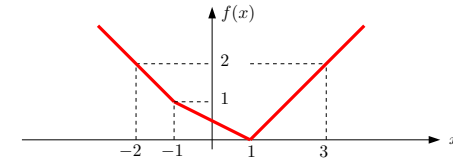
pro následující případy: (a) $\mathbf{c} = (-1, 0, 1)$, (b) $\mathbf{c} = (0, 1, 0)$, (c) $\mathbf{c} = (0, 0, -1)$.

11.2. Vyřešte úvahou tyto jednoduché úlohy LP a napište co nejjednodušší výraz pro optimální hodnotu. Správnost výsledku dokažte. Vektor $\mathbf{c} \in \mathbb{R}^n$ a číslo $k \in \mathbb{N}$, $1 \leq k \leq n$, jsou dány.

- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} = 1\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} \leq 1\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid -1 \leq \mathbf{1}^T \mathbf{x} \leq 1\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} = k\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{1}^T \mathbf{x} = k\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{1}^T \mathbf{x} \leq k\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid 0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq 1\}$
- (*) $\max\{\mathbf{c}^T \mathbf{x} \mid -\mathbf{y} \leq \mathbf{x} \leq \mathbf{y}, \mathbf{1}^T \mathbf{y} = k, \mathbf{y} \leq \mathbf{1}\}$

11.3. Pokud to dokážete, převed'te na LP.

- $\min\{|x_1| + |x_2| \mid 2x_1 - x_2 \geq 1, -x_1 + 2x_2 \geq 1\}$
- $\max\{|x_1 - c_1| + \dots + |x_n - c_n| \mid a_1 x_1 + \dots + a_n x_n \geq b\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, |\mathbf{d}^T \mathbf{x}| \leq 1, \mathbf{x} \geq \mathbf{0}\}$
- $\min_{\mathbf{x} \in \mathbb{R}^n} \sum_{l=1}^L \max_{k=1}^K (c_{kl}^T \mathbf{x} + d_{kl})$
- $\min_{\mathbf{x} \in \mathbb{R}^n} \sum_{i=1}^m f(\mathbf{a}_i^T \mathbf{x} - b_i)$, kde funkce f je definována obrázkem



- $\min\{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_1 \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\|_\infty \leq 1\}$
- $\min\{\|\mathbf{x}\|_1 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}\}$
- $\min\{\|\mathbf{x}\|_1 \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_\infty \leq 1\}$
- $\min_{\mathbf{x} \in \mathbb{R}^n} (\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_1 + \|\mathbf{x}\|_\infty)$

11.4. Máme algoritmus (černou skříňku) na řešení LP, kterou můžeme zavolat i vícekrát. S pomocí tohoto algoritmu vyřešte úlohu $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$?

11.5. Dokažte nebo vyvrat'te následující rovnosti. Zde $\mathbf{c} \in \mathbb{R}^n$ a $\mathbf{A} \in \mathbb{R}^{m \times n}$ jsou dány, $\|\cdot\|$ je libovolná norma, a optimalizuje se přes proměnné $\mathbf{x} \in \mathbb{R}^n$.

- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| = 1\} = \max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| \leq 1\}$
- $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| = 1\} = \min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| \leq 1\}$
- $\max\{\|\mathbf{A}\mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| = 1\} = \max\{\|\mathbf{A}\mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| \leq 1\}$

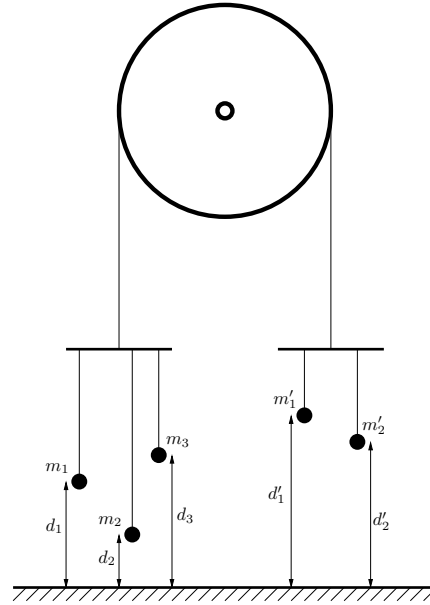
Nápověda: Inspirujte se úvahou v §11.1.1.

11.6. Pochopte kód v Matlabu, který nakreslí graf funkce $f(\mathbf{x}) = \max_{i=1}^k (c_i^T \mathbf{x} + d_i)$ pro $\mathbf{x} \in \mathbb{R}^2$:

```
k = 200; N = 40;
cd = randn(3,k);
x1 = ones(N,1)*linspace(-1,1,N); x2 = linspace(-1,1,N)*ones(1,N);
x = [x1(:)'; x2(:)']; x(3,:) = 1;
meshc(x1,x2,reshape(max(cd'*x,[],1),[N N])); axis vis3d
```

11.7. Hledáme největší hyperkouli $B(\mathbf{a}, r) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}\|_2 \leq r\}$, která se vejde do mnohostěnu $P = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$. Tedy hledáme maximální r za podmínky $B(\mathbf{a}, r) \subseteq P$, kde optimalizujeme přes proměnné (\mathbf{a}, r) . Vyjádřete jako LP.

11.8. Máme kladku s provazem, jehož oba konce končí hákem. Na levém háku visí n závaží na provázcích, přičemž i -té závaží má tíhu m_i a jeho výška nad zemí je d_i , pro $i = 1, \dots, n$. Na pravém háku visí n' závaží na provázcích, přičemž i -té závaží má tíhu m'_i a jeho výška nad zemí je d'_i , pro $i = 1, \dots, n'$. Výšky d_i a d'_i se měří v poloze, kdy jsou oba háky ve stejné výšce nad zemí. Kladka se pohybuje bez tření, provaz a provázky jsou nekonečně ohebné, provázky a háky mají nulovou hmotnost. Obrázek ukazuje příklad pro $n = 3$, $n' = 2$.



Soustava má jediný stupeň volnosti daný otáčením kladky. Označme jako x výšku levého háku nad bodem, kdy jsou oba háky ve stejné výšce – tedy pro $x = 0$ jsou oba háky ve stejné výšce a pro $x > 0$ bude levý hák o $2x$ výše než pravý hák. V závislosti na x každé závaží buď visí nad zemí (pak je jeho potenciální energie rovna m_i krát výška nad zemí) nebo leží na zemi (pak je jeho potenciální energie nulová). Soustava bude v rovnováze při minimální celkové potenciální energii.

- Napište vzorec pro celkovou potenciální energii soustavy jako funkci x .
- Napište lineární program, jehož optimum je rovno minimální potenciální energii soustavy. Není-li to možné, vysvětlete.

11.9. Uvažujme Příklad 11.5 takto pozměněný: pán má pomocníka, kterému platí 10 Kč za každý kg vyrobeného zboží (je jedno, zda to jsou lupínky nebo hranolky). Ovšem pokud se toho vyrobí hodně, chce pomocník větší plat, protože musí zůstat přesčas. Tak za každý kg nad 20 kg vyrobeného zboží si nechá připlatit dalších 10 Kč, a za každý kg nad 30 kg vyrobeného zboží si nechá připlatit dalších 20 Kč (tedy za každý kg nad 30 kg vyrobeného zboží dostane $10 + 10 + 20 = 40$ Kč). Kolik má pán vyrobit lupínků a hranolků, aby měl co největší denní zisk (tj. tržbu z prodeje minus plat pomocníkovi)? Zformulujte jako LP.

11.10. Veverka před zimou potřebuje přerovnat zásoby oříšků. Stávající zásoby má v m jamkách, přičemž i -tá jamka má souřadnice $\mathbf{p}_i \in \mathbb{R}^2$ a je v ní a_i oříšků. Potřebuje je přenosit do n nových připravených jamek, přičemž j -tá jamka má souřadnice $\mathbf{q}_j \in \mathbb{R}^2$ a na konci v ní bude y_j oříšků. Veverka unese najednou jen jeden oříšek. Necht' x_{ij} označuje celkový počet oříšků přenesených ze staré jamky i do nové jamky j . Uvažujte dvě úlohy:

- Čísla y_j jsou dána. Hledají se taková čísla x_{ij} , aby se veverka vykonala co nejméně práce, kde práce na přenesení jednoho oříšku je přímo úměrná vzdálenosti (vzdušnou čarou). Běh bez oříšku se za práci nepovažuje.
- Hledají se čísla x_{ij} a y_j tak, aby veverka vykonala co nejméně práce a navíc byly v nových jamkách oříšky rozloženy co nejrovnoměrněji, čímž minimalizuje škodu způsobenou případnou krádeží. Přesněji, aby rozdíl mezi největším a nejmenším z čísel y_j byl menší než dané číslo t .

Formulujte obě úlohy jako LP. Předpokládejte, že počty oříšků jsou nezáporná reálná čísla, ač ve skutečnosti mohou být pouze nezáporná celá čísla.

Nápověda a řešení

- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\} = \sum_{i=1}^n \max\{0, c_i\}$, tedy optimální hodnota je součet kladných čísel c_i .
Důkaz: Ukažme, že optimum se nabývá pro takové \mathbf{x} , že $x_i = 0$ pro $c_i < 0$ a $x_i = 1$ pro $c_i > 0$ (pro $c_i = 0$ je x_i libovolné). Kdyby to tak totiž nebylo, mohli bychom číslo $\mathbf{c}^T \mathbf{x}$ zvětšit zmenšením nějakého x_i pro $c_i < 0$ nebo zvětšením pro $c_i > 0$. Tedy \mathbf{x} by nebyl optimální argument.
 - $\sum_{i=1}^n |c_i|$. Dokáže se podobně.
 - $\max_{i=1}^n c_i$
 - $\max_{i=1}^n \max\{0, c_i\} = \max\{0, \max_{i=1}^n c_i\}$
 - Když $c_i = a$ pro každé i (tj. všechna c_i jsou stejná), tak optimální hodnota je $|a|$. Jinak je úloha neomezená.
 - Nápověda: substituujte $y_i = x_i - x_{i-1}$
 - $\min\{z_1 + z_2 \mid x_1, x_2, z_1, z_2 \in \mathbb{R}, 2x_1 - x_2 \geq 1, -x_1 + 2x_2 \geq 1, x_1 \leq z_1, x_2 \leq z_2, -x_1 \leq z_1, -x_2 \leq z_2\}$
 - Nejde.
 - $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{d}^T \mathbf{x} \leq 1, -\mathbf{d}^T \mathbf{x} \leq 1, \mathbf{x} \geq \mathbf{0}\}$
 - $\min\{\mathbf{1}^T \mathbf{z} \mid \mathbf{c}_{kl}^T \mathbf{x} + d_{kl} \leq z_l (\forall k, l), -\mathbf{c}_{kl}^T \mathbf{x} - d_{kl} \leq z_l (\forall k, l), \mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{R}^L\}$ (analogické §11.1.1)
 - $\min\{\mathbf{1}^T \mathbf{z} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{R}^m, -\mathbf{z} \leq \mathbf{A}\mathbf{x} - \mathbf{b} \leq \mathbf{z}, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}\}$
 - $\min\{\mathbf{1}^T \mathbf{y} + z \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m, z \in \mathbb{R}, -\mathbf{z} \leq \mathbf{A}\mathbf{x} - \mathbf{b} \leq \mathbf{z}, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}\}$
 - Postupem uvedeným v §11.1 nedokážeme převést na jedinou úlohu LP. Ale lze vyřešit vypočtením dvou úloh LP: optimální hodnota je $\max\{A, -B\}$, kde $A = \max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ a $B = \max\{-\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$.
 - $E(x) = \sum_{i=1}^n m_i \max(d_i + x, 0) + \sum_{i=1}^{n'} m'_i \max(d'_i - x, 0)$
 - $\min\{\sum_{i=1}^n m_i z_i + \sum_{i=1}^{n'} m'_i z'_i \mid x, z_i, z'_i \in \mathbb{R}, z_i \geq d_i + x, z'_i \geq d'_i - x, z_i \geq 0, z'_i \geq 0\}$
- 11.9. Stejná úloha jako v Příkladu 11.5, jen účelová funkce se změní na $120l + 76h - f(l + h)$ kde $f(t) = \max\{10t, 200 + 20(t - 20), 400 + 40(t - 30)\}$. To převedeme na LP

$$\begin{array}{ll} \max & 120l + 76h - z \\ \text{za podmíněk} & 10l + 10h - z \leq 0 \\ & 20l + 20h - z \leq 200 \\ & 40l + 40h - z \leq 800 \\ & 2l + 1.5h \leq 100 \\ & 0.4l + 0.2h \leq 16 \\ & l, h \geq 0 \end{array}$$

Kapitola 12

Konvexní množiny a mnohostěny

12.1 Konvexní množiny

Množina $X \subseteq \mathbb{R}^n$ se nazývá **konvexní**, jestliže

$$\mathbf{x} \in X, \mathbf{y} \in X, 0 \leq \alpha \leq 1 \implies \alpha \mathbf{x} + (1 - \alpha) \mathbf{y} \in X. \quad (12.1)$$

Množina $\{\alpha \mathbf{x} + (1 - \alpha) \mathbf{y} \mid 0 \leq \alpha \leq 1\}$ je úsečka spojující body \mathbf{x} a \mathbf{y} (viz Příklad 3.6). Definice tedy říká, že množina je konvexní, jestliže s každými dvěma body obsahuje i úsečku, která je spojuje. Obrázek ukazuje příklad konvexní a nekonvexní množiny v \mathbb{R}^2 :



Konvexní množinu lze definovat i abstraktněji. **Konvexní kombinace** vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ je jejich lineární kombinace $\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$ taková, že $\alpha_1 + \dots + \alpha_k = 1$ a $\alpha_1, \dots, \alpha_k \geq 0$. Množina je konvexní právě tehdy, když je uzavřená vůči konvexním kombinacím (neboli každá konvexní kombinace vektorů z množiny leží v množině). Lze dokázat indukci, že tato definice je ekvivalentní (12.1). Všimněte si, že $\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}$ pro $0 \leq \alpha \leq 1$ je konvexní kombinací dvou vektorů \mathbf{x}, \mathbf{y} , neboť $\alpha + (1 - \alpha) = 1$, $\alpha \geq 0$, $1 - \alpha \geq 0$.

Konvexní obal vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je množina všech jejich konvexních kombinací, značíme

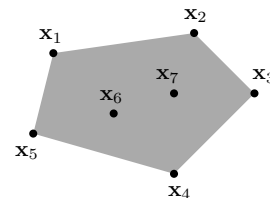
$$\text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} = \{\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k \mid \alpha_1 + \dots + \alpha_k = 1, \alpha_1, \dots, \alpha_k \geq 0\}. \quad (12.2)$$

Jak ale definovat konvexní obal množiny s *nekonečným* počtem prvků, např. pravém obrázku výše? Nelze použít definice (12.2), neboť není jasné, co znamená součet $\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$ pro nekonečný počet vektorů. Konvexní obal libovolné (konečné či nekonečné) množiny $X \subseteq \mathbb{R}^n$ definujeme jako průnik všech konvexních množin, které množinu obsahují, tedy

$$\text{conv } X = \bigcap \{Y \subseteq \mathbb{R}^n \mid Y \supseteq X, Y \text{ konvexní}\}.$$

Tato definice je poněkud abstraktní a nepraktická. Zdánlivě jasněji můžeme $\text{conv } X$ definovat jako ‘nejmenší’ konvexní množinu, která obsahuje množinu X . Ovšem zde musíme dát pozor na to, co přesně znamená slovo ‘nejmenší’, neboť ne pro každé dvě množiny lze říci, která z nich je menší (lze to říci jen pro množiny, jež jsou podmnožinou jedna druhé). Přesně, ‘nejmenší’ znamená *minimální prvek* v částečném uspořádání podmnožin relací inkluze (\subseteq).

Obrázek ukazuje konvexní obal konečné (vlevo) a nekonečné (vpravo) množiny pro $n = 2$:



Věta 12.1. Průnik (konečně či nekonečně mnoha) konvexních množin je konvexní množina.

Důkaz. Stačí dokázat pro dvě množiny, pro více množin věta plyne z asociativity operace průnik. Necht' $X, Y \subseteq \mathbb{R}^n$ jsou konvexní. Necht' $\mathbf{x}, \mathbf{y} \in X \cap Y$, tedy $\mathbf{x}, \mathbf{y} \in X$ a $\mathbf{x}, \mathbf{y} \in Y$. Proto pro $0 \leq \alpha \leq 1$ je bod $\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}$ také v X i Y , tedy je v $X \cap Y$. \square

Sjednocení konvexních množin ale nemusí být konvexní množina.

12.2 Čtyři kombinace a čtyři obaly

Konvexní kombinace je lineární kombinace, jejíž koeficienty splňují omezení $\alpha_1 + \dots + \alpha_k = 1$ a $\alpha_1, \dots, \alpha_k \geq 0$. Všimněte si, že když vynecháme druhé omezení, dostaneme afinní kombinaci (viz §3.5). Podle toho, které ze dvou omezení vyžadujeme, dostaneme čtyři druhy kombinací. Udělejme si v nich nyní pořádek.

Vážený součet $\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$ vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ se nazývá jejich

| | | | |
|------------------------------|----------|--|--|
| lineární kombinace , | jestliže | $\alpha_1, \dots, \alpha_k \in \mathbb{R}$. | |
| afinní kombinace , | jestliže | $\alpha_1, \dots, \alpha_k \in \mathbb{R}$, | $\alpha_1 + \dots + \alpha_k = 1$. |
| nezáporná kombinace , | jestliže | $\alpha_1, \dots, \alpha_k \in \mathbb{R}$, | $\alpha_1, \dots, \alpha_k \geq 0$. |
| konvexní kombinace , | jestliže | $\alpha_1, \dots, \alpha_k \in \mathbb{R}$, | $\alpha_1 + \dots + \alpha_k = 1$, $\alpha_1, \dots, \alpha_k \geq 0$. |

Množina, která je uzavřená vůči

| | | |
|------------------------|-----------|-----------------------------|
| lineárním kombinacím, | se nazývá | lineární podprostor. |
| afinním kombinacím, | se nazývá | afinní podprostor. |
| nezáporným kombinacím, | se nazývá | konvexní kužel. |
| konvexním kombinacím, | se nazývá | konvexní množina. |

K tomu, co již znáte, přibyl pojem *nezáporné kombinace* a konvexního kuželu.

Lineární [afinní, *nezáporný*, konvexní] **obal** vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je množina všech jejich lineárních [afinních, *nezáporných*, konvexních] kombinací. Obecněji, lineární [afinní, *nezáporný*, konvexní] obal množiny $X \subseteq \mathbb{R}^n$ je průnik všech lineárních podprostorů [afinních podprostorů, konvexních kuželů, konvexních množin] obsahující množinu X .

Příklad 12.1. Mějme tři body v \mathbb{R}^3 , které neleží v jedné rovině s počátkem. Jejich lineární obal je celé \mathbb{R}^3 . Jejich afinní obal je rovina jimi procházející. Jejich *nezáporný* obal je nekonečný trojboký hranol, jehož vrchol je v počátku a jehož hrany jsou tři polopřímky určené počátkem a danými body. Jejich konvexní obal je trojúhelník jimi určený. \square

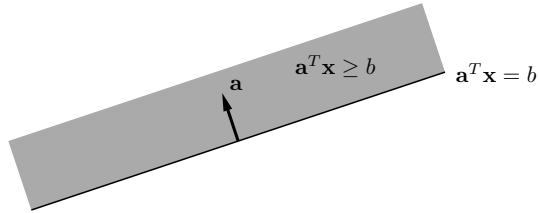
Jako cvičení si nakreslete lineární, afinní, *nezáporný* a konvexní obal náhodně zvolených k vektorů v \mathbb{R}^n pro všech devět případů $k, n \in \{1, 2, 3\}$.

12.3 Konvexní mnohostěny

Poloprostor je množina

$$\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} \geq b \} \quad (12.3)$$

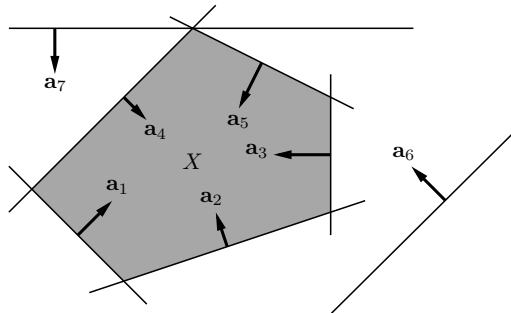
pro nějaké $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$. Jeho hranice je nadrovina $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b \}$. Vektor $\mathbf{a} \in \mathbb{R}^n$ je normála této nadroviny. Obrázek ilustruje tyto pojmy pro $n = 2$:



Konvexní mnohostěn (angl. *polyhedron*) je průnik konečně mnoha poloprostorů. Je to tedy množina

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T \mathbf{x} \geq b_i, i = 1, \dots, m \} = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} \geq \mathbf{b} \}, \quad (12.4)$$

kde $\mathbf{a}_1^T, \dots, \mathbf{a}_m^T \in \mathbb{R}^n$ jsou řádky matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $b_1, \dots, b_m \in \mathbb{R}$ jsou složky vektoru $\mathbf{b} \in \mathbb{R}^m$. Nerovnost (a příslušný poloprostor) $\mathbf{a}_i^T \mathbf{x} \geq b_i$ je **redundantní**, jestliže jejím vypuštěním se množina (12.4) nezmění. Obrázek ukazuje příklad pro $n = 2$ a $m = 7$ (poloprostory 6 a 7 jsou redundantní):



Poloprostor je očividně konvexní množina, proto dle Věty 12.1 je konvexní mnohostěn konvexní množina. Všimněte si, že konvexní mnohostěn nemusí být omezený.

Příklad 12.2. Množina X z Příkladu 11.1 je konvexní mnohostěn. \square

Příklad 12.3. Příklady 'jednoduchých' konvexních mnohostěňů v \mathbb{R}^n :

- prázdná množina \emptyset
- celý prostor \mathbb{R}^n
- každý afinní podprostor (např. bod, přímka, rovina, nadrovina)
- polopřímka $\{ \mathbf{x} + \alpha \mathbf{v} \mid \alpha \geq 0 \}$

- poloprostor
- panel $\{ \mathbf{x} \in \mathbb{R}^n \mid b_1 \leq \mathbf{a}^T \mathbf{x} \leq b_2 \}$
- hyperkrychle $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_\infty \leq 1 \} = \{ \mathbf{x} \in \mathbb{R}^n \mid -1 \leq x_i \leq 1, i = 1, \dots, n \}$
- simplex, což je konvexní obal $n + 1$ afinně nezávislých bodů
- standardní simplex $\{ \mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^n x_i \leq 1 \}$
- pravděpodobnostní simplex $\{ \mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^n x_i = 1 \}$ (množina všech rozdělení pravděpodobnosti diskrétní náhodné proměnné)
- zobecněný osmistěn $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i| \leq 1 \}$. \square

Příklad 12.4. Koule v \mathbb{R}^n pro $n \geq 2$ je průnikem nekonečně mnoha poloprostorů $\mathbf{a}^T \mathbf{x} \leq 1$ pro všechna $\|\mathbf{a}\|_2 = 1$. Je to konvexní množina, ale není to konvexní mnohostěn (protože počet poloprostorů není konečný). \square

12.3.1 Stěny konvexního mnohostěnu

Nechť $X \subseteq \mathbb{R}^n$ je konvexní mnohostěn a necht' \mathbf{a}, b jsou takové, že $X \subseteq \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} \geq b \}$. Pak množina

$$F = X \cap \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b \} \quad (12.5)$$

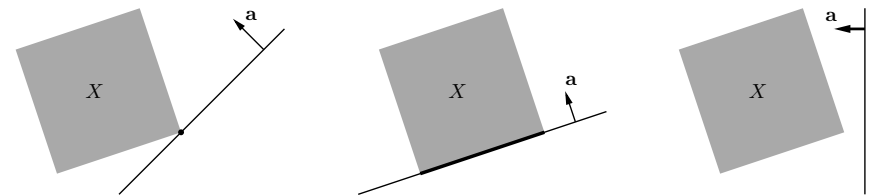
se nazývá **stěna** mnohostěnu X . Z Věty 12.1 plyne, že každá stěna konvexního mnohostěnu je sama o sobě konvexní mnohostěn.

Všimněte si, že každý mnohostěn je sám svoji stěnou, tedy $F = X$ (např. pro $\mathbf{a} = 0$ a $b = 0$). Každý mnohostěn má prázdnou stěnu $F = \emptyset$ (např. pro $\mathbf{a} = 0$ a $b = 1$).

Dimenze stěny je dimenze jejího afinního obalu (zopakujte si pojem afinního obalu z §12.2 a dimenze afinního podprostoru z §3.5). Stěny některých dimenzí mají jméno:

- stěna dimenze 0 se nazývá **vrchol**,
- stěna dimenze 1 se nazývá **hrana**,
- stěna dimenze $n - 1$ se nazývá **faseta** (angl. *facet*, zatímco *face* znamená stěna).

Obrázky ilustrují definici stěny pro $n = 2$. Vlevo je stěna dimenze 0 (vrchol), vpravo stěna dimenze 1 (hrana), vpravo prázdná stěna:



V Příkladu 11.1 je množina F na prvním i druhém obrázku stěna polyhedru X .

Bez důkazu uvedeme ekvivalentní definici stěny, která předpokládá, že mnohostěn je ve tvaru (12.4). Pro $I \subseteq \{1, \dots, m\}$ budeme jako \mathbf{A}_I označovat matici tvořenou řádky $\mathbf{a}_i^T, i \in I$, a jako \mathbf{b}_I budeme označovat vektor tvořený prvky $b_i, i \in I$.

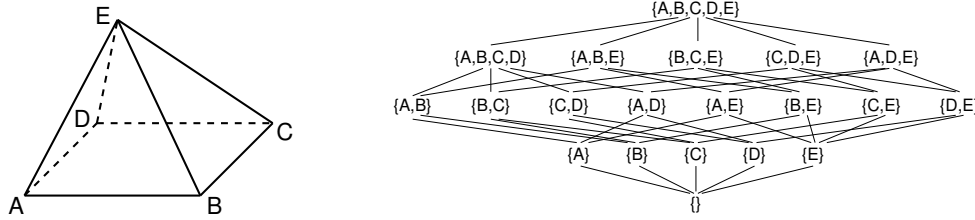
Věta 12.2. $F \subseteq X$ je neprázdná stěna konvexního mnohostěnu (12.4) právě tehdy, když

$$F = \{ \mathbf{x} \in X \mid \mathbf{A}_I \mathbf{x} = \mathbf{b}_I \} \quad (12.6)$$

pro nějakou podmnožinu indexů $I \subseteq \{1, \dots, m\}$.

Stěny konvexního mnohostěnu tvoří množinu částečně uspořádanou množinovou inkluzí, tzv. *stěnový svaz*. Stěnový svaz popisuje 'kombinatorickou strukturu' mnohostěnu.

Příklad 12.5. Necht' mnohostěn (12.4) je pyramida v \mathbb{R}^3 , na levém obrázku:



Tento mnohostěn je průnikem pěti poloprostorů (předpokládáme, že žádný poloprostor není redundantní), tedy $m = 5$ a $n = 3$. Necht' omezení $\mathbf{a}_i^T \mathbf{x} \geq b_i$ pro $i = 1, 2, 3, 4, 5$ je poloprostor, jehož hranicí je polorovina určená po řadě body ABCD, ABE, BCE, CDE, ADE. Pro $I = \{1\}$ je množina (12.6) faceta ABCD. Pro $I = \{1, 2\}$ je množina (12.6) hrana AB. Pro $I = \{1, 2, 3\}$ je množina (12.6) vrchol B.

Obrázek vpravo je diagram¹ znázorňující stěnový svaz. V d -tém řádku diagramu jsou vypsány všechny stěny dimenze d (počítáno zdola). Stěna F dimenze d je spojena čárkou se stěnou G dimenze $d - 1$ jestliže $G \subseteq F$. □

12.3.2 Jak byste vypsali všechny vrcholy konvexního mnohostěnu?

Z Věty 12.2 plyne (po troše přemýšlení), že následující dvě tvrzení jsou ekvivalentní:

- Bod $\mathbf{x} \in \mathbb{R}^n$ je vrchol mnohostěnu (12.4).
- Existuje množina $I \subseteq \{1, \dots, m\}$ tak, že soustava $\mathbf{a}_i^T \mathbf{x} = b_i$, $i \in I$, má právě jedno řešení \mathbf{x} a toto řešení navíc patří do mnohostěnu (12.4), tj. splňuje nerovnice $\mathbf{a}_i^T \mathbf{x} \geq b_i$, i, \dots, m .

Toto nám dovoluje formulovat algoritmus, který vypíše všechny vrcholy mnohostěnu (12.4). Najdeme všechny množiny $I \subseteq \{1, \dots, m\}$ takové, že soustava $\mathbf{a}_i^T \mathbf{x} = b_i$, $i \in I$ má právě jedno řešení a toto řešení navíc splňuje soustavu $\mathbf{a}_i^T \mathbf{x} \geq b_i$, i, \dots, m . Každé takové řešení je vrchol.

12.3.3 Dvě reprezentace konvexního mnohostěnu

Následující věta je hluboká a uvádíme ji bez důkazu. Pro neomezené konvexní mnohostěny platí podobná věta, trochu složitější, kterou neuvádíme.

Věta 12.3. *Konvexní obal konečně mnoha bodů je omezený konvexní mnohostěn. Obráceně, omezený konvexní mnohostěn je konvexním obalem svých vrcholů.*

Máme tedy dvě reprezentace omezeného konvexního mnohostěnu:

¹Přesněji, je to *Hasseho diagram* částečně uspořádané množiny, který máte znát.

- **H-representace:** průnik konečně mnoha poloprostorů ('H' jako *half-space*)
- **V-representace:** konvexní obal konečně mnoha bodů ('V' jako *vertex*)

Přechod od jedné reprezentace ke druhé může být výpočetně těžký či prakticky nemožný. Důvodem je to, že mnohostěn definovaný jako průnik malého počtu poloprostorů může mít velmi velký počet vrcholů (přesně: je-li m počet poloprostorů, tak počet vrcholů nebude omezen žádnou polynomiální funkcí proměnné m). Naopak, mnohostěn s malým počtem vrcholů může mít velký počet facet. V tom případě by algoritmus, který převádí H-representaci na V-representaci nebo naopak, musel vydat velmi dlouhý výstup. V algoritmu v §12.3.2 se to projevuje tím, že musíme projít všechny podmnožiny množiny $\{1, \dots, m\}$, kterých je 2^m .

Příklad 12.6. Uvažujme následující konvexní mnohostěny v \mathbb{R}^n (viz Příklad 12.3):

- Simplex má $n + 1$ vrcholů a $n + 1$ facet.
- Hyperkrychle má $2n$ facet a 2^n vrcholů.
- Zobecněný osmistěn má $2n$ vrcholů a 2^n facet. □

12.4 Cvičení

12.1. Odpovězte, zda následující množiny jsou konvexní a odpověď dokažte z definice konvexní množiny:

- interval $[a, b] \subseteq \mathbb{R}$, kde $a \leq b$
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{C}\mathbf{x} = \mathbf{d} \}$
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{A}\mathbf{x} \leq 1 \}$, kde \mathbf{A} je pozitivně semidefinitní
- \mathbb{Z} (množina celých čísel)
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \max\{x_1, \dots, x_n\} \geq 0 \}$

12.2. Které z následujících množin jsou konvexní? Nemusíte dokazovat z definice, stačí uvést přesvědčivý argument. Množinu si nakreslete pro případ $n = 1$ a $n = 2$.

- $\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1 \}$
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i \geq 1 \}$
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}, \sum_{i=1}^n x_i = 1 \}$
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}, \sum_{i=1}^n x_i \leq 1 \}$
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 = 1 \}$
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 < 1 \}$
- $\{ (x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, xy = 1 \}$
- $\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 2 \} \cap \{ (x, y) \in \mathbb{R}^2 \mid (x - 1)^2 + y^2 \leq 2 \}$

12.3. Bude Věta 12.1 platit, pokud v ní sousloví 'konvexní množina' nahradíme souslovím 'lineární podprostor' (příp. 'afinní podprostor', 'konvexní kužel')? Kladnou i zápornou odpověď dokažte.

12.4. Které z následujících množin jsou konvexní mnohostěny? Pokud je množina konvexní mnohostěn, dokažte ji vyjádřit ve tvaru $\{ \mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b} \}$ (tj. jako průnik poloprostorů)?

- $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}, \sum_i x_i a_i = b, \sum_i x_i a_i^2 = c \}$, kde a_i, b, c jsou dané skaláry

- b) $\{ \mathbf{C}\mathbf{x} \mid \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} = 1 \}$, kde matice \mathbf{C} je dána
 c) $\{ \mathbf{C}\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\|_2 \leq 1 \}$, kde matice \mathbf{C} je dána
 d) $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}\|_2 \leq \|\mathbf{x} - \mathbf{b}\|_2 \}$, kde \mathbf{a}, \mathbf{b} jsou dány

12.5. Mějme konvexní mnohostěn ve tvaru $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \geq \mathbf{b} \}$.

- a) Napište kód v Matlabu, který vypíše vrcholy mnohostěnu pomocí postupu v §12.3.2.
 b) Vypíše vrcholy mnohostěnu pro

$$\mathbf{A} = \begin{bmatrix} -1 & -2 \\ 3 & -1 \\ -1 & 1 \\ 2 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -14 \\ 0 \\ -2 \\ 4 \end{bmatrix}.$$

12.6. Mějme vektory $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$. Pro každé $i = 1, \dots, m$ definujeme množinu

$$X_i = \{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}_i\|_2 \leq \|\mathbf{x} - \mathbf{a}_j\|_2, j \neq i \}.$$

Ukažte, že množiny X_1, \dots, X_m jsou konvexní mnohostěny. Ukažte, že tyto množiny tvoří rozklad (zopakujte si, co je to rozklad množiny) množiny \mathbb{R}^n . Sjednocení hranic těchto množin se nazývá *Voronoiův diagram*. Nakreslete si ho pro $n = 2$ a $m = 4$ pro různé konfigurace bodů $\mathbf{a}_1, \dots, \mathbf{a}_4$.

Nápověda a řešení

- 12.1.a) Konvexní, protože pro libovolné $\alpha \in [0, 1]$ je $\alpha a + (1 - \alpha)b \in [a, b]$.
 12.1.b) Konvexní.
 12.1.c) Konvexní.
 12.1.d) Nekonvexní. Např. pro $x = 1, y = 2, \alpha = \frac{1}{2}$ číslo $\alpha x + (1 - \alpha)y = 1.5$ není celé.
 12.1.e) Nekonvexní.
 12.2.a) nadrovina, konvexní
 12.2.b) poloprostor, konvexní
 12.2.c) průnik poloprostorů a nadroviny, konvexní mnohostěn
 12.2.d) průnik poloprostorů, konvexní
 12.2.e) sféra, není konvexní
 12.2.f) koule bez hranice, konvexní
 12.2.g) graf jedné větve hyperboly, není konvexní
 12.2.h) průnik dvou koulí, konvexní
 12.5.b) Zde jsou vrcholy a odpovídající množiny $I \subseteq \{1, \dots, 4\}$:

| | | | | |
|--------------|------------|------------|-------------------------------|------------|
| I | $\{1, 2\}$ | $\{1, 3\}$ | $\{2, 4\}$ | $\{3, 4\}$ |
| \mathbf{x} | $(2, 6)$ | $(6, 4)$ | $(\frac{4}{3}, \frac{12}{5})$ | $(2, 0)$ |

Kapitola 13

Simplexová metoda

Zde popíšeme algoritmus na řešení úloh lineárního programování zvaný **simplexová metoda**.

Zapomeňme prozatím na účelovou funkci a zkoumejme množinu přípustných řešení LP ve tvaru

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}, \quad (13.1)$$

kde $\mathbf{A} \in \mathbb{R}^{m \times n}$ je široká ($m < n$) matice s hodnotami m , tedy její řádky jsou lineárně nezávislé.

Soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ má nekonečně mnoho řešení. Položíme-li však $n - m$ složek vektoru \mathbf{x} rovno nule (tedy učiníme-li $n - m$ z podmínek $\mathbf{x} \geq \mathbf{0}$ aktivních), soustava má nejvýše jedno řešení. Tato úvaha vede k následujícím definicím:

- Množina $J \subseteq \{1, 2, \dots, n\}$ se nazývá **báze** úlohy, pokud $|J| = m$ a sloupce matice \mathbf{A} s indexy J jsou lineárně nezávislé. Tedy sloupce J tvoří regulární matici $m \times m$.
- Vektor \mathbf{x} je **bázové řešení** příslušné bázi J , pokud $\mathbf{A}\mathbf{x} = \mathbf{b}$ a $x_j = 0$ pro $j \notin J$.
- Bázové řešení \mathbf{x} je **přípustné**, pokud $\mathbf{x} \geq \mathbf{0}$.
- Bázové řešení \mathbf{x} je **degenerované**, pokud má méně než m nenulových složek.
- Dvě báze jsou **sousední**, pokud mají $m - 1$ společných prvků.

Protože matice \mathbf{A} má hodnotu m , existuje aspoň jedna báze a každé bázi přísluší právě jedno bázové řešení. Bázové řešení však může příslušet více než jedné bázi, což se stane právě tehdy, když je toto bázové řešení degenerované.

Příklad 13.1. Necht' je soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ dána tabulkou (blokovou maticí)

$$[\mathbf{A} \ \mathbf{b}] = \begin{bmatrix} -1 & 1 & 3 & 1 & 0 & 2 & 1 \\ 1 & 0 & 4 & 0 & 1 & 4 & 4 \\ -1 & 0 & 4 & 1 & 1 & 4 & 2 \end{bmatrix}. \quad (13.2)$$

- $J = \{2, 3, 5\}$ není báze, protože sloupce 2, 3, 5 matice \mathbf{A} jsou lineárně závislé.
- $J = \{1, 4, 5\}$ je báze, protože tyto sloupce jsou lineárně nezávislé. Bázové řešení $\mathbf{x} = (x_1, x_2, \dots, x_6)$ příslušné bázi J se najde řešením soustavy

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} \quad (13.3)$$

a položením $x_2 = x_3 = x_6 = 0$. Soustava (13.3) má právě jedno řešení, neboť její matice je regulární. Dostaneme $\mathbf{x} = (3, 0, 0, 4, 1, 0)$. Toto bázové řešení je přípustné. Není degenerované, protože má $m = 3$ nenulových složek.

- $J = \{1, 2, 4\}$ je báze. Bázové řešení je $\mathbf{x} = (4, -1, 0, 6, 0, 0)$. Je nepřipustné, protože $x_2 < 0$.
- $J = \{3, 4, 5\}$ je báze. Bázové řešení $\mathbf{x} = (0, 0, 1, -2, 0, 0)$ je degenerované, protože má méně než $m = 3$ nenulových složek.
- Stejně bázové řešení $\mathbf{x} = (0, 0, 1, -2, 0, 0)$ dostaneme volbou báze $J = \{3, 4, 6\}$. Vidíme, že degenerované bázové řešení odpovídá více než jedné bázi.
- Báze $\{1, 4, 5\}$ a $\{2, 4, 5\}$ jsou sousední, protože mají společné dva prvky $\{4, 5\}$. Báze $\{1, 4, 5\}$ a $\{2, 4, 6\}$ nejsou sousední. \square

Mnohostěn (13.1) je speciální tvar mnohostěnu (12.4). Dle definice mnohostěnu se všechna minima lineární funkce na konvexním mnohostěnu (13.1) nabývají na nějaké stěně tohoto mnohostěnu. Pokud tato stěna obsahuje alespoň jeden vrchol mnohostěnu (to nebude platit jen pro velmi jednoduché mnohostěny (13.1), zamyslete se nad tím!), pak se alespoň jedno minimum nabývá i v tomto vrcholu.

Z Věty 12.2 plyne, že přípustná bázová řešení jsou vrcholy mnohostěnu (13.1). Dále lze ukázat (důkaz vynecháme), že dvojice sousedních bází odpovídají buď jedinému (degenerovanému) vrcholu nebo dvojici vrcholů spojených hranou.

To nám dovoluje navrhnout naivní algoritmus na řešení LP: uděláme výčet všech přípustných bázových řešení a nalezneme to s nejlepší hodnotou účelové funkce. Tento algoritmus není praktický, protože bázových řešení je exponenciálně mnoho. Simplexová metoda je vylepšení tohoto přístupu: přechází mezi sousedními bázemi tak, že bázová řešení jsou stále přípustná (tedy přechází po hranách mnohostěnu) a účelová funkce se zlepšuje (nebo aspoň nezhoršuje).

13.1 Stavební kameny algoritmu

Zde vysvětlíme jednotlivé stavební kameny simplexové metody, které nakonec v §13.2 spojíme v celý algoritmus.

13.1.1 Přejít od aktuální standardní bázi

Simplexová metoda pracuje pouze se *standardními* bázemi, tj. sloupce J jsou (permutované) vektory standardní báze. To má výhodu v tom, že (i) nemusíme kontrolovat, zda jsou sloupce J lineárně nezávislé a (ii) nenulové složky bázového řešení \mathbf{x} jsou rovny přímo složkám vektoru \mathbf{b} . Na počátku algoritmu se předpokládá, že matice \mathbf{A} obsahuje standardní bázi.

Z lineární algebry známe *ekvivalentní řádkové úpravy* soustavy $\mathbf{Ax} = \mathbf{b}$: libovolný řádek tabulky $[\mathbf{A} \ \mathbf{b}]$ můžeme vynásobit nenulovým číslem a můžeme k němu přičíst lineární kombinaci ostatních řádků. Tyto úpravy nemění množinu řešení soustavy.

Ukážeme, jak přejít od aktuální standardní báze J k sousední standardní bázi, tedy nějaký sloupec $j' \in J$ bázi opustí a nějaký sloupec $j \notin J$ do báze vstoupí. Necht' i je řádek, ve kterém je $a_{ij'} = 1$. Prvek (i, j) matice se nazývá **pivot** (angl. znamená *čep*). Necht' $a_{ij} \neq 0$. Chceme nastavit pivot a_{ij} na jedničku, vynulovat prvky nad i pod pivotem, a nezměnit přitom sloupce $J \setminus \{j'\}$. Toho se dosáhne těmito ekvivalentními řádkovými úpravami:

1. Vyděl řádek i číslem a_{ij} .

2. Pro každé $i' \neq i$ odečti $a_{i'j}$ -násobek řádku i od řádku i' .

Ríkáme, že jsme provedli *ekvivalentní úpravu kolem pivotu* s indexy (i, j) .

Příklad 13.2. Mějme soustavu

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccccc|c} 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & \boxed{1} & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \end{array} \right] \quad (13.4)$$

se (standardní) bází $J = \{1, 4, 5\}$. Vidíme ihned odpovídající bázové řešení, $\mathbf{x} = (3, 0, 0, 4, 1, 0)$.

Chceme nahradit bázový sloupec $j' = 1$ nebázovým sloupcem $j = 2$, tedy přejít k sousední bázi $\{2, 4, 5\}$. Máme $i = 2$, tedy pivot je prvek a_{22} (v tabulce orámován). Ekvivalentními řádkovými úpravami musíme docílit, aby pivot byl roven jedné a prvky nad ním a pod ním byly nulové. Při tom smíme změnit sloupec 1, ale sloupce 4 a 5 se změnit nesmějí. Toho se docílí vydělením řádku 2 číslem a_{22} (což zde nemá žádný efekt, protože náhodou $a_{22} = 1$) a pak přičtením vhodných násobků řádku 2 k ostatním řádkům. Výsledek:

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccccc|c} -2 & 0 & 0 & 1 & 0 & 0 & -2 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 1 & 0 & 4 & 0 & 1 & 4 & 4 \end{array} \right].$$

Nyní sloupce $\{2, 4, 5\}$ tvoří standardní bázi. \square

13.1.2 Kdy je sousední bázové řešení přípustné?

Uvedeným způsobem můžeme od aktuální báze přejít k libovolné sousední bázi. Přitom nové bázové řešení může nebo nemusí být přípustné. Je-li aktuální bázové řešení přípustné, jak poznáme, zda i nové bázové řešení bude přípustné?

Protože nenulové složky bázového řešení \mathbf{x} jsou rovny složkám vektoru \mathbf{b} , bázové řešení je přípustné právě tehdy, když $\mathbf{b} \geq \mathbf{0}$. Necht' v aktuální tabulce je $\mathbf{b} \geq \mathbf{0}$. Provedme ekvivalentní úpravu kolem pivotu (i, j) . Hledáme podmínky na (i, j) , za kterých bude i po úpravě $\mathbf{b} \geq \mathbf{0}$.

Po ekvivalentní úpravě kolem pivotu (i, j) se vektor \mathbf{b} změní takto (viz §13.1.1):

- b_i se změní na b_i/a_{ij} ,
- pro každé $i' \neq i$ se $b_{i'}$ změní na $b_{i'} - a_{i'j}b_i/a_{ij}$.

Tato čísla musejí být nezáporná. To nastane právě tehdy, když platí následující podmínky:

$$a_{ij} > 0, \quad (13.5a)$$

$$\text{pro každé } i' \neq i \text{ platí } a_{i'j} \leq 0 \text{ nebo } \frac{b_i}{a_{ij}} \leq \frac{b_{i'}}{a_{i'j}}, \quad (13.5b)$$

kde 'nebo' je užito v nevylučovacím smyslu. Podmínka (13.5a) je zřejmá. Podmínka (13.5b) je ekvivalentní podmínce $b_{i'} - a_{i'j}b_i/a_{ij} \geq 0$, neboť $a_{ij} > 0$, $b_i \geq 0$, $b_{i'} \geq 0$ (rozmyslete!).

Příklad 13.3. Uvažujme opět tabulku (13.4).

- Ekvivalentní úprava okolo pivotu $(i, j) = (3, 2)$ nepovede k přípustnému bázovému řešení, neboť $a_{ij} = -1 < 0$, což porušuje podmínku (13.5a).
- Ekvivalentní úprava okolo pivotu $(i, j) = (2, 2)$ nepovede k přípustnému bázovému řešení, neboť pro $i' = 1$ je $a_{i'j} > 0$ a $\frac{3}{1} > \frac{4}{2}$, tedy podmínka (13.5b) je porušena.
- Ekvivalentní úprava okolo pivotu $(i, j) = (3, 6)$ povede k přípustnému bázovému řešení. Podmínky (13.5) jsou splněny, neboť $a_{ij} = 2 > 0$ a $\frac{1}{2} \leq \frac{4}{4}$, $\frac{1}{2} \leq \frac{3}{2}$. \square

13.1.3 Co když je celý sloupec nekladný?

Jestliže jsou všechny prvky v nějakém nebázovém sloupci j nekladné, víme z podmínky (13.5a), že tento sloupec se nemůže stát bázovým. Platí ale navíc, že souřadnice x_j bodu \mathbf{x} se může libovolně zvětšovat a bod \mathbf{x} přesto zůstane v mnohostěnu X . Tedy existuje polopřímka s počátkem v \mathbf{x} ležící celá v mnohostěnu X . Tedy mnohostěn X je neomezený.

Příklad 13.4. Necht' $[\mathbf{A} \ \mathbf{b}]$ je tabulka

$$\begin{array}{cccccc|c} 0 & -2 & 6 & 1 & 0 & 4 & 4 \\ 1 & -1 & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \\ \hline \mathbf{x} = & 3 & 0 & 0 & 4 & 1 & 0 \end{array}$$

s bází $\{1, 4, 5\}$. Pod tabulkou je napsáno bázové řešení \mathbf{x} . Když se x_2 bude libovolně zvětšovat, změnu lze kompenzovat současným zvětšováním bázových proměnných x_1, x_4, x_5 tak, že vektor \mathbf{Ax} zůstane nezměněn a tedy roven \mathbf{b} . Konkrétně, vektor pro každé $\alpha \geq 0$ bude vektor $\mathbf{x} = (3, 0, 0, 4, 1, 0) + \alpha(1, 1, 0, 2, 1, 0)$ splňovat $\mathbf{Ax} = \mathbf{b}$ a $\mathbf{x} \geq \mathbf{0}$. \square

13.1.4 Ekvivalentní úpravy účelového řádku

Dosud jsme prováděli ekvivalentní řádkové úpravy pouze na soustavě $\mathbf{Ax} = \mathbf{b}$ a účelové funkce si nevšimli. Tyto úpravy lze rozšířit na celou úlohu LP včetně účelové funkce. Nebudeme účelovou funkcí uvažovat ve tvaru $\mathbf{c}^T \mathbf{x}$, ale v mírně obecnějším tvaru $\mathbf{c}^T \mathbf{x} - d$. Tedy řešíme LP

$$\min\{\mathbf{c}^T \mathbf{x} - d \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}. \quad (13.6)$$

Úlohu budeme reprezentovat **simplexovou tabulkou**

$$\begin{bmatrix} \mathbf{c}^T & d \\ \mathbf{A} & \mathbf{b} \end{bmatrix}. \quad (13.7)$$

Přičteme k účelovému řádku $[\mathbf{c}^T \ d]$ libovolnou lineární kombinaci $\mathbf{y}^T [\mathbf{A} \ \mathbf{b}]$ ostatních řádků $[\mathbf{A} \ \mathbf{b}]$, kde \mathbf{y} jsou koeficienty lineární kombinace. Ukážeme, že tato úprava zachová hodnotu účelové funkce $\mathbf{c}^T \mathbf{x} - d$ pro každé \mathbf{x} splňující $\mathbf{Ax} = \mathbf{b}$. Nový účelový řádek bude

$$[\mathbf{c}^T \ d] + \mathbf{y}^T [\mathbf{A} \ \mathbf{b}] = [\mathbf{c}^T + \mathbf{y}^T \mathbf{A} \ d + \mathbf{y}^T \mathbf{b}].$$

Nová účelová funkce bude tedy

$$(\mathbf{c}^T + \mathbf{y}^T \mathbf{A})\mathbf{x} - (d + \mathbf{y}^T \mathbf{b}) = \mathbf{c}^T \mathbf{x} - d + \mathbf{y}^T (\mathbf{Ax} - \mathbf{b}).$$

Ale to je rovno $\mathbf{c}^T \mathbf{x} - d$ pro každé \mathbf{x} splňující $\mathbf{Ax} = \mathbf{b}$.

13.1.5 Co udělá přechod k sousední bázi s účelovou funkcí?

Necht' J je standardní báze. Přičteme k účelovému řádku takovou lineární kombinaci ostatních řádků, aby pro všechna $j \in J$ bylo $c_j = 0$ (novému vektoru \mathbf{c} se pak říká *reduované ceny*). Protože bázové řešení \mathbf{x} je v nebázových sloupcích nulové, znamená to $\mathbf{c}^T \mathbf{x} = 0$. Tedy hodnota účelové funkce $\mathbf{c}^T \mathbf{x} - d$ v bázovém řešení \mathbf{x} je rovna jednoduše $-d$.

Navíc je snadno vidět, co udělá přechod k nové bázi s účelovou funkcí. Necht' j' je sloupec opouštějící bázi a j je sloupec vstupující do báze. Při přechodu k nové bázi se číslo $x_{j'}$ stane nulovým a číslo x_j se zvětší z nuly na kladné (nebo se nezmění). Protože $c_{j'} = 0$, číslo $\mathbf{c}^T \mathbf{x} - d$ při $c_j \geq 0$ stoupne (nebo se nezmění) a při $c_j \leq 0$ klesne (nebo se nezmění).

Příklad 13.5. Mějme úlohu se simplexovou tabulkou

$$\begin{bmatrix} \mathbf{c}^T & d \\ \mathbf{A} & \mathbf{b} \end{bmatrix} = \begin{array}{cccccc|c} 1 & -2 & -3 & -1 & 2 & 1 & 4 \\ 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \end{array},$$

kde $J = \{1, 4, 5\}$. Složky vektoru \mathbf{c} v bázových sloupcích vynulujeme tak, že k účelovému řádku přičteme první řádek, odečteme druhý řádek, a odečteme dvojnásobek třetího řádku:

$$\begin{array}{cccccc|c} 0 & 1 & -2 & 0 & 0 & -1 & 3 \\ 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \\ \hline \mathbf{x} = & 3 & 0 & 0 & 4 & 1 & 0 \end{array},$$

Pod tabulku jsme napsali bázové řešení \mathbf{x} . Nyní je $\mathbf{c}^T \mathbf{x} = 0$, a tedy hodnota účelové funkce v bázovém řešení je $\mathbf{c}^T \mathbf{x} - d = -d = -3$.

Dejme tomu, že chceme přidat do báze nebázový sloupec 2 a vyloučit z ní některý z bázových sloupců $\{1, 4, 5\}$. Po tomto přechodu se x_2 stane kladné nebo zůstane nulové a jedna ze složek x_1, x_4, x_5 se vynuluje. Protože $c_1 = c_4 = c_5 = 0$, změna x_1, x_4, x_5 se na účelové funkci neprojeví a ta se změní o $c_2 x_2$. Kritériem tedy stoupne nebo zůstane stejné, protože $c_2 = 1 > 0$. \square

Pokud v některém sloupci j je $c_j \leq 0$ a $a_{ij} \leq 0$ pro všechna i , pak můžeme proměnnou x_j libovolně zvětšovat (viz §13.1.3) a účelovou funkci libovolně zmenšovat. Úloha je tedy neomezená.

13.2 Základní algoritmus

Spojením popsaných stavebních kamenů dostaneme iteraci simplexového algoritmu na řešení úlohy (13.6). Iterace přejde k sousední standardní bázi takové, že bázové řešení zůstane přípustné a účelová funkce se zmenší nebo alespoň nezmění. Vstupem i výstupem iterace je simplexová tabulka (13.7) s těmito vlastnostmi:

- podmnožina sloupců \mathbf{A} tvoří standardní bázi J ,
- bázové řešení odpovídající této bázi je přípustné, $\mathbf{b} \geq \mathbf{0}$,
- složky vektoru \mathbf{c} v bázových sloupcích jsou nulové, $c_j = 0$ pro $j \in J$.

Iterace se provede v těchto krocích:

1. Vyber index j pivotu tak, aby $c_j < 0$ (§13.1.5).
2. Vyber index i pivotu podle podmínek (13.5). Z těchto podmínek plyne (promyslete!)

$$i \in \operatorname{argmin}_{i' | a_{i'j} > 0} \frac{b_{i'}}{a_{i'j}}, \quad (13.8)$$

kde $\operatorname{argmin}_{i' | a_{i'j} > 0}$ označuje, že se minimalizuje přes všechna i' splňující $a_{i'j} > 0$.

- Udělej ekvivalentní úpravu okolo pivotu (i, j) (§13.1.1).
- Udělej ekvivalentní úpravu účelového řádku, která vynuluje c_j v novém báзовém sloupci j (§13.1.5).

Algoritmus, který opakuje uvedenou iteraci, nazveme **základní simplexový algoritmus**. Algoritmus končí, když už nelze iteraci provést. To nastane z jednoho z těchto důvodů:

- Všechny koeficienty c_j jsou nezáporné. Účelovou funkci nelze zlepšit a jsme v optimu.
- V některém sloupci j je $c_j < 0$ a $a_{ij} \leq 0$ pro všechna i . Úloha je neomezená.

Výběr indexů (i, j) pivotu v krocích 1 a 2 nemusí být jednoznačný, tedy může být více sloupců j s vhodným znaménkem c_j a více řádků i může splňovat podmínky (13.5) (tedy může být více argumentů minima v podmínce (13.8)). Algoritmus, který vybírá jediný pivot z těchto možností, se nazývá **pivotové pravidlo**.

Zřídka se algoritmus může dostat do stavu, kdy cyklicky prochází stále stejnou množinou bází, které odpovídají jedinému degenerovanému báзовému řešení a tedy účelová funkce se nemění. Tomuto problému **cyklení** se dá zabránit použitím vhodného pivotového pravidla (nejznámější je *Blandovo anticyklické pravidlo*), které ale popisovat nebudeme.

Příklad 13.6. Vyřešte lineární program (13.6) simplexovou metodou, když výchozí simplexová tabulka (13.7) je

$$\begin{array}{cccccc|c} 0 & -2 & 1 & 0 & 0 & -3 & 0 \\ 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & \boxed{2} & 1 \end{array}$$

Báze je $J = \{1, 4, 5\}$ a báзовé řešení $\mathbf{x} = (3, 0, 0, 4, 1, 0)$.

Účelový řádek budeme nazývat nultý, ostatní pak první, druhý atd. První iterace simplexového algoritmu se provede v těchto krocích:

- Vybereme sloupec j , který vstoupí do báze. To může být libovolný sloupec, který má v nultém řádku záporné číslo. Je rozumné vzít nejmenší takové číslo, zde -3 , tedy $j = 6$.
- Vybereme řádek i pivotu dle (13.8) nalezením argumentu minima z čísel $\frac{4}{4}, \frac{3}{2}, \frac{1}{2}$. Bude tedy $i = 3$. Výsledný pivot je označen rámečkem. Všimněte si, že řádek $i = 3$ má v aktuální bázi jedničku ve sloupci 5, sloupec 5 tedy bázi opustí.
- Uděláme ekvivalentní úpravu okolo pivotu (i, j) a zároveň vynulujeme číslo c_j . Neboli chceme, aby se z pivotu a_{ij} stala jednička a nad i pod pivotem byly nuly, a to včetně nultého řádku. Tedy nejprve třetí řádek vydělíme dvěma a potom k nultému řádku přičteme trojnásobek třetího řádku, od prvního řádku odečteme čtyřnásobek třetího řádku, a od druhého řádku odečteme dvojnásobek třetího řádku. Všimněte si: k žádnému řádku nikdy nepřičítáme násobky jiného řádku než pivotového. Výsledek:

$$\begin{array}{cccccc|c} 0 & -3.5 & 2.5 & 0 & 1.5 & 0 & 1.5 \\ 0 & \boxed{4} & 4 & 1 & -2 & 0 & 2 \\ 1 & 2 & 2 & 0 & -1 & 0 & 2 \\ 0 & -0.5 & 0.5 & 0 & 0.5 & 1 & 0.5 \end{array}$$

Na konci první iterace máme bázi $J = \{1, 4, 6\}$, báзовé řešení $\mathbf{x} = (2, 0, 0, 2, 0, 0.5)$, a hodnotu účelové funkce $-d = -1.5$.

Druhá iterace: pivot je ve sloupci $j = 2$. Jeho řádek najdeme dle (13.8) porovnáním čísel $\frac{2}{4}, \frac{2}{2}$, tedy $i = 1$. Výsledek druhé iterace:

$$\begin{array}{cccccc|c} 0 & 0 & 6 & 0.875 & -0.25 & 0 & 3.25 \\ 0 & 1 & 1 & 0.25 & -0.5 & 0 & 0.5 \\ 1 & 0 & 0 & -0.5 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0.125 & \boxed{0.25} & 1 & 0.75 \end{array}$$

Výsledek třetí iterace:

$$\begin{array}{cccccc|c} 0 & 0 & 7 & 1 & 0 & 1 & 4 \\ 0 & 1 & 3 & 0.5 & 0 & 2 & 2 \\ 1 & 0 & 0 & -0.5 & 0 & 0 & 1 \\ 0 & 0 & 4 & 0.5 & 1 & 4 & 3 \end{array}$$

Protože všechna čísla v účelovém řádku jsou nezáporná, algoritmus končí. Úloha má optimální řešení s hodnotou -4 v bodě $(x_1, x_2, x_3, x_4, x_5, x_6) = (1, 2, 0, 0, 3, 0)$. \square

Příklad 13.7. Necht' simplexová tabulka (13.7) je

$$\begin{array}{cccccc|c} -2 & 6 & 1 & 0 & 0 & 0 \\ -1 & -1 & -1 & 1 & 0 & 2 \\ \boxed{2} & -1 & -2 & 0 & 1 & 1 \end{array}$$

Tabulka po první iteraci je

$$\begin{array}{cccccc|c} 0 & 5 & -1 & 0 & 1 & 1 \\ 0 & -1.5 & -2 & 1 & 0.5 & 2.5 \\ 1 & -0.5 & -1 & 0 & 0.5 & 0.5 \end{array}$$

Podle nultého řádku by další pivot měl být ve třetím sloupci. Ale čísla a_{i3} jsou všechna záporná (viz §13.1.3). Tedy úloha je neomezená. V nové tabulce je vidět, že můžeme zvětšovat x_3 libovolně a kompenzovat to vhodným nárůstem x_1 a x_4 . Jelikož $c_1 = c_4 = 0$, změny x_1 a x_4 se na účelové funkci neprojeví a jediný vliv na ní bude mít x_3 , které ho bude libovolně zmenšovat. \square

13.3 Inicializace algoritmu

Na začátku základního simplexového algoritmu musí být úloha zadána ve tvaru

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}, \quad (13.9)$$

kde matice \mathbf{A} obsahuje standardní bázi a $\mathbf{b} \geq \mathbf{0}$. Ukážeme, jak lze obecnou úlohu LP převést na tento tvar.

Někdy je převod snadný. Pokud má úloha tvar $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ a platí $\mathbf{b} \geq \mathbf{0}$, přidáním slacků úlohu převedeme na $\min\{\mathbf{c}^T \mathbf{u} \mid \mathbf{Ax} + \mathbf{u} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}, \mathbf{u} \geq \mathbf{0}\}$. Tato úloha simplexovou tabulkou

$$\begin{bmatrix} \mathbf{c}^T & \mathbf{0} & \mathbf{0} \\ \mathbf{A} & \mathbf{I} & \mathbf{b} \end{bmatrix},$$

ve které sloupce příslušné proměnným \mathbf{u} tvoří standardní bázi.

Příklad 13.8. Vyřešte simplexovým algoritmem:

$$\begin{aligned} \min \quad & -3x_1 - x_2 - 3x_3 \\ \text{za podmíněk} \quad & 2x_1 + x_2 + x_3 \leq 2 \\ & x_1 + 2x_2 + 3x_3 \leq 5 \\ & 2x_1 + 2x_2 + x_3 \leq 6 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

Přidáme slackové proměnné $u_1, u_2, u_3 \geq 0$, abychom omezení uvedli do tvaru rovností:

$$\begin{aligned} \min \quad & -3x_1 - x_2 - 3x_3 \\ \text{za podmíněk} \quad & 2x_1 + x_2 + x_3 + u_1 = 2 \\ & x_1 + 2x_2 + 3x_3 + u_2 = 5 \\ & 2x_1 + 2x_2 + x_3 + u_3 = 6 \\ & x_1, x_2, x_3, u_1, u_2, u_3 \geq 0 \end{aligned}$$

Zde je výchozí simplexová tabulka:

$$\begin{array}{cccccc|c} -3 & -1 & -3 & 0 & 0 & 0 & 0 \\ \hline 2 & 1 & 1 & 1 & 0 & 0 & 2 \\ 1 & 2 & 3 & 0 & 1 & 0 & 5 \\ 2 & 2 & 1 & 0 & 0 & 1 & 6 \end{array}$$

□

13.3.1 Dvofázová simplexová metoda

Pokud je úloha zadána v obecném tvaru, operacemi z §11.1 ji lze vždy převést do tvaru (13.9). Vynásobením vhodných řádků záporným číslem vždy zajistíme $\mathbf{b} \geq \mathbf{0}$, matice \mathbf{A} ale nemusí obsahovat standardní bázi. Máme dokonce vážnější problém: není vůbec jasné, zda úloha (13.9) je přípustná. V tomto případě nejdříve vyřešíme *pomocnou úlohu* LP, která najde *nějaké* (ne nutně optimální) přípustné řešení. Z něj pak získáme standardní bázi. Pomocná úloha je

$$\min \{ \mathbf{1}^T \mathbf{u} \mid \mathbf{A} \mathbf{x} + \mathbf{u} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}, \mathbf{u} \geq \mathbf{0} \} \quad (13.10)$$

a má simplexovou tabulku

$$\begin{bmatrix} \mathbf{0} & \mathbf{1}^T & \mathbf{0} \\ \mathbf{A} & \mathbf{I} & \mathbf{b} \end{bmatrix}.$$

Pro libovolné $\mathbf{u} \geq \mathbf{0}$ je $\mathbf{1}^T \mathbf{u} \geq 0$, přičemž $\mathbf{1}^T \mathbf{u} = 0$ právě tehdy, když $\mathbf{u} = \mathbf{0}$. Tedy úloha (13.9) je přípustná právě tehdy, je-li optimální hodnota úlohy (13.10) rovna 0. Na počátku tvoří sloupce příslušné proměnným \mathbf{u} standardní bázi, lze tedy na ní pustit základní simplexový algoritmus. Ten může skončit dvěma způsoby:

- Pokud je optimum větší než 0, pak úloha (13.9) je nepřipustná.
- Pokud je optimum rovno 0, pak úloha (13.9) je přípustná. Pokud není optimální řešení (\mathbf{x}, \mathbf{u}) úlohy (13.10) degenerované, po skončení simplexového algoritmu jsou všechny bazové proměnné kladné. Protože $\mathbf{u} = \mathbf{0}$, proměnné \mathbf{u} budou tedy nebázové. Proto mezi sloupce příslušnými proměnným \mathbf{x} bude existovat standardní báze.

Pokud je optimální řešení (\mathbf{x}, \mathbf{u}) úlohy (13.10) degenerované, některé proměnné \mathbf{u} mohou být na konci algoritmu bazové. Pak je nutno udělat dodatečné úpravy kolem pivotů ve sloupcích příslušných bazovým proměnným \mathbf{u} , abychom tyto sloupce dostali z báze ven.

Nalezení nějakého přípustného řešení v pomocné úloze (13.10) se nazývá **první fáze** a řešení původní úlohy pak **druhá fáze** algoritmu, mluvíme tedy o **dvofázové simplexové metodě**.

Příklad 13.9. Řešte

$$\begin{aligned} \min \quad & -20x_1 - 30x_2 - 40x_3 \\ \text{za podmíněk} \quad & 3x_1 + 2x_2 + x_3 = 10 \\ & x_1 + 2x_2 + 2x_3 = 15 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

Máme sice $\mathbf{b} \geq \mathbf{0}$, ale není jasné, zda existuje přípustné \mathbf{x} , tím méně není vidět standardní báze. Provedeme první fázi algoritmu. Pomocná úloha bude

$$\begin{aligned} \min \quad & u_1 + u_2 \\ \text{za podmíněk} \quad & 3x_1 + 2x_2 + x_3 + u_1 = 10 \\ & x_1 + 2x_2 + 2x_3 + u_2 = 15 \\ & x_1, x_2, x_3, u_1, u_2 \geq 0 \end{aligned}$$

s tabulkou

$$\begin{array}{cccccc|c} 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ \hline 3 & 2 & 1 & 1 & 0 & 0 & 10 \\ 1 & 2 & 2 & 0 & 1 & 0 & 15 \end{array}$$

Sloupce nad přidanými proměnnými tvoří standardní bázi, můžeme tedy na pomocnou úlohu pustit základní simplexový algoritmus. Po vynulování ceny nad bazovými proměnnými budou kroky algoritmu vypadat takto:

$$\begin{array}{cccccc|c} -4 & -4 & -3 & 0 & 0 & 0 & -25 \\ \hline 3 & \boxed{2} & 1 & 1 & 0 & 0 & 10 \\ 1 & 2 & 2 & 0 & 1 & 0 & 15 \\ \hline 2 & 0 & -1 & 2 & 0 & 0 & -5 \\ 1.5 & 1 & 0.5 & 0.5 & 0 & 0 & 5 \\ -2 & 0 & \boxed{1} & -1 & 1 & 0 & 5 \\ \hline 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 2.5 & 1 & 0 & 1 & -0.5 & 0 & 2.5 \\ -2 & 0 & 1 & -1 & 1 & 0 & 5 \end{array}$$

Optimum je rovno 0, tedy původní úloha je přípustná. Proměnné u_1, u_2 jsou nebázové a tedy rovny nule, bazové proměnné jsou x_2, x_3 . Ted' tedy můžeme začít druhou fází (řešení původní úlohy) s počáteční tabulkou

$$\begin{array}{cccc|c} -20 & -30 & -40 & 0 & 0 \\ \hline 2.5 & 1 & 0 & 0 & 2.5 \\ -2 & 0 & 1 & 0 & 5 \end{array}$$

□

13.4 Cvičení

13.1. Najděte všechny báze a básová řešení mnohostěnu $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ pro

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 2 & -1 & 4 & 1 & 4 \end{array} \right].$$

Která z nich jsou přípustná? Která z nich jsou degenerovaná?

13.2. V tabulce zakroužkujte všechny pivoty takové, že ekvivalentní úprava kolem nich povede k přípustnému básovému řešení:

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccccc|c} 0 & 2 & 6 & 1 & 0 & -4 & 3 & 0 & 4 \\ 1 & 1 & -3 & 0 & 0 & 2 & 3 & 0 & 3 \\ 0 & -1 & 1 & 0 & 1 & -2 & -3 & 0 & 1 \\ 0 & -2 & 2 & 0 & 0 & 2 & -1 & 1 & 1 \end{array} \right]$$

13.3. Zapište lineární program

$$\begin{array}{ll} \min & -x_1 \quad - \quad x_4 - 3x_5 \\ \text{za podmínek} & 2x_1 \quad + \quad x_4 + x_5 + x_6 = 2 \\ & -x_1 + x_2 \quad + \quad 2x_4 + 3x_5 = 4 \\ & 2x_1 \quad + \quad x_3 + 2x_4 - x_5 = 6 \\ & x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{array}$$

do simplexové tabulky. Předpokládejte, že aktuální báze je $\{2, 3, 6\}$. Jaké je aktuální básové řešení? Je toto básové řešení přípustné. Je degenerované? Pokud je to možné, udělejte jeden krok simplexového algoritmu. Pokud to možné není, vysvětlete proč.

13.4. Vyřešte simplexovou metodou:

$$\begin{array}{ll} \max & 2x_1 - x_2 - 3x_3 \\ \text{za podmínek} & -2x_1 - x_2 + x_3 \leq 2 \\ & -x_1 + 2x_2 - 3x_3 \leq 5 \\ & -2x_1 - 4x_2 + x_3 \leq 6 \\ & x_1, x_2, x_3 \geq 0 \end{array}$$

13.5. Vyřešte simplexovou metodou (navzdory tomu, že lze řešit úvahou):

$$\begin{array}{ll} \max & 6x_1 + 9x_2 + 5x_3 + 9x_4 \\ \text{za podmínek} & x_1 + x_2 + x_3 + x_4 = 1 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{array}$$

13.6. Necht' úloha (13.6) má více než jedno optimální řešení. Jak se to pozná v simplexové tabulce? Navrhnete algoritmus, jehož výstupem bude výčet všech optimálních básových řešení.

13.7. Mějme lineární program

$$\begin{array}{ll} \min & 2x_1 \quad - \quad 3x_3 + x_4 \\ \text{za podmínek} & x_1 - x_2 - x_3 \geq 0 \\ & -x_1 + 2x_2 - 3x_3 \leq 5 \\ & 2x_1 - x_2 - x_3 + 2x_4 = 6 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{array}$$

Inicializujte co nejjednodušším způsobem základní simplexový algoritmus. Vyřešte tímto algoritmem. Nepoužívejte dvoufázovou metodu.

13.8. Upravte do vhodného tvaru a vyřešte dvoufázovou simplexovou metodou:

$$\begin{array}{ll} \max & 3x_1 - 4x_2 \\ \text{za podmínek} & -2x_1 - 5x_2 \leq 10 \\ & 3x_1 + x_2 \leq 3 \\ & -2x_1 + x_2 \leq -2 \\ & x_1 \geq 0 \\ & x_2 \leq -1 \end{array}$$

Nápověda a řešení

| | | | | | | |
|-------|--------------|------------|------------|------------|------------|------------|
| | J | $\{1, 2\}$ | $\{1, 3\}$ | $\{1, 4\}$ | $\{2, 3\}$ | $\{3, 4\}$ |
| 13.1. | \mathbf{x} | $(1, -1)$ | $(0, 1)$ | $(1, 2)$ | $(0, 1)$ | $(1, 0)$ |
| | příp. | ne | ano | ano | ano | ano |
| | degen. | ne | ano | ne | ano | ano |

13.4. Úloha je neomezená kvůli prvnímu sloupci.

13.7. Např. otočíme znaménko prvního omezení, třetí omezení vydělíme dvěma, přidáme slackové proměnné pro první a druhé omezení. Pak vynulujeme ceny nad básovými sloupci. Iterace algoritmu:

| | | | | | | | | | | | | | | | | | | | | |
|----|------|----------|---|---|---|----|------------|---|---|---|-----|---|----|---|---|---|---|---|---|----|
| 1 | 0.5 | -2.5 | 0 | 0 | 0 | -3 | -1.5 | 3 | 0 | 0 | 2.5 | 0 | -3 | 0 | 3 | 0 | 3 | 4 | 0 | 6 |
| -1 | 1 | <u>1</u> | 0 | 1 | 0 | 0 | -1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 2 | 2 | 0 | 0 | 6 |
| -1 | 2 | -3 | 0 | 0 | 1 | 5 | -4 | 5 | 0 | 0 | 3 | 1 | 5 | 0 | 5 | 0 | 8 | 7 | 1 | 29 |
| 1 | -0.5 | -0.5 | 1 | 0 | 0 | 3 | <u>0.5</u> | 0 | 0 | 1 | 0.5 | 0 | 3 | 1 | 0 | 0 | 2 | 1 | 0 | 6 |

Výsledek: $(x_1, x_2, x_3, x_4) = (6, 0, 6, 0)$, hodnota optima -6 .

13.8. Optimum je $(x_1, x_2) = (25, -36)/13$.

Kapitola 14

Dualita v lineárním programování

Ke každé úloze LP lze sestavit podle jistého postupu jinou úlohu LP. Novou úlohu nazýváme **duální**, původní úlohu nazýváme **primární** či **přímou**. Konstrukce je symetrická (viz Cvičení 14.1): duální úloha k duální úloze je původní úloha. Tedy má smysl říkat, že primární a duální úloha jsou *navzájem* duální. Dvojice duálních úloh je svázána zajímavými vztahy.

14.1 Konstrukce duální úlohy

K úloze LP v obecném tvaru (viz §11.1) se duální úloha získá dle tohoto postupu:

$$\begin{array}{ll}
 \min \sum_{j \in J} c_j x_j & \max \sum_{i \in I} b_i y_i \\
 \text{za podm. } \sum_{j \in J} a_{ij} x_j = b_i & \text{za podm. } y_i \in \mathbb{R}, \quad i \in I_0 \\
 \sum_{j \in J} a_{ij} x_j \geq b_i & y_i \geq 0, \quad i \in I_+ \\
 \sum_{j \in J} a_{ij} x_j \leq b_i & y_i \leq 0, \quad i \in I_- \\
 x_j \in \mathbb{R} & \sum_{i \in I} a_{ij} y_i = c_j, \quad j \in J_0 \\
 x_j \geq 0 & \sum_{i \in I} a_{ij} y_i \leq c_j, \quad j \in J_+ \\
 x_j \leq 0 & \sum_{i \in I} a_{ij} y_i \geq c_j, \quad j \in J_-
 \end{array}$$

V levém sloupci je primární úloha, v prostředním sloupci je z ní vytvořená duální úloha. V pravém sloupci jsou množiny indexů pro obě úlohy: $I = \{1, \dots, m\} = I_0 \cup I_+ \cup I_-$ je indexová množina primárních omezení a duálních proměnných, $J = \{1, \dots, n\} = J_0 \cup J_+ \cup J_-$ je indexová množina primárních proměnných a duálních omezení.

Všimněte si symetrie: i -tému primárnímu omezení $\sum_j a_{ij} x_j \geq b_i$ odpovídá duální proměnná $y_i \geq 0$. Opačně, j -tá primární proměnná $x_j \geq 0$ odpovídá j -tému duálnímu omezení $\sum_i a_{ij} y_i \leq c_j$. Podobně pro ostatní řádky.

Pro speciální tvary LP se dvojice duálních úloh přehledněji napíše v maticové formě. Např. pro $I_0 = I_- = J_0 = J_- = \emptyset$ obdržíme

$$\begin{array}{ll}
 \min \mathbf{c}^T \mathbf{x} & \max \mathbf{b}^T \mathbf{y} \\
 \text{za podm. } \mathbf{A} \mathbf{x} \geq \mathbf{b} & \text{za podm. } \mathbf{y} \geq \mathbf{0} \\
 \mathbf{x} \geq \mathbf{0} & \mathbf{A}^T \mathbf{y} \leq \mathbf{c}
 \end{array} \quad (14.1)$$

14.2 Věty o dualitě

Následující věty platí pro obecný tvar LP, ale důkazy uděláme pouze pro speciální tvar (14.1). V důkazech si všimněte, že $\mathbf{b}^T \mathbf{y} = \mathbf{y}^T \mathbf{b}$ a $\mathbf{A}^T \mathbf{y} \leq \mathbf{c}$ je totéž jako $\mathbf{y}^T \mathbf{A} \leq \mathbf{c}^T$.

Věta 14.1 (o slabé dualitě). *Necht' \mathbf{x} je přípustné primární řešení a \mathbf{y} přípustné duální řešení. Pak $\mathbf{c}^T \mathbf{x} \geq \mathbf{b}^T \mathbf{y}$.*

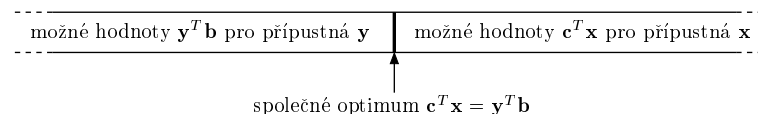
Důkaz. Díky přípustnosti \mathbf{x} a \mathbf{y} platí $\mathbf{y}^T \mathbf{A} \leq \mathbf{c}^T$ a $\mathbf{x} \geq \mathbf{0}$, z čehož plyne $\mathbf{y}^T \mathbf{A} \mathbf{x} \leq \mathbf{c}^T \mathbf{x}$. Podobně, díky přípustnosti \mathbf{x} a \mathbf{y} platí $\mathbf{A} \mathbf{x} \geq \mathbf{b}$ a $\mathbf{y} \geq \mathbf{0}$, z čehož plyne $\mathbf{y}^T \mathbf{A} \mathbf{x} \geq \mathbf{y}^T \mathbf{b}$. Z toho

$$\mathbf{c}^T \mathbf{x} \geq \mathbf{y}^T \mathbf{A} \mathbf{x} \geq \mathbf{y}^T \mathbf{b}. \quad (14.2)$$

□

Věta 14.2 (o silné dualitě). *Primární úloha má optimální řešení právě tehdy, když má duální úloha optimální řešení. Má-li primární úloha optimální řešení \mathbf{x} a duální úloha optimální řešení \mathbf{y} , platí $\mathbf{c}^T \mathbf{x} = \mathbf{b}^T \mathbf{y}$.*

Důkaz věty o silné dualitě není jednoduchý a vynecháme jej. Věty o slabé a silné dualitě mají jasnou interpretaci: pro přípustná \mathbf{x} a \mathbf{y} není hodnota duální účelové funkce nikdy větší než hodnota primární účelové funkce a tyto hodnoty se potkají ve společném optimu:



Z věty o silné dualitě ihned plyne věta o slabé dualitě (proč?), proto se někdo může ptát, proč větu o slabé dualitě uvádíme. Je to proto, že její důkaz je o mnoho jednodušší.

Uvědomte si, že věta o komplementaritě neříká, že rovnost $\mathbf{c}^T \mathbf{x} = \mathbf{b}^T \mathbf{y}$ vůbec někdy nastane. Uveďme ještě jeden jednoduchý důsledek slabé duality, který je opět slabší než silná dualita.

Důsledek 14.3. *Necht' \mathbf{x} je přípustné primární řešení a \mathbf{y} je přípustné duální řešení. Necht' $\mathbf{c}^T \mathbf{x} = \mathbf{b}^T \mathbf{y}$. Potom \mathbf{x} a \mathbf{y} jsou zároveň optimální řešení.*

Důkaz. Pro libovolné primární přípustné řešení \mathbf{x}' plyne z věty o slabé dualitě $\mathbf{b}^T \mathbf{y} \leq \mathbf{c}^T \mathbf{x}'$. Z předpokladu máme $\mathbf{c}^T \mathbf{x} = \mathbf{b}^T \mathbf{y}$. Z toho plyne $\mathbf{c}^T \mathbf{x} \leq \mathbf{c}^T \mathbf{x}'$. Jelikož toto platí pro každé přípustné \mathbf{x}' , řešení \mathbf{x} musí být optimální.

Optimalita \mathbf{y} se dokáže symetricky. □

Věta 14.4 (o komplementaritě). *Necht' \mathbf{x} je přípustné primární řešení a \mathbf{y} přípustné duální řešení. Pak $\mathbf{c}^T \mathbf{x} = \mathbf{b}^T \mathbf{y}$ právě tehdy, když zároveň platí tyto dvě podmínky:*

$$y_i \left(\sum_{j \in J} a_{ij} x_j - b_i \right) = 0 \quad \forall i \in I, \quad (14.3a)$$

$$x_j \left(\sum_{i \in I} a_{ij} y_i - c_j \right) = 0 \quad \forall j \in J. \quad (14.3b)$$

Všimněte si, co podmínky (14.3) říkají: na každém řádku ve dvojici duálních úloh je vždy alespoň jedno omezení aktivní, buď primární nebo duální (příčemž omezení typu rovnosti bereme vždy za aktivní).

Důkaz. Pro libovolné vektory $\mathbf{u}, \mathbf{v} \geq \mathbf{0}$ platí

$$\forall i (u_i = 0 \text{ nebo } v_i = 0) \iff \forall i (u_i v_i = 0) \iff \mathbf{u}^T \mathbf{v} = 0.$$

Protože \mathbf{x} a \mathbf{y} jsou přípustné, podmínky (14.3) je tedy možno psát jako

$$\mathbf{y}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) = 0, \quad (14.4a)$$

$$(\mathbf{c}^T - \mathbf{y}^T \mathbf{A})\mathbf{x} = 0, \quad (14.4b)$$

neboli

$$\mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{A}\mathbf{x} = \mathbf{y}^T \mathbf{b}. \quad (14.5)$$

Vztah (14.5) zjevně implikuje $\mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{b}$. Obráceně, $\mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{b}$ implikuje (14.5), neboť' jsme dříve ukázali, že pro přípustné \mathbf{x}, \mathbf{y} platí (14.2). \square

Příklad 14.1. Mějme dvojici navzájem duálních úloh LP:

| | |
|---|---|
| $\begin{aligned} \min \quad & 2x_1 + 5x_2 + 6x_3 = 5.4 \\ \mathbf{3} = \quad & 2x_1 + x_2 + 2x_3 \geq 3 \\ \mathbf{2.4} = \quad & x_1 + 2x_2 + 2x_3 \geq 1 \\ \mathbf{3} = \quad & x_1 + 3x_2 + x_3 \geq 3 \\ -\mathbf{0.6} = \quad & -x_1 + x_2 - 2x_3 \geq -1 \\ \mathbf{1.2} = \quad & x_1 \geq 0 \\ \mathbf{0.6} = \quad & x_2 \geq 0 \\ \mathbf{0} = \quad & x_3 \geq 0 \end{aligned}$ | $\begin{aligned} \max \quad & 3y_1 + y_2 + 3y_3 - y_4 = 5.4 \\ \mathbf{0.2} = \quad & y_1 \geq 0 \\ \mathbf{0} = \quad & y_2 \geq 0 \\ \mathbf{1.6} = \quad & y_3 \geq 0 \\ \mathbf{0} = \quad & y_4 \geq 0 \\ \mathbf{2} = \quad & 2y_1 + y_2 + y_3 - y_4 \leq 2 \\ \mathbf{5} = \quad & y_1 + 2y_2 + 3y_3 + y_4 \leq 5 \\ \mathbf{2} = \quad & 2y_1 + 2y_2 + y_3 - 2y_4 \leq 6 \end{aligned}$ |
|---|---|

Spočetli jsme optimální řešení obou úloh a dosadili tato řešení do účelových funkcí a do omezení. Hodnoty optimálních řešení $\mathbf{x}^* = (1.2, 0.6, 0)$ a $\mathbf{y}^* = (0.2, 0, 1.6)$ a hodnoty omezení a účelových funkcí v optimech jsou napsané tučně před/za rovnítky. Dle věty o silné dualitě se optima rovnají. Vezmeme-li libovolný řádek (kromě účelového), je na něm alespoň jedno z obou omezení aktivní. Např. ve druhém řádku je primární omezení $2x_1 + x_2 + 2x_3 \geq 3$ aktivní a duální omezení $y_1 \geq 0$ je neaktivní. Podle věty o komplementaritě se nemůže stát, že by na některém řádku byly obě omezení zároveň neaktivní (mohou být obě ale zároveň aktivní, což zde nenastává, ale může to nastat v případě degenerace). \square

Předložíme-li přípustná primární a duální řešení taková, že se účelové funkce rovnají, dokázali jsme optimalitu obou úloh. Pro velké úlohy to může být nejsnadnější důkaz optimality.

Máme-li duální optimální řešení, jak z něj co nejlevněji spočítat primární optimální řešení? Obecně je k tomu nutno vyřešit soustavu lineárních nerovnic (což není o moc snadnější než vyřešit lineární program). Někdy ale postačí vyřešit soustavu rovnic.

Příklad 14.2. Je dána primární úloha z Příkladu 14.1. Zkuste dokázat bez použití algoritmu na řešení LP, že $\mathbf{x} = (x_1, x_2, x_3) = (1.2, 0.6, 0)$ je optimální řešení primární úlohy (příčemž optimální duální řešení \mathbf{y} není zadáno).

Pomocí věty o komplementaritě zkusíme z daného optimálního \mathbf{x} zkusíme spočítat optimální \mathbf{y} . Protože jsou druhé a čtvrté primární omezení neaktivní, z komplementarity plyne

$y_2 = y_4 = 0$. Protože $x_1 > 0$ a $x_2 > 0$, z komplementarity musí být první a druhé duální omezení aktivní. Máme tedy soustavu lineárních rovnic

$$\begin{aligned} 2y_1 + y_3 &= 2 \\ y_1 + 3y_3 &= 5 \end{aligned} \quad (14.6)$$

kteřá má jediné řešení $(y_1, y_3) = (0.2, 1.6)$. Tedy $\mathbf{y} = (0.2, 0, 1.6, 0)$. Toto duální řešení je přípustné (tj. splňuje všechna duální omezení). Protože se hodnota primární účelové funkce v bodě \mathbf{x} rovná hodnotě duální účelové funkce v bodě \mathbf{y} , musejí být \mathbf{x} a \mathbf{y} optimální řešení.

Tento postup nemusí vést vždy k cíli. Pokud by duální úloha by měla nekonečně mnoho optimálních řešení, soustava (14.6) by měla nekonečně mnoho řešení (měla by např. více proměnných než neznámých). Z nich by bylo nutno vybrat přípustná duální řešení, tedy $\mathbf{y} \geq \mathbf{0}$. Museli bychom tedy řešit soustavu rovnic a nerovnic. \square

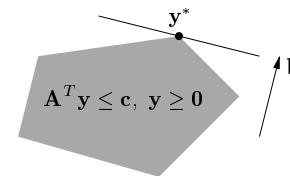
Zkoumejme, jak se změní optimální hodnota úlohy $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, jestliže nepatrně změním pravé strany omezení \mathbf{b} . Odpověď je snadno vidět v duálu.

Věta 14.5 (o stínových cenách). Necht' funkce $f: \mathbb{R}^m \rightarrow \mathbb{R}$ je definována jako

$$f(\mathbf{b}) = \min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\} = \max\{\mathbf{b}^T \mathbf{y} \mid \mathbf{A}^T \mathbf{y} \leq \mathbf{c}, \mathbf{y} \geq \mathbf{0}\}.$$

Jestliže má duální úloha pro dané \mathbf{b} jediné optimální řešení \mathbf{y}^* , pak je funkce f v bodě \mathbf{b} diferencovatelná a platí $f'(\mathbf{b}) = \mathbf{y}^{*T}$, neboli $\partial f(\mathbf{b})/\partial b_i = y_i^*$.

Důkaz. Je-li \mathbf{y}^* duální optimální řešení pro dané \mathbf{b} , je $f(\mathbf{b}) = \mathbf{b}^T \mathbf{y}^*$. Jelikož je toto optimální řešení jediné, nabývá se ve vrcholu mnohostěnu přípustných řešení $\{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{A}^T \mathbf{y} \leq \mathbf{c}, \mathbf{y} \geq \mathbf{0}\}$, viz obrázek:



Změníme-li nepatrně \mathbf{b} , optimální duální řešení \mathbf{y}^* se nezmění a zůstane jediné (toto odůvodnění není zcela rigorózní, ale geometricky je dostatečně názorné). Tedy při malé změně vektoru \mathbf{b} je hodnota optima stále rovna $f(\mathbf{b}) = \mathbf{b}^T \mathbf{y}^*$. Derivací získáme $f'(\mathbf{b}) = \mathbf{y}^{*T}$. \square

Zdůrazněme předpoklad jednoznačnosti optimálního řešení. Kdyby množina duálních optimálních řešení byla ne jediný vrchol, ale stěna vyšší dimenze, po malé změně vektoru \mathbf{b} by se optimální stěna mohla stát vrcholem a funkce f by tedy v bodě \mathbf{b} nebyla diferencovatelná.

Protože \mathbf{b} je zároveň vektor pravých stran primární úlohy, optimální duální proměnné \mathbf{y}^* vyjadřují citlivost optima primární úlohy na změnu pravých stran primárních omezení $\mathbf{A}\mathbf{x} \geq \mathbf{b}$. Interpretujeme-li naše LP jako optimální výrobní plán (11.6) (pozor, liší se obrácenou nerovností v omezení), pak hodnota y_i^* říká, jak by se náš výdělek zvětšil, kdybychom trochu uvolnili omezení na výrobní zdroje $\mathbf{a}_i^T \mathbf{x} \leq b_i$. V ekonomii se proto duálním proměnným říká **stínové ceny** primárních omezení.

Všimněte si, že věta o stínových cenách je ve shodě s větou o komplementaritě. Pokud $y_i^* = 0$, je $\mathbf{a}_i^T \mathbf{x} < b_i$, tedy malá změna b_i nemá na optimum vliv.

Příklad 14.3. Necht' je známo, že duální úloha v Příkladu 14.1 má jediné optimální řešení. Stínová cena prvního primárního omezení $2x_1 + x_2 + 2x_3 \geq 3$ je $y_1 = 0.2$. Změňme pravou stranu $b_1 = 3$ tohoto omezení o malou hodnotu $h = 0.01$ a zkoumejme, jak se změní optimum. Tato změna nezmění argument \mathbf{y}^* duálního optima, pouze jeho hodnotu $\mathbf{b}^T \mathbf{y}^*$. Podle silné duality hodnota primárního optima musí být rovna hodnotě duálního optima (argument \mathbf{x}^* primárního optima se může nějak změnit, to nás ale nezajímá). Dvojice úloh tedy bude vypadat takto:

$$\begin{array}{ll} \min & 2x_1 + 5x_2 + 6x_3 = \mathbf{5.402} \\ & 2x_1 + x_2 + 2x_3 \geq 3.01 \\ & x_1 + 2x_2 + 2x_3 \geq 1 \\ & x_1 + 3x_2 + x_3 \geq 3 \\ & -x_1 + x_2 - 2x_3 \geq -1 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \\ & x_3 \geq 0 \end{array} \quad \begin{array}{ll} \max & 3.01y_1 + y_2 + 3y_3 - y_4 = \mathbf{5.402} \\ & 0.2 = y_1 \\ & 0 = y_2 \\ & 1.6 = y_3 \\ & 0 = y_4 \\ & 2 = 2y_1 + y_2 + y_3 - y_4 \leq 2 \\ & 5 = y_1 + 2y_2 + 3y_3 + y_4 \leq 5 \\ & 2 = 2y_1 + 2y_2 + y_3 - 2y_4 \leq 6 \end{array}$$

V okolí bodu $\mathbf{b} = (3, 1, 3, -1)$, ve kterém se nemění optimální \mathbf{y}^* , bude $f(\mathbf{b}) = \mathbf{b}^T \mathbf{y}^*$ a tedy hodnota společného optima se změní o $h \cdot \partial f(\mathbf{b}) / \partial b_1 = h \cdot y_1 = 0.2 \cdot 0.01$ na 5.402. \square

14.3 Příklady na konstrukci a interpretaci duálních úloh

Dualita umožňuje *vhled* do řešeného problému, často velmi netriviální. Abychom danou úlohu (fyzikální, ekonomickou či jinou) popsanou lineárním programem pochopili do hloubky, je třeba pochopit význam nejen primární úlohy, ale i duální úlohy a jejich vztahu.

Příklad 14.4 (Ekonomická interpretace duality). Vrat' se k Příkladu 11.5 o výrobci lupínků a hranolků z brambor a oleje. Napišme k této úloze duální úlohu:

$$\begin{array}{ll} \max & 120l + 76h \\ \text{za podm.} & 2l + 1.5h \leq 100 \\ & 0.4l + 0.2h \leq 16 \\ & l \geq 0 \\ & h \geq 0 \end{array} \quad \begin{array}{ll} \min & 100a + 16b \\ \text{za podm.} & a \geq 0 \\ & b \geq 0 \\ & 2a + 0.4b \geq 120 \\ & 1.5a + 0.2b \geq 76 \end{array}$$

Přijde překupník a chce koupit od výrobce jeho zásoby brambor a oleje. Překupník řeší tuto otázku: Jaké nejnižší ceny musím nabídnout, aby mi výrobce své zásoby prodal? Tvrdíme, že toto je význam duální úlohy.

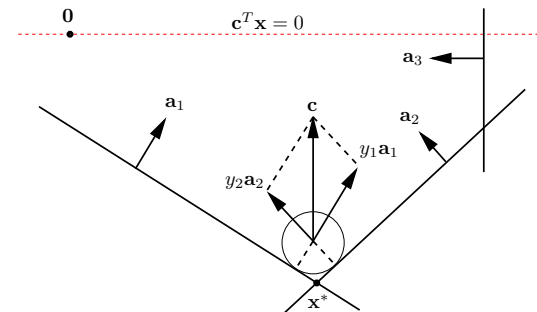
Vskutku, necht' a, b označují nabízenou cenu za jednotku brambor a oleje. Překupník chce minimalizovat celkovou cenu za suroviny $100a + 16b$. Musí být $2a + 0.4b \geq 120$, neboť jinak by výrobci více vyplatilo vyrobit ze všech brambor a oleje lupínky a prodat je, než prodat suroviny. Ze stejného důvodu musí být $1.5a + 0.2b \geq 76$. Optimální duální řešení je $a = 32$ a $b = 140$.

Toto je další důvod (kromě Věty 14.5), proč se optimálním duálním proměnným někdy říká *stínové ceny* odpovídajících primárních omezení. Např. stínová cena brambor je 32 Kč/kg. \square

Příklad 14.5 (Fyzikální interpretace duality). Uvažujme dvojici duálních úloh

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \in \mathbb{R}^n\} = \max\{\mathbf{b}^T \mathbf{y} \mid \mathbf{A}^T \mathbf{y} = \mathbf{c}, \mathbf{y} \geq \mathbf{0}\}.$$

Uvažujme následující 'analogový počítač'. Mějme mnohostěn tvořený třemi poloprostory $\mathbf{a}_i^T \mathbf{x} \geq b_i$ a vektor \mathbf{c} mířící svisle vzhůru:



Hoďme do mnohostěnu malý míček, na který působí tíhová síla $-\mathbf{c}$. Míček s pozicí \mathbf{x} má potenciální energii $\mathbf{c}^T \mathbf{x}$. Míček se bude pohybovat do té doby, než nalezneme místo s nejmenší potenciální energií, což je nejnižší vrchol \mathbf{x}^* . Tedy \mathbf{x}^* je řešením primární úlohy.

V bodě \mathbf{x}^* je míček v klidu a proto pro něj platí rovnováha sil: tíha $-\mathbf{c}$ se vyrovnává silami stěn. Tedy existují skaláry y_i^* tak, že $\mathbf{c} = \sum_i y_i^* \mathbf{a}_i = \mathbf{A}^T \mathbf{y}^*$. Musí být $y_i^* \geq 0$, protože stěny působí silou jen dovnitř mnohostěnu, ne ven.

Pokud $\mathbf{a}_i^T \mathbf{x}^* > b_i$, míček se i -té stěny nedotýká. V tom případě je síla stěny na míček nutně nulová, $y_i^* = 0$. Proto pro každé i platí $y_i^* (\mathbf{a}_i^T \mathbf{x}^* - b_i) = 0$, což jsou podmínky (14.3). Dle věty o komplementaritě tedy je $\mathbf{c}^T \mathbf{x}^* = \mathbf{b}^T \mathbf{y}^*$.

To ale říká, že potenciální energie míčku v bodě \mathbf{x}^* je $\mathbf{b}^T \mathbf{y}^*$. Proč tomu tak je? Tato energie se rovná práci, která by se vykonala posunutím míčku do počátku $\mathbf{0}$. Posunujeme míček takto: nejdříve posuneme první stěnu rovnoběžně tak, aby procházela počátkem, pak druhou stěnu, až budou všechny stěny, kterých se míček dotýká, procházet počátkem. Při tomto posouvání se síly stěn působící na míček nemění. Vzdálenost stěny i od počátku je $b_i / \|\mathbf{a}_i\|_2$ (viz Cvičení .a). Síla stěny i působící na míček je $y_i^* \mathbf{a}_i$. Síla je ve směru posouvání, tedy vykonaná práce je $(b_i / \|\mathbf{a}_i\|_2) \cdot \|y_i^* \mathbf{a}_i\|_2 = b_i y_i^*$. Po odtlačení všech stěn, jichž se míček dotýká, do počátku, tedy vykonáme práci $\sum_i b_i y_i^* = \mathbf{b}^T \mathbf{y}^*$.

Zdůrazněme, že tato úvaha *nedokazuje* žádnou z vět o dualitě. Předpokládá totiž platnost fyzikálních zákonů, které nelze matematicky dokázat ale pouze experimentálně pozorovat. \square

Příklad 14.6. Mějme úlohu

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}\} = \min\left\{\sum_{i=1}^n c_i x_i \mid \sum_{i=1}^n x_i = 1, x_i \geq 0\right\},$$

kde $\mathbf{c} = (c_1, \dots, c_n)$ je dáno a optimalizuje se přes $\mathbf{x} = (x_1, \dots, x_n)$. Napište duální úlohu a interpretujte věty o silné dualitě a komplementaritě.

Úvahou snadno vidíme (viz Příklad 11.2), že optimální hodnota je $\min_{i=1}^n c_i$ a nabývá se ve vektoru \mathbf{x} jehož všechny složky jsou nulové kromě složek příslušných minimálnímu c_i . Pokud je více minimálních prvků c_i , optimální \mathbf{x} není dáno jednoznačně. Např. pro $\mathbf{c} = (1, 3, 1, 2)$ bude optimálním řešením každé $\mathbf{x} = (x_1, 0, x_3, 0)$ pro $x_1, x_3 \geq 0$ splňující $x_1 + x_3 = 1$.

Podle návodu na konstrukci duální úlohy dostaneme duál

$$\max\{y \in \mathbb{R} \mid y \mathbf{1} \leq \mathbf{c}\} = \max\{y \in \mathbb{R} \mid y \leq c_i, i = 1, \dots, n\}.$$

Neboli hledá se největší číslo y , které je menší než všechna čísla c_i . Takové číslo y se rovná minimu z čísel c_i . Tedy platí silná dualita.

Podmínky komplementarity říkají, že v optimech bude alespoň jedno z odpovídající dvojice primární-duální omezení aktivní. Dvojice omezení $\sum_i x_i = 1$, $y \in \mathbb{R}$ splňuje podmínky komplementarity triviálně. Dvojice omezení $x_i \geq 0$, $y \leq c_i$ je splňuje právě tehdy, když je splněna aspoň jedna z rovností $x_i = 0$, $y = c_i$. To znamená:

- Pokud je v duálu $y < c_i$, v primáru musí být $x_i = 0$. To je ale jasné, protože $y < c_i$ znamená, že c_i není nejmenší ze složek vektoru \mathbf{c} a tudíž (dle úvahy výše) mu v primáru nemůžeme přiřadit nenulovou váhu x_i .
- Obráceně, pokud je v primáru $x_i > 0$, musí být v duálu $y = c_i$. To je jasné, protože pokud jsme v primáru přiřadili číslu c_i nenulovou váhu, musí být nejmenší. \square

Příklad 14.7. Z §11.3 víme, že optimální argument úlohy

$$\begin{aligned} \min_{x \in \mathbb{R}} \sum_{i=1}^n |x - a_i| &= \min \left\{ \sum_{i=1}^n z_i \mid z_i \in \mathbb{R}, x \in \mathbb{R}, -z_i \leq x - a_i \leq z_i \right\} \\ &= \min \{ \mathbf{1}^T \mathbf{z} \mid \mathbf{z} \in \mathbb{R}^n, x \in \mathbb{R}, -\mathbf{z} \leq \mathbf{1}x - \mathbf{a} \leq \mathbf{z} \} \end{aligned} \quad (14.7)$$

je medián z čísel a_1, \dots, a_n . Napište duální úlohu a co nejvíce ji zjednodušte. Úvahou nalezněte optimální hodnotu primární a duální úlohy a ověřte, že se (dle silné duality) rovnají.

Rychlý způsob jak vytvořit duální úlohu je podle předpisu v §14.1, to ovšem vyžaduje zkušenost a snadno se v tom udělá chyba. Zdlouhavější avšak bezpečnější způsob je přes maticovou formu. Primární a duální úlohu napíšeme v maticovém tvaru, kde zvolíme názvy matic tak, aby nekolidovaly s (14.7):

$$\begin{array}{ll} \min & \mathbf{h}^T \mathbf{u} \\ \text{za podm.} & \mathbf{F}\mathbf{u} \geq \mathbf{g} \\ & \mathbf{u} \in \mathbb{R}^{1+n} \end{array} \quad \begin{array}{ll} \max & \mathbf{g}^T \mathbf{v} \\ \text{za podm.} & \mathbf{v} \geq \mathbf{0} \\ & \mathbf{F}^T \mathbf{v} = \mathbf{h} \end{array}$$

Matrice zvolíme tak, aby primární úloha odpovídala úloze (14.7):

$$\mathbf{F} = \begin{bmatrix} \mathbf{1} & \mathbf{I} \\ -\mathbf{1} & \mathbf{I} \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} \mathbf{a} \\ -\mathbf{a} \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} x \\ \mathbf{z} \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix}.$$

Vektor duálních proměnných \mathbf{v} jsme zároveň rozdělili na dva bloky \mathbf{p}, \mathbf{q} , odpovídající blokům matic \mathbf{F} a \mathbf{g} . Vynásobním matic přepíšeme duální úlohu do tvaru (ověřte!)

$$\begin{aligned} & \max \{ \mathbf{a}^T (\mathbf{p} - \mathbf{q}) \mid \mathbf{1}^T (\mathbf{p} - \mathbf{q}) = 0, \mathbf{p} + \mathbf{q} = \mathbf{1}, \mathbf{p} \geq \mathbf{0}, \mathbf{q} \geq \mathbf{0} \} \\ &= \max \left\{ \sum_{i=1}^n a_i (p_i - q_i) \mid \sum_{i=1}^n (p_i - q_i) = 0, p_i + q_i = 1, p_i \geq 0, q_i \geq 0 \right\} \end{aligned} \quad (14.8)$$

Úlohu (14.8) lze dále zjednodušit substitucí

$$2p_i = 1 + t_i, \quad 2q_i = 1 - t_i.$$

Po této substitucí je $p_i - q_i = t_i$ a podmínka $p_i + q_i = 1$ je splněna automaticky. Podmínka $\sum_i (p_i - q_i) = 0$ odpovídá podmínce $\sum_i t_i = 0$. Podmínka $p_i \geq 0$ odpovídá $t_i \geq -1$ a podmínka $q_i \geq 0$ odpovídá $t_i \leq 1$. Duální úloha s novými proměnnými $\mathbf{t} \in \mathbb{R}^n$ je tedy

$$\max \left\{ \sum_{i=1}^n a_i t_i \mid \sum_{i=1}^n t_i = 0, -1 \leq t_i \leq 1 \right\} = \max \{ \mathbf{a}^T \mathbf{t} \mid \mathbf{1}^T \mathbf{t} = 0, -\mathbf{1} \leq \mathbf{t} \leq \mathbf{1} \}. \quad (14.9)$$

Primární úloha (14.7) a duální úloha (14.9) spolu zdánlivě vůbec nesouvisejí – avšak podle silné duality jejich optimální hodnoty musí být stejné! Zkusme pochopit, proč tomu tak je.

Nejprve si všimneme, že optimální hodnota primární úlohy (14.7) se nezmění, posuneme-li čísla a_1, \dots, a_n o libovolnou konstantu $b \in \mathbb{R}$. To je jasné, neboť medián x se posune o stejnou konstantu a je $|(x - b) - (a_i - b)| = |x - a_i|$. Totéž platí pro duální úlohu (14.9), neboť díky podmínce $\sum_i t_i = 0$ je $\sum_i (a_i - b)t_i = \sum_i a_i t_i$. Proto bez ztráty obecnosti můžeme zvolit $b = \text{median}_i a_i$, neboli posunout body tak, že jejich medián bude $x = 0$.

Nyní je primární optimální hodnota rovna jednoduše $\sum_i |x - a_i| = \sum_i |a_i|$. Protože kladných a záporných čísel a_i je stejný počet, duální úloha nabývá optima v takovém vektoru \mathbf{t} , kde $t_i = -1$ pro $a_i < 0$ a $t_i = 1$ pro $a_i > 0$ (což splňuje podmínku $\sum_i t_i = 0$). Tedy duální optimální hodnota je také $\sum_i a_i t_i = \sum_i |a_i|$. \square

14.4 Cvičení

14.1. Ukažte pro dvojici úloh v §14.1, že duál duálu se rovná původní úloze.

14.2. Napište duální úlohu a podmínky komplementarity k následujícím úlohám. Pokud úloha není LP, nejdříve převed'te na LP (dle §11.1). Výslednou duální úlohu co nejvíce zjednodušte, příp. převed'te do skalární formy, je-li skalární forma výstižnější.

- $\min_{x \in \mathbb{R}} \max_{i=1}^n |a_i - x|$ (střed intervalu)
- úloha (11.11)
- úloha (11.13)
- dopravní problém (11.8)
- všechny úlohy ze Cvičení 11.2
- úloha vzniklá ve Cvičení 11.8
- $\min_{\mathbf{x} \in \mathbb{R}^n} \max_{i=1}^m (\mathbf{a}_i^T \mathbf{x} + b_i)$ (viz §11.1.1)

(*) Pro každou úlohu se pokuste nalézt význam duální úlohy, podobně jako v Příkladu 14.6.

14.3. Dokažte bez užití algoritmu na řešení LP, že $\mathbf{x} = (1, 1, 1, 1)$ je optimální řešení úlohy

$$\begin{aligned} & \min \quad [47 \quad 93 \quad 17 \quad -93] \mathbf{x} \\ & \text{za podm.} \quad \begin{bmatrix} -1 & -6 & 1 & 3 \\ -1 & -2 & 7 & 1 \\ 0 & 3 & -10 & -1 \\ -6 & -11 & -2 & 12 \\ 1 & 6 & -1 & -3 \end{bmatrix} \mathbf{x} \leq \begin{bmatrix} -3 \\ 5 \\ -8 \\ -7 \\ 4 \end{bmatrix} \end{aligned}$$

Nápověda a řešení

14.1. Pravou úlohu nejdříve převedeme na tvar levé úlohy a pak k ní napíšeme duální úlohu. Ukážeme jen pro tvar (14.1):

$$\begin{array}{ll} -\min & (-\mathbf{b})^T \mathbf{y} \\ \text{za podm.} & (-\mathbf{A}^T) \mathbf{y} \geq -\mathbf{c} \\ & \mathbf{y} \geq \mathbf{0} \end{array} \quad \begin{array}{ll} -\max & (-\mathbf{c})^T \mathbf{x} \\ \text{za podm.} & \mathbf{x} \geq \mathbf{0} \\ & (-\mathbf{A}) \mathbf{x} \leq -\mathbf{b} \end{array} \quad (14.10)$$

Levá úloha (14.10) je ekvivalentní pravé úloze (14.1), pravá úloha (14.10) je ekvivalentní levé úloze (14.1).

$$14.2.d) \max \left\{ \sum_{i=1}^m a_i p_i + \sum_{j=1}^n b_j q_j \mid p_i \in \mathbb{R}, q_i \in \mathbb{R}, p_i + q_j \leq c_{ij} \right\}$$

$$14.2.f) \text{ Duál: } \max \left\{ \sum_{i=1}^n y_i d_i + \sum_{i=1}^{n'} y'_i d'_i \mid \sum_{i=1}^n y_i = \sum_{i=1}^{n'} y'_i, y_i \leq m_i, y'_i \leq m'_i, y_i, y'_i \geq 0 \right\}.$$

Podmínky komplementarity: $z_i(y_i - m_i) = 0, z'_i(y'_i - m'_i) = 0, (z_i - d_i - x)y_i = 0, (z'_i - d'_i + x)y'_i = 0.$

Kapitola 15

Konvexní funkce

Definice 15.1. Funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je **konvexní** na konvexní množině $X \subseteq \mathbb{R}^n$, jestliže

$$\mathbf{x} \in X, \mathbf{y} \in X, 0 \leq \alpha \leq 1 \implies f(\alpha \mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y}). \quad (15.1)$$

Funkce f je **konkávnní** na množině X , jestliže je funkce $-f$ konvexní na množině X .

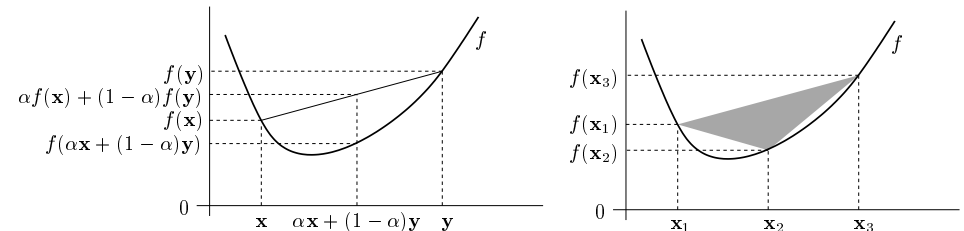
Rozlišujte pojem *konvexní množina* a *konvexní funkce*, jde o různé věci. Dále si všimněte, že X musí být konvexní množina – pojem konvexní funkce na nekonvexní množině není definován. Pokud X je celý definiční obor funkce f , odkaz na X můžeme vynechat a říkáme pouze, že funkce f je konvexní.

Podmínku (15.1) lze zobecnit pro více než dva body: funkce f je konvexní právě tehdy, když

$$\mathbf{x}_1, \dots, \mathbf{x}_k \in X, \alpha_1, \dots, \alpha_k \geq 0, \alpha_1 + \dots + \alpha_k = 1 \implies f(\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k) \leq \alpha_1 f(\mathbf{x}_1) + \dots + \alpha_k f(\mathbf{x}_k). \quad (15.2)$$

Podmínka (15.2) (známá jako **Jensenova nerovnost**) zjevně implikuje (15.1) a indukcí lze dokázat, že to platí i naopak. Porovnejte s definicí lineárního zobrazení (3.4)!

Geometrický význam podmínky (15.1) je ten, že úsečka spojující body $(\mathbf{x}, f(\mathbf{x}))$ a $(\mathbf{y}, f(\mathbf{y}))$ leží nad grafem funkce (viz levý obrázek). Geometrický význam podmínky (15.2) je ten, že konvexní mnohostěn vybarvený šedě (viz pravý obrázek) leží nad grafem funkce. Podrobně rozmyslete, jak tyto geometrické interpretace odpovídají výrazům (15.1) a (15.2)!



Příklad 15.1. Dokažme z Definice 15.1, že funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ definovaná jako $f(\mathbf{x}) = \max_{i=1}^n x_i$

je konvexní. Máme dokázat, že pro každé \mathbf{x}, \mathbf{y} a $0 \leq \alpha \leq 1$ platí

$$f(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}) = \max_i (\alpha x_i + (1 - \alpha) y_i) \quad (15.3a)$$

$$\leq \max_i \alpha x_i + \max_i (1 - \alpha) y_i \quad (15.3b)$$

$$= \alpha \max_i x_i + (1 - \alpha) \max_i y_i \quad (15.3c)$$

$$= \alpha f(\mathbf{x}) + (1 - \alpha) f(\mathbf{y}) \quad (15.3d)$$

kde rovnost (15.3c) plyne z nezápornosti čísel α a $1 - \alpha$.

Nerovnost (15.3b) plyne z toho, že pro každé $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}$ platí

$$\max_i (a_i + b_i) \leq \max_i a_i + \max_i b_i. \quad (15.4)$$

Nerovnost (15.4) dokážeme například takto. Necht' i^*, j^*, k^* jsou indexy, ve kterých se nabývají maxima, tedy $a_{i^*} + b_{i^*} = \max_i (a_i + b_i)$, $a_{j^*} = \max_i a_i$, $b_{k^*} = \max_i b_i$. Proto $a_{i^*} \leq a_{j^*}$ a $b_{i^*} \leq b_{k^*}$. Tedy $\max_i (a_i + b_i) = a_{i^*} + b_{i^*} \leq a_{j^*} + b_{k^*} = \max_i a_i + \max_i b_i$. \square

Příklad 15.2. Dokažme z Definice 15.1, že funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ definovaná jako $f(\mathbf{x}) = \min_{i=1}^n x_i$ není konvexní. Např. volba $n = 2$, $\mathbf{x} = (0, 2)$, $\mathbf{y} = (2, 0)$, $\alpha = \frac{1}{2}$ nespĺňuje (15.1), neboť

$$f((\mathbf{x} + \mathbf{y})/2) = f(1, 1) = 1 > (f(\mathbf{x}) + f(\mathbf{y}))/2 = (0 + 0)/2 = 0. \quad \square$$

Použitím Jensenovy nerovnosti na vhodnou konvexní funkci lze získat mnoho užitečných nerovností.

Příklad 15.3. Funkce \log je konkávní na \mathbb{R}_{++} . Napišme pro tuto funkci Jensenovu nerovnost (15.2) (jelikož funkce je konkávní a ne konvexní, musíme v Jensenově nerovnosti obrátit znaménko nerovnosti), ve které položíme $\alpha_1 = \dots = \alpha_n = \frac{1}{n}$:

$$\log \frac{x_1 + \dots + x_n}{n} \geq \frac{\log x_1 + \dots + \log x_n}{n}$$

kde x_1, \dots, x_n jsou kladné. Vezmeme-li exponenciálu každé strany, dostaneme

$$\frac{x_1 + \dots + x_n}{n} \geq (x_1 \dots x_n)^{1/n}.$$

Tato známá nerovnost říká, že aritmetický průměr není menší než geometrický. \square

Příklad 15.4. Uveďme často potkávané jednoduché konvexní či konkávní funkce:

1. Exponenciála $f(x) = e^{ax}$ je konvexní na \mathbb{R} , pro libovolné $a \in \mathbb{R}$.
2. Mocnina $f(x) = x^a$ je na \mathbb{R}_{++} konvexní pro $a \geq 1$ nebo $a \leq 0$ a konkávní pro $0 \leq a \leq 1$.
3. Mocnina absolutní hodnoty $f(x) = |x|^a$ je pro $a \geq 1$ konvexní na \mathbb{R} (speciálně: absolutní hodnota $|x|$ je konvexní).
4. Logaritmus $f(x) = \log x$ je konkávní na \mathbb{R}_{++} .
5. Záporná entropie $f(x) = x \log x$ je konvexní na \mathbb{R}_{++} (nebo i na \mathbb{R}_+ , pokud dodefinujeme $0 \log 0 = 0$, což se často dělá, protože $\lim_{x \rightarrow 0^+} x \log x = 0$).
6. Afinní funkce $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b$ je zároveň konvexní i konkávní.

7. Kvadratická forma $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ je konvexní pro \mathbf{A} pozitivně semidefinitní, konkávní pro \mathbf{A} negativně semidefinitní, a nekonvexní a nekonkávní pro \mathbf{A} indefinitní (viz Příklad 15.5).
8. Maximum složek $f(\mathbf{x}) = \max_{i=1}^n x_i = \max\{x_1, \dots, x_n\}$ je konvexní na \mathbb{R}^n .
9. Log-sum-exp funkce $f(\mathbf{x}) = \log(e^{x_1} + \dots + e^{x_n})$ je konvexní. Tato funkce se někdy nazývá *měkké maximum*, neboť funkce

$$f_t(\mathbf{x}) = f(t\mathbf{x})/t = \log(e^{tx_1} + \dots + e^{tx_n})/t$$

se pro $t \rightarrow +\infty$ blíží funkci $\max_{i=1}^n x_i$ (dokažte výpočtem limity!).

10. Geometrický průměr $f(\mathbf{x}) = (x_1 \dots x_n)^{1/n}$ je konkávní na \mathbb{R}_+^n .
11. Každá norma (viz Definice 11.1) je konvexní funkce, neboť pro každé $0 \leq \alpha \leq 1$ máme

$$\|\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}\| \leq \|\alpha \mathbf{x}\| + \|(1 - \alpha) \mathbf{y}\| = \alpha \|\mathbf{x}\| + (1 - \alpha) \|\mathbf{y}\|,$$

kde nerovnost plyne z trojúhelníkové nerovnosti a rovnost z homogenity.

Nakreslete či představte si vrstevnice a grafy těchto funkcí (v případě více proměnných pro $n = 1$ a $n = 2$)! \square

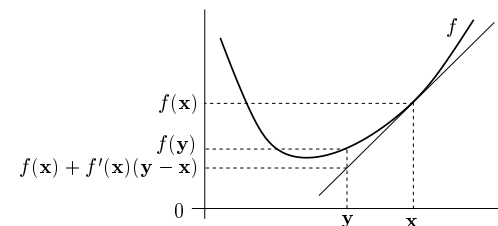
15.1 Konvexita diferencovatelných funkcí

Konvexní funkce nemusí být v každém bodě diferencovatelná (uvažte např. funkci $f(x) = |x|$). Pokud je ale funkce jednou či dvakrát diferencovatelná, její konvexitu lze snadněji než pomocí podmínky (15.1) (které se někdy říká podmínka nultého řádu) charakterizovat pomocí derivací. Následující dvě věty uvedeme bez důkazů.

Věta 15.1 (Podmínka prvního řádu). Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je diferencovatelná na konvexní množině $X \subseteq \mathbb{R}^n$. Funkce f je konvexní na množině X právě tehdy, když

$$\mathbf{x} \in X, \mathbf{y} \in X \implies f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x}).$$

To znamená, že Taylorův polynom prvního řádu funkce f v každém bodě $\mathbf{x} \in X$ (viz (8.13b)) je všude (tj. pro každé \mathbf{y}) menší nebo roven funkci f :



Věta 15.2 (Podmínka druhého řádu). Necht' $X \subseteq \mathbb{R}^n$ je konvexní množina, která má pouze vnitřní body. Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je dvakrát diferencovatelná na X . Funkce f je konvexní na množině X právě tehdy, když v každém bodě $\mathbf{x} \in X$ je Hessova matice $f''(\mathbf{x})$ pozitivně semidefinitní.

Příklad 15.5. Necht' $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$, kde \mathbf{A} je symetrická pozitivně semidefinitní. Ukažme konvexitu této funkce třemi způsoby:

- Dokažme konvexitu z Věty 15.2. To je triviální, protože Hessián je $f''(\mathbf{x}) = 2\mathbf{A}$ a tedy je pozitivně semidefinitní.
- Dokažme konvexitu z Věty 15.1. Protože $f'(\mathbf{x}) = 2\mathbf{x}^T \mathbf{A}$, máme dokázat, že

$$\mathbf{y}^T \mathbf{A} \mathbf{y} \geq \mathbf{x}^T \mathbf{A} \mathbf{x} + 2\mathbf{x}^T \mathbf{A} (\mathbf{y} - \mathbf{x}).$$

To jde upravit na $\mathbf{x}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{A} \mathbf{y} + \mathbf{y}^T \mathbf{A} \mathbf{y} \geq 0$. Platí¹

$$\mathbf{x}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{A} \mathbf{y} + \mathbf{y}^T \mathbf{A} \mathbf{y} = (\mathbf{x} - \mathbf{y})^T \mathbf{A} (\mathbf{x} - \mathbf{y}), \quad (15.5)$$

což je nezáporné pro každé \mathbf{x}, \mathbf{y} , protože \mathbf{A} je pozitivně semidefinitní.

- Dokážme konvexitu z Definice 15.1. Musíme dokázat, že pro každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ a $0 \leq \alpha \leq 1$ platí (15.1), tedy

$$[\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}]^T \mathbf{A} [\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}] \leq \alpha \mathbf{x}^T \mathbf{A} \mathbf{x} + (1 - \alpha) \mathbf{y}^T \mathbf{A} \mathbf{y}$$

Po roznásobení a převedení všech členů na jednu stranu upravujeme:

$$\begin{aligned} (\alpha - \alpha^2) \mathbf{x}^T \mathbf{A} \mathbf{x} - 2\alpha(1 - \alpha) \mathbf{x}^T \mathbf{A} \mathbf{y} + ((1 - \alpha) - (1 - \alpha)^2) \mathbf{y}^T \mathbf{A} \mathbf{y} &\geq 0 \\ \alpha(1 - \alpha) (\mathbf{x}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{A} \mathbf{y} + \mathbf{y}^T \mathbf{A} \mathbf{y}) &\geq 0. \end{aligned}$$

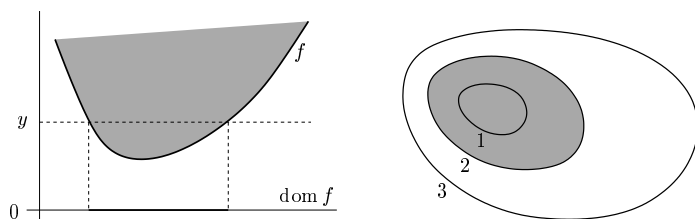
Výraz $\alpha(1 - \alpha)$ je pro každé $0 \leq \alpha \leq 1$ nezáporný. Nezápornost výrazu (15.5) jsme již ukázali. \square

15.2 Vztah konvexní funkce a konvexní množiny

Zopakujte si pojmy vrstevnice a graf funkce z §1.1.3! Zavedeme dva podobné pojmy, které se liší pouze nahrazením rovností nerovnostmi. Pro funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ definujeme:

- **Epigraf** funkce je množina $\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid f(\mathbf{x}) \leq y\}$.
- **Subkontura**² výšky y je množina $\{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \leq y\}$.

Levý obrázek znázorňuje subkonturu výšky y a epigraf funkce $\mathbb{R} \rightarrow \mathbb{R}$, pravý obrázek subkonturu výšky 2 funkce $\mathbb{R}^2 \rightarrow \mathbb{R}$:



Existují těsné vztahy mezi konvexitou funkce a konvexitou jejího epigrafu a subkontur (což jsou množiny), dané následujícími větami.

¹ Všimněte si, že pro $n = 1$ a $\mathbf{A} = 1$ se rovnost (15.5) zjednoduší na známé $x^2 - 2xy + y^2 = (x - y)^2$.

² Slovo 'subkontura' je pokus o český překlad anglického 'sublevel set'.

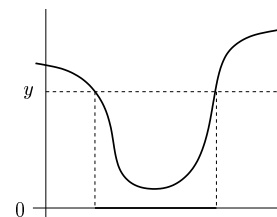
Věta 15.3. Je-li f konvexní funkce, pak je každá subkontura této funkce konvexní množina.

Důkaz. Předpokládejme, že body \mathbf{x}_1 a \mathbf{x}_2 patří do subkontury, tedy $f(\mathbf{x}_1) \leq y$ a $f(\mathbf{x}_2) \leq y$. Pro každé $0 \leq \alpha \leq 1$ platí

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \leq \alpha y + (1 - \alpha) y = y,$$

kde první nerovnost plyne z konvexity funkce a druhá z nerovností $f(\mathbf{x}_1) \leq y$, $f(\mathbf{x}_2) \leq y$. Tedy bod $\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$ patří do subkontury, která je proto konvexní množina. \square

Obrácená implikace ve Větě 15.3 neplatí: snadno najdeme funkci, která není konvexní a jejíž každá subkontura je konvexní množina³. Příklad je na obrázku:



Věta 15.4. Funkce f je konvexní právě tehdy, když její epigraf je konvexní množina.

Důkaz. Předpokládejme, že funkce f je konvexní. Vezměme dva body (\mathbf{x}_1, y_1) a (\mathbf{x}_2, y_2) z epigrafu, tedy $f(\mathbf{x}_1) \leq y_1$ a $f(\mathbf{x}_2) \leq y_2$. Pro každé $0 \leq \alpha \leq 1$ platí

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \leq \alpha y_1 + (1 - \alpha) y_2,$$

kde první nerovnost plyne z konvexity funkce a druhá nerovnost z $f(\mathbf{x}_1) \leq y_1$ a $f(\mathbf{x}_2) \leq y_2$. Tedy bod $\alpha(\mathbf{x}_1, y_1) + (1 - \alpha)(\mathbf{x}_2, y_2)$ patří do epigrafu, který je proto konvexní množina.

Předpokládejme, že epigraf je konvexní množina. Tedy pokud body (\mathbf{x}_1, y_1) a (\mathbf{x}_2, y_2) patří do epigrafu, pak také bod $\alpha(\mathbf{x}_1, y_1) + (1 - \alpha)(\mathbf{x}_2, y_2)$ patří do epigrafu pro každé $0 \leq \alpha \leq 1$. Volbou $y_1 = f(\mathbf{x}_1)$ a $y_2 = f(\mathbf{x}_2)$ máme

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha y_1 + (1 - \alpha) y_2 = \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2),$$

proto je funkce f konvexní. \square

15.3 Operace zachovávající konvexitu funkcí

Operace zachovávající konvexitu funkcí umožňují z jednoduchých konvexních funkcí získat složitější. Konvexitu složitější funkce je často snadnější dokázat pohodlněji pomocí těchto operací než z Definice 15.1 nebo Vět 15.1 a 15.2.

Jsou-li $g_1, \dots, g_k: \mathbb{R}^n \rightarrow \mathbb{R}$ konvexní funkce a $\alpha_1, \dots, \alpha_k \geq 0$, je snadné dokázat z Definice 15.1 (proved'te!), že také funkce

$$f = \alpha_1 g_1 + \dots + \alpha_k g_k \quad (15.6)$$

³ Funkce, jejíž každá subkontura je konvexní množina, se nazývá *kvazikonvexní*. Kvazikonvexní funkce nejsou zdaleka tak hezké jako konvexní funkce.

je konvexní. Speciálně, jsou-li f a g konvexní funkce, pak $f + g$ je konvexní.

Zkoumejme nyní složenou funkci $f(\mathbf{x}) = (h \circ \mathbf{g})(\mathbf{x}) = h(\mathbf{g}(\mathbf{x}))$, kde $\mathbb{R}^n \xrightarrow{\mathbf{g}} \mathbb{R}^m \xrightarrow{h} \mathbb{R}$. Obecně *neplatí* ani v případě $m = n = 1$, že konvexitá funkcí g a h zaručuje konvexitu funkce f . Nutné a postačující podmínky pro konvexitu složené funkce jsou obecně dosti komplikované a nebudeme je uvádět. Uvedeme jen nejdůležitější případ, kdy \mathbf{g} je afinní zobrazení.

Věta 15.5. *Necht' funkce $h: \mathbb{R}^m \rightarrow \mathbb{R}$ je konvexní. Necht' $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{b} \in \mathbb{R}^m$. Pak funkce $f(\mathbf{x}) = h(\mathbf{A}\mathbf{x} + \mathbf{b})$ je konvexní.*

Důkaz. Pro každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ a $0 \leq \alpha \leq 1$ platí

$$\begin{aligned} f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) &= h(\mathbf{A}[\alpha\mathbf{x} + (1-\alpha)\mathbf{y}] + \mathbf{b}) \\ &= h(\alpha(\mathbf{A}\mathbf{x} + \mathbf{b}) + (1-\alpha)(\mathbf{A}\mathbf{y} + \mathbf{b})) \\ &\leq \alpha h(\mathbf{A}\mathbf{x} + \mathbf{b}) + (1-\alpha)h(\mathbf{A}\mathbf{y} + \mathbf{b}) \\ &= \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y}). \end{aligned} \quad \square$$

Příklad 15.6. Funkce $f: \mathbb{R}^{2n} \rightarrow \mathbb{R}$ daná vzorcem $f(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$, kde $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, je konvexní funkce argumentu $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{2n}$. Ve Větě 15.5 vezmeme $\mathbf{A} = [\mathbf{I} \quad -\mathbf{I}] \in \mathbb{R}^{n \times (2n)}$ a $\mathbf{b} = \mathbf{0}$. \square

Nejzajímavější operace zachovávající konvexitu funkcí je maximum.

Věta 15.6. *Necht' I je libovolná množina a $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$, $i \in I$, jsou konvexní funkce. Pak*

$$f(\mathbf{x}) = \max_{i \in I} g_i(\mathbf{x}) \quad (15.7)$$

je konvexní funkce, kde předpokládáme, že pro každé \mathbf{x} maximum existuje⁴.

Důkaz. Postupujeme podobně jako v Příkladu 15.1:

$$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) = \max_{i \in I} g_i(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \quad (15.8a)$$

$$\leq \max_{i \in I} [\alpha g_i(\mathbf{x}) + (1-\alpha)g_i(\mathbf{y})] \quad (15.8b)$$

$$\leq \max_{i \in I} [\alpha g_i(\mathbf{x})] + \max_{i \in I} [(1-\alpha)g_i(\mathbf{y})] \quad (15.8c)$$

$$= \alpha \max_{i \in I} g_i(\mathbf{x}) + (1-\alpha) \max_{i \in I} g_i(\mathbf{y}) \quad (15.8d)$$

$$= \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y}). \quad (15.8e)$$

Nerovnost (15.8b) plyne z konvexity funkcí g_i . Nerovnost (15.8c) plyne z nerovnosti (15.4), kterou jsme sice dokázali jen pro konečnou množinu I ale zřejmě platí i pro nekonečnou I . Rovnost (15.8d) plyne z nezápornosti α a $1 - \alpha$. \square

Uved'me ještě jiný důkaz Věty 15.6, který využívá pojem epigrafu.

⁴ Pokud pro nějaké \mathbf{x} množina $\{g_i(\mathbf{x}) \mid i \in I\}$ nemá největší prvek (což se může stát jen tehdy, je-li množina I nekonečná), můžeme maximum v (15.7) nahradit supremem a věta stále platí.

Důkaz. Epigraf funkce f je množina

$$\begin{aligned} \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n, \max_{i \in I} g_i(\mathbf{x}) \leq y\} &= \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n, (\forall i \in I)(g_i(\mathbf{x}) \leq y)\} \\ &= \bigcap_{i \in I} \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n, g_i(\mathbf{x}) \leq y\}, \end{aligned}$$

kde jsme využili ekvivalence (11.5). Vidíme, že epigraf funkce f průnik epigrafů funkcí g_i .

Protože funkce g_i jsou konvexní, dle Věty 15.4 jsou jejich epigrafy konvexní množiny. Dle Věty 12.1 je průnik konvexních množin konvexní množina. Tedy epigraf funkce (15.7) je konvexní množina. Dle Věty 15.4 je tedy funkce f konvexní. \square

Příklad 15.7. Necht' $f(\mathbf{x}) = \max_{i=1}^n x_i$ je maximum ze složek \mathbf{x} . Konvexitu této funkce jsme dokázali z Definice 15.1, nicméně dokažme ji z Věty 15.6. Máme $g_i(\mathbf{x}) = x_i$. Funkce g_i jsou lineární, tedy konvexní. Tedy funkce $f(\mathbf{x}) = \max_{i=1}^n g_i(\mathbf{x})$ je konvexní. \square

Příklad 15.8. Funkce

$$f(\mathbf{x}) = \max_{i=1}^k (\mathbf{a}_i^T \mathbf{x} + b_i)$$

je maximum afinních funkcí. Tuto funkci jsme již potkali v §11.1.1. Protože afinní funkce jsou konvexní, je i jejich maximum konvexní. \square

Příklad 15.9. Necht' $C \subseteq \mathbb{R}^n$ je libovolná (ne nutně konvexní) množina. Funkce

$$f(\mathbf{x}) = \max_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|$$

udává vzdálenost bodu \mathbf{x} od nejbližšího bodu množiny C (zde předpokládáme, že maximum existuje). Dle Věty 15.5 je pro každé pevné \mathbf{y} výraz $\|\mathbf{x} - \mathbf{y}\|$ konvexní funkcí \mathbf{x} . Tedy výraz $\|\mathbf{x} - \mathbf{y}\|$ lze chápat jako množinu konvexních funkcí \mathbf{x} indexovaných indexem \mathbf{y} (můžeme označit $\|\mathbf{x} - \mathbf{y}\| = g_{\mathbf{y}}(\mathbf{x})$). Jelikož f je maximum těchto funkcí, je i funkce f konvexní. \square

Příklad 15.10. Mějme funkci

$$f(\mathbf{c}) = \max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b}\},$$

kteřá vyjadřuje závislost optimální hodnoty daného lineárního programu na vektoru \mathbf{c} (viz §11). Máme $f(\mathbf{c}) = \max_{\mathbf{x} \in X} \mathbf{c}^T \mathbf{x}$ a $X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$ (zde předpokládáme, že pro každé \mathbf{c} maximum existuje, neboli množina X je neprázdná a omezená). Je-li \mathbf{x} pevné, je $\mathbf{c}^T \mathbf{x}$ lineární funkce vektoru \mathbf{c} . Funkce f je tedy maximum nekonečného množství lineárních funkcí, tedy je konvexní. \square

Příklad 15.11. Necht' $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$, $b_1, \dots, b_n \in \mathbb{R}$ a $\mathbf{w} = (w_1, \dots, w_n) \in \mathbb{R}^n$ je vektor nezáporných vah. Přibližné řešení soustavy $\mathbf{a}_i^T \mathbf{x} = b_i$, $i = 1, \dots, n$, ve smyslu *vážených nejmenších čtverců* (viz §5.4) znamená vypočítat

$$f(\mathbf{w}) = \min_{\mathbf{x} \in \mathbb{R}^m} \sum_{i=1}^n w_i (\mathbf{a}_i^T \mathbf{x} - b_i)^2,$$

kde jsme označili hodnotu výsledného minima jako funkci vektoru vah. Funkce f je konkávní, protože je minimem lineárních funkcí. \square

15.4 Cvičení

15.1. Pro každou funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ dokažte z Definice 15.1, které z těchto čtyř tvrzení platí: funkce je konvexní, konkávní, konvexní i konkávní, ani konvexní ani konkávní.

- $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b$
- $f(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$
- $f(\mathbf{x}) =$ aritmetický průměr čísel x_1, \dots, x_n
- $f(\mathbf{x}) =$ median $_{i=1}^n x_i$ (medián čísel x_1, \dots, x_n)

15.2. Pro každou funkci dokažte, které z těchto čtyřech tvrzení platí: funkce je konvexní, konkávní, konvexní i konkávní, ani konvexní ani konkávní. Můžete to dokázat buď z Definice 15.1, pomocí derivací, nebo pomocí operací zachovávajících konvexitu.

- $f(x) = e^{x^2}$
- $f(x) = e^{-x^2}$
- $f(x, y) = |x - y|$
- $f(x, y) = -y$
- $f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2^2$
- $f(\mathbf{x}) = \sum_{i=1}^n x_i \log x_i$ na množině \mathbb{R}_{++}^n
- $f(\mathbf{x}) = \sum_{i=1}^k \log(b_i - \mathbf{a}_i^T \mathbf{x})$ na množině $X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T \mathbf{x} < b_i, i = 1, \dots, k\}$
- $f(\mathbf{x}) = \min_{i=1}^n |x_i|$
- $f(\mathbf{x}) = \max_{i=1}^n x_i + \min_{i=1}^n x_i$
- $f(\mathbf{x}) = \max_{i=1}^n x_i - \min_{i=1}^n x_i$
- (\star) $f(\mathbf{x}) =$ součet k největších čísel x_1, \dots, x_n (kde $k \leq n$ je dáno)

15.3. Robustní prokládání přímkou množinou bodů $(\mathbf{x}_i, y_i) \in (\mathbb{R}^n \times \mathbb{R})$, $i = 1, \dots, m$ vyžaduje minimalizaci funkce

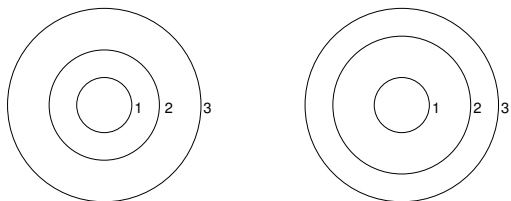
$$f(\mathbf{a}, b) = \sum_{i=1}^m \max\{-\mathbf{a}^T \mathbf{x}_i + b + y_i - \varepsilon, 0, \mathbf{a}^T \mathbf{x}_i + b - y_i - \varepsilon\},$$

kde $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$. Dokažte, že $f(\mathbf{a}, b)$ je konvexní funkce.

15.4. Je dána funkce $f(x) = -\cos x$ a množina $X = [-\pi, +\pi]$ (kde $[\cdot]$ značí uzavřený interval). Zakroužkujte pravdivá tvrzení (může jich být i více):

- Funkce f je na množině X konvexní.
- Funkce f je na množině X konkávní.
- Funkce f není na množině X ani konvexní ani konkávní.

15.5. Každý z obrázků zobrazuje některé vrstevnice funkce dvou proměnných a jejich výšky. Je možné, aby funkce, která má tyto vrstevnice, byla konvexní? Dokažte z Definice 15.1.



15.6. Co je subkontura výšky 2 funkce jedné proměnné $f(x) = x^2 - x$?

15.7. Zkuste dokázat z Definice 15.1 konvexitu či konkavitu funkcí z Příkladu 15.4. Jestliže to nesvedete, dokažte jejich konvexitu či konkavitu pomocí Vět 15.1 a 15.2.

15.8. Dokažte, že účelové funkce vystupující v následujících úlohách jsou nekonvexní:

- Příklad 10.6
- Cvičení 10.3

Nápověda a řešení

15.1.a) Konvexní i konkávní, nerovnost (15.1) platí s rovností.

15.1.b) Je konvexní, není konkávní.

15.1.c) Konvexní i konkávní, nerovnost (15.1) platí s rovností.

15.1.d) Pro $n \leq 2$ konvexní i konkávní, pro $n > 2$ ani konvexní ani konkávní.

15.5. V Definici 15.1 zvolte \mathbf{x}, \mathbf{y} na vrstevnicích výšky 1 a 3. Zvolte chytře α . Odpověď: ne, ano.

15.7. Interval $[-1, 2]$.

Kapitola 16

Konvexní optimalizační úlohy

Najít globální minimum funkce na množině je obvykle mnohem těžší než najít *nějaké* lokální minimum. Mohli bychom si myslet, že globální minimum najdeme tak, že najdeme všechna lokální minima a vybereme to, pro které je účelová funkce nejmenší. Problém je v tom, že lokálních minim může být velmi mnoho.

Příklad 16.1. Řešme úlohu

$$\max\{\mathbf{x}^T \mathbf{x} \mid \mathbf{x} \in [-1, 1]^n\}, \quad (16.1)$$

tedy maximalizujeme konvexní funkci $\mathbf{x}^T \mathbf{x}$ na hyperkrychli $[-1, 1]^n$. Je jasné (nakreslete si obrázek pro $n = 2$, tedy pro čtverec!), že funkce má lokální maximum v každém vrcholu hyperkrychle. Jelikož hyperkrychle má 2^n vrcholů (viz §12.3.3), úloha má 2^n lokálních maxim.

V tomto symetrickém případě globální maximum snadno najdeme úvahou. Uvažme však mírnou modifikaci úlohy:

$$\min\{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x} \in [-1, 1]^n\}. \quad (16.2)$$

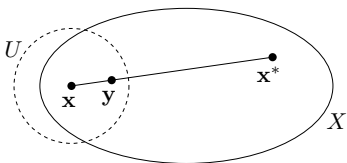
□

kde $\mathbf{A} \in \mathbb{Z}^{n \times n}$ (tedy matice má celočíselné prvky). Nalezení globálního maxima této úlohy je pravděpodobně velmi těžké, v následujícím smyslu.

Řekneme, že daný algoritmus řeší úlohu (16.2) v *polynomiálním čase*¹, jestliže existuje polynom p takový, že algoritmus pro každou matici \mathbf{A} skončí v čase menším než $p(L)$, kde L je počet bitů potřebných k zápisu všech prvků a_{ij} matice \mathbf{A} v binárním kódu. Algoritmus, který by řešil úlohu v polynomiálním čase, není znám.

Věta 16.1. *Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je konvexní na konvexní množině $X \subseteq \mathbb{R}^n$. Pak každé lokální minimum funkce f na množině X je zároveň globální.*

Důkaz. Necht' \mathbf{x} je lokálním minimem f na X , viz obrázek:



Dle Definice 9.2 tedy existuje okolí U bodu \mathbf{x} tak, že $f(\mathbf{x}) \leq f(\mathbf{y})$ pro všechna $\mathbf{y} \in U \cap X$. Necht' ale \mathbf{x} není globální minimum, tedy existuje $\mathbf{x}^* \in X$ takové, že $f(\mathbf{x}^*) < f(\mathbf{x})$. Ukážeme, že to vede ke sporu. Můžeme totiž zvolit $0 < \alpha < 1$ tak, že bod $\mathbf{y} = \alpha \mathbf{x} + (1 - \alpha)\mathbf{x}^*$ leží v U . Protože je množina X konvexní, leží bod \mathbf{y} zároveň i v X . Máme

$$f(\mathbf{y}) = f(\alpha \mathbf{x} + (1 - \alpha)\mathbf{x}^*) \leq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{x}^*) < \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{x}) = f(\mathbf{x}).$$

Ale tvrzení $f(\mathbf{y}) < f(\mathbf{x})$ je ve sporu s předpokladem, že \mathbf{x} je lokální minimum. □

Věta 16.1 tedy definuje třídu úloh, pro které nám stačí najít libovolné lokální minimum, abychom našli globální minimum. Úloze, ve které minimalizujeme konvexní funkci na konvexní množině, se říká **konvexní optimalizační úloha**.

Uvažujme nyní obecnou úlohu spojité optimalizace ve standardním tvaru (1.9),

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\} \quad (16.3)$$

neboli

$$\begin{aligned} &\min f(x_1, \dots, x_n) \\ &\text{za podmíněk } \begin{cases} g_i(x_1, \dots, x_n) \leq 0, & i = 1, \dots, m \\ h_i(x_1, \dots, x_n) = 0, & i = 1, \dots, l \end{cases} \end{aligned}$$

kde $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $(g_1, \dots, g_m) = \mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $(h_1, \dots, h_l) = \mathbf{h}: \mathbb{R}^n \rightarrow \mathbb{R}^l$. Množina přípustných řešení této úlohy je konvexní, jestliže funkce f, g_1, \dots, g_m jsou konvexní a funkce h_1, \dots, h_l jsou afinní (tedy zobrazení \mathbf{h} je afinní). Tato množina je totiž průnik množin $\{\mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \leq 0\}$ (které jsou konvexní, neboť jsou to subkontury konvexní funkce g_i) a $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$ (což je afinní podprostor, tedy také konvexní).

Podmínka, že funkce f, g_1, \dots, g_m jsou konvexní a zobrazení \mathbf{h} je afinní, je postačující ale nikoliv nutná pro konvexitu množiny přípustných řešení.

Příklad 16.2. Uvažujme dvě ekvivalentní definice téže množiny

$$\{\mathbf{x} \in \mathbb{R}^2 \mid x_1/(1 + x_2^2) \leq 0, (x_1 + x_2)^2 = 0\} = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 \leq 0, x_1 + x_2 = 0\}.$$

Oba tvary definují stejnou množinu (proč?). V prvním tvaru funkce $g(\mathbf{x}) = x_1/(1 + x_2^2)$ není konvexní a funkce $h(\mathbf{x}) = (x_1 + x_2)^2$ není afinní. Přesto je množina konvexní, což je vidět ze druhého tvaru. □

Úloze tvaru (16.3), ve které jsou funkce f, g_1, \dots, g_m konvexní a zobrazení \mathbf{h} afinní, říkáme **konvexní optimalizační úloha ve standardním tvaru**.

16.1 Ekvivalentní transformace úlohy

Dvě úlohy ve tvaru (16.3) nazveme **ekvivalentní**, když se z množiny optimálních řešení jedné dá 'snadno' (v lineárním čase) získat množina optimálních řešení druhé. **Ekvivalentní transformace** je pak každá transformace úlohy, jejímž výsledkem je úloha ekvivalentní. Dále uvedeme příklady ekvivalentních transformací. U každé poznamenejme, zda zachovává konvexitu úlohy.

¹Zde se dotýkáme *teorie složitosti algoritmů*, kterou budete brát až později.

Změna proměnných

Nechť $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ je bijektivní zobrazení (viz §1.1.2). Pak úloha (16.3) je ekvivalentní úloze

$$\min\{f(\varphi(\mathbf{x})) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{g}(\varphi(\mathbf{x})) \leq \mathbf{0}, \mathbf{h}(\varphi(\mathbf{x})) = \mathbf{0}\}.$$

Tato transformace nemusí zachovat konvexitu úlohy.

Monotónní transformace účelové funkce

Nechť $\psi: \mathbb{R} \rightarrow \mathbb{R}$ je monotónní rostoucí funkce. Pak

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in X\} = \min\{\psi(f(\mathbf{x})) \mid \mathbf{x} \in X\}.$$

Tato transformace nemusí zachovat konvexitu funkce f .

Příklad 16.3. Tuto transformaci jsme již několikrát použili v nejmenších čtvercích. Máme minimalizovat např. funkci $f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2$, ale zvolíme $\psi(y) = y^2$ a minimalizujeme funkci $\psi(f(\mathbf{x})) = \|\mathbf{Ax} - \mathbf{b}\|_2^2 = (\mathbf{Ax} - \mathbf{b})^T(\mathbf{Ax} - \mathbf{b})$. Nová funkce má výhodu, že je na rozdíl od staré diferencovatelná, a to při zachování konvexity. \square

Slackové proměnné

Podobně jako v lineárním programování, úloha (16.3) je ekvivalentní úloze

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{s} \in \mathbb{R}^m, \mathbf{s} \geq \mathbf{0}, \mathbf{g}(\mathbf{x}) + \mathbf{s} = \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}.$$

Transformace zachová konvexitu úlohy jen v případě, kdy $\mathbf{g}(\mathbf{x}) + \mathbf{s}$ je afinní zobrazení vektoru (\mathbf{x}, \mathbf{s}) , tedy kdy zobrazení \mathbf{g} je afinní,

Epigrafový tvar

Úlohu (16.3) je ekvivalentní úloze

$$\min\{y \mid \mathbf{x} \in \mathbb{R}^n, y \in \mathbb{R}, f(\mathbf{x}) - y \leq 0, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}.$$

Tato transformace zachovává konvexitu úlohy. Plyne z ní důležité pozorování, že každou optimalizační úlohu lze převést na úlohu s lineární účelovou funkcí.

16.2 Třídy konvexních optimalizačních úloh

Optimalizační úlohy ve tvaru (16.3) se taxonomizují podle druhu funkcí f, g_i, h_i . Pro každou třídu existují specializované algoritmy schopné najít lokální minimum².

16.2.1 Lineární programování (LP)

V lineárním programování jsou všechny funkce f, g_i, h_i afinní. Jde tedy v jistém smyslu o nej-jednodušší případ konvexní optimalizační úlohy. Přesto jsme viděli v Kapitole 11, že již tento jednoduchý případ má velmi mnoho aplikací.

² Viz např. <http://www.neos-guide.org>.

16.2.2 Kvadratické programování (QP)

V kvadratickém programování jsou funkce g_i, h_i afinní a funkce f je kvadratická konvexní, tedy $f(\mathbf{x}) = \mathbf{x}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{x} + c$ kde \mathbf{A} je pozitivně semidefinitní (viz Příklad 15.5).

Příklad 16.4. Při řešení soustavy ve smyslu nejmenších čtverců (§5.1) počítáme konvexní QP bez omezení $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_2^2$. Tuto úlohu lze všelijak modifikovat, např. můžeme přidat omezení $\mathbf{c} \leq \mathbf{x} \leq \mathbf{d}$, tj. každá proměnná x_j musí být v intervalu $[c_j, d_j]$. To vede na konvexní QP s omezeními. \square

Příklad 16.5. Hledání řešení přeuročené lineární soustavy s nejmenší normou (§5.2) vede na úlohu $\min\{\|\mathbf{x}^T \mathbf{x} \mid \mathbf{Ax} = \mathbf{b}\}$, což je konvexní QP. \square

Příklad 16.6. Chceme spočítat vzdálenost $d(P_1, P_2) = \min\{\|\mathbf{x}_1 - \mathbf{x}_2\|_2 \mid \mathbf{x}_1 \in P_1, \mathbf{x}_2 \in P_2\}$ dvou konvexních mnohostěnů $P_1 = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}_1 \mathbf{x} \leq \mathbf{b}_1\}$, $P_2 = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}_2 \mathbf{x} \leq \mathbf{b}_2\}$. Úloha vede na QP

$$d(P_1, P_2)^2 = \min\{\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2 \mid \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n, \mathbf{A}_1 \mathbf{x}_1 \leq \mathbf{b}_1, \mathbf{A}_2 \mathbf{x}_2 \leq \mathbf{b}_2\}. \quad \square$$

Příklad 16.7. Je dáno m bodů v \mathbb{R}^n , z nichž každý patří do jedné ze dvou tříd, označených -1 a 1 . Jinými slovy, je dána množina dvojic $(\mathbf{x}_i, y_i) \in \mathbb{R}^n \times \{-1, 1\}$ pro $i = 1, \dots, m$. V úloze lineární klasifikace hledáme nadrovinu, která odděluje body z obou tříd. Tedy hledáme $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$ takové, aby

$$\begin{aligned} \mathbf{a}^T \mathbf{x}_i - b &< 0 & \text{pro } y_i = -1, \\ \mathbf{a}^T \mathbf{x}_i - b &> 0 & \text{pro } y_i = 1, \end{aligned}$$

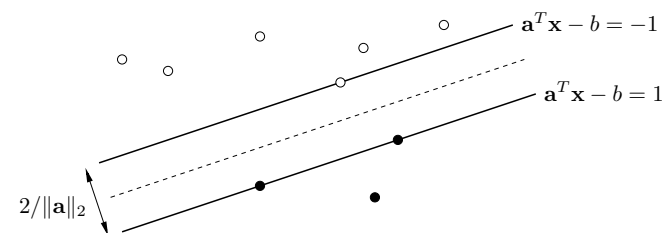
což lze napsat jako

$$y_i(\mathbf{a}^T \mathbf{x}_i - b) > 0, \quad i = 1, \dots, m. \quad (16.4)$$

Označme $\varepsilon_i = y_i(\mathbf{a}^T \mathbf{x}_i - b)$ a vydělme vektor (\mathbf{a}, b) kladným číslem $\min_{i=1}^m \varepsilon_i$. Pak soustavu (16.4) můžeme ekvivalentně psát jako

$$y_i(\mathbf{a}^T \mathbf{x}_i - b) \geq 1, \quad i = 1, \dots, m. \quad (16.5)$$

Hledáme-li libovolnou oddělovací nadrovinu, stačí nám najít libovolné řešení soustavy nerovnic (16.5). Soustava ale navíc říká, že body jsou odděleny pásem $\{\mathbf{x} \in \mathbb{R}^n \mid -1 \geq \mathbf{a}^T \mathbf{x} - b \geq 1\}$:



Snadno spočítáme (srov. Cvičení .a), že šířka pásu je $2/\|\mathbf{a}\|_2$. V úloze support vector machine (SVM) najít rozdělovací nadrovinu která maximalizuje šířku pásu, tedy minimalizuje $\|\mathbf{a}\|_2^2 = \mathbf{a}^T \mathbf{a}$ za podmínek (16.5). To je úloha QP. \square

16.2.3 Kvadratické programování s kvadratickými omezeními (QCQP)

Obecnější variantou je *kvadratické programování s kvadratickými omezeními* (QCQP, *quadratically constrained quadratic programming*), kde funkce f, g_i jsou kvadratické konvexní a funkce h_i jsou afinní.

16.2.4 Programování na kuželu druhého řádu (SOCP)

V úloze *programování na kuželu druhého řádu* (SOCP, *second-order cone programming*) jsou funkce f, h_i afinní a funkce g_i mají tvar

$$g_i(\mathbf{x}) = \|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\|_2 - (\mathbf{c}_i^T \mathbf{x} + d_i). \quad (16.6)$$

Tedy úloha SOCP má tvar (vynecháváme afinní omezení $h_i(\mathbf{x}) = 0$)

$$\begin{aligned} \min \quad & \mathbf{e}^T \mathbf{x} \\ \text{za podmíněk} \quad & \|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\|_2 \leq \mathbf{c}_i^T \mathbf{x} + d_i, \quad i = 1, \dots, m. \end{aligned}$$

Funkce g_i jsou konvexní (neboť norma je konvexní funkce, dále viz Věta 15.5). Podmínku $g_i(\mathbf{x}) \leq 0$ lze psát také jako $(\mathbf{A}_i \mathbf{x} + \mathbf{b}_i, \mathbf{c}_i^T \mathbf{x} + d_i) \in K_2^n$, kde konvexní množina

$$K_2^n = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 \leq y\}$$

je epigraf Euklidovské normy $\|\cdot\|_2$ (která je konvexní funkce, viz Příklad 15.4), kterému se také říká *kužel druhého řádu*.

Pro $\mathbf{A}_i = \mathbf{0}$ se podmínka $g_i(\mathbf{x}) \leq 0$ stane lineární nerovnicí. Pro $\mathbf{c}_i = \mathbf{0}$ se podmínka $g_i(\mathbf{x}) \leq 0$ po umocnění na druhou stane konvexní kvadratická. Tedy LP a QCQP jsou speciální případy SOCP.

Příklad 16.8. Jsou dány body $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ a chceme minimalizovat funkci

$$f(\mathbf{x}) = \sum_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\|_2 \quad (16.7)$$

přes $\mathbf{x} \in \mathbb{R}^n$. Řešení této úlohy je známo jako *geometrický medián*. Pro $n = 1$ se funkce redukuje na $f(x) = \sum_{i=1}^m |x - a_i|$, jejímž minimem je obyčejný medián.

Úloha je konvexní, ale nelze ji převést na LP, QP ani QCQP. Po zavedení pomocných proměnných z_i (podobná úprava jako v §11.1.1) ji lze formulovat jako SOCP:

$$\begin{aligned} \min \quad & z_1 + \dots + z_m \\ \text{za podmíněk} \quad & \|\mathbf{x} - \mathbf{a}_i\|_2 \leq z_i, \quad i = 1, \dots, m. \end{aligned}$$

Pro případ $n = 2$ má úloha jednoduchý mechanický model³. Do vodorovného prkna vyvrtáme díry o souřadnicích \mathbf{a}_i . Každou dírou provlečeme provázek. Provázky jsou nahoře svázané uzlem do jednoho bodu a dole mají závaží o stejné hmotnosti. Poloha uzlu je \mathbf{x} . Hodnota $f(\mathbf{x})$ je potenciální energie soustavy a ustálený stav odpovídá minimu $f(\mathbf{x})$. □

³Toto mechanické zařízení, známé jako *Varignon frame*, se v minulosti opravdu používalo na řešení úlohy. Úloha má bohatou historii, je známa také jako Fermat-Weberův problém.

16.2.5 Semidefinitní programování (SDP)

Věta 16.2. Množina všech pozitivně semidefinitních matic rozměru $n \times n$ je konvexní kužel.

Důkaz. Necht' pro $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ a $\mathbf{x} \in \mathbb{R}^n$ je $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$ a $\mathbf{x}^T \mathbf{B} \mathbf{x} \geq 0$. Pak pro každé $\alpha, \beta \geq 0$

$$\mathbf{x}^T (\alpha \mathbf{A} + \beta \mathbf{B}) \mathbf{x} = \alpha \mathbf{x}^T \mathbf{A} \mathbf{x} + \beta \mathbf{x}^T \mathbf{B} \mathbf{x} \geq 0. \quad \square$$

Konvexní kužel je konvexní množina. To umožňuje formulovat třídu konvexních úloh známou jako *semidefinitní programování* (SDP). Jednou z možných formulací je

$$\begin{aligned} \min \quad & \mathbf{C} \cdot \mathbf{X} \\ \text{za podmíněk} \quad & \mathbf{X} \text{ je pozitivně semidefinitní} \\ & \mathbf{A}_i \cdot \mathbf{X} = b_i, \quad i = 1, \dots, m. \end{aligned} \quad (16.8)$$

kde matice $\mathbf{C}, \mathbf{A}_i \in \mathbb{R}^{n \times n}$ a skaláry b_i jsou dány a optimalizujeme přes pozitivně semidefinitní matice $\mathbf{X} \in \mathbb{R}^{n \times n}$. Operace

$$\mathbf{A} \cdot \mathbf{X} = \sum_{i=1}^n \sum_{j=1}^n A_{ij} X_{ij}$$

označuje skalární součin matic.

SDP je velmi obecná třída konvexních úloh. Lze ukázat, že konvexní úlohy LP, QP, QCQP i SOCP jsou speciální případy SDP. Pro ilustraci ukážeme, že pokud matice \mathbf{C}, \mathbf{A}_i jsou diagonální, úloha (16.8) se redukuje na LP. V tom případě v součinech $\mathbf{C} \cdot \mathbf{X}$ a $\mathbf{A}_i \cdot \mathbf{X}$ nediagonální prvky matice \mathbf{X} nehrají žádnou roli. Diagonální matice je pozitivně semidefinitní právě tehdy, když všechny její prvky jsou nezáporné (viz Cvičení 6.19). Tedy úloha (16.8) se redukuje na

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq \mathbf{0}, \mathbf{a}_i^T \mathbf{x} = b_i \forall i = 1, \dots, m \},$$

kde vektory $\mathbf{c}, \mathbf{a}_i \in \mathbb{R}^n$ jsou diagonály matic \mathbf{C}, \mathbf{A}_i .

Některé konvexní úlohy nepatří do žádné z uvedených tříd.

Příklad 16.9. *Analytický střed soustavy nerovnic* $X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} \geq \mathbf{b} \}$ je bod, který maximalizuje funkci

$$f(\mathbf{x}) = \sum_{i=1}^m \log(\mathbf{a}_i^T \mathbf{x} - b_i) \quad (16.9)$$

na vnitřku množiny X , kde \mathbf{a}_i^T jsou řádky matice \mathbf{A} . Pokud je mnohostěn X neprázdný a omezený, úloha má maximum a toto maximum je jediné.

Každou nerovnici $\mathbf{a}_i^T \mathbf{x} \geq b_i$ vydělme nezáporným číslem tak, aby $\|\mathbf{a}_i\|_2 = 1$. Nyní $\mathbf{a}_i^T \mathbf{x} - b_i$ je vzdálenost bodu \mathbf{x} od nadroviny $\mathbf{a}_i^T \mathbf{x} = b_i$. Číslo $f(\mathbf{x})$ je tedy logaritmus součinu vzdáleností bodu \mathbf{x} od nadrovin. Analytický střed tedy minimalizuje součin těchto vzdáleností. □

16.3 Cvičení

16.1. Významnou vlastností konvexních funkcí je to, že každé lokální minimum funkce je zároveň globální (Věta 16.1). Ne každá funkce s touto vlastností je ovšem konvexní. Člověk by si mohl myslet, že součet dvou funkcí (ne nutně konvexních) s touto vlastností bude mít tuto vlastnost také. Je toto tvrzení pravdivé? Odpověď dokažte.

16.2. Dokažte, že množina optimálních řešení konvexní optimalizační úlohy je konvexní.

16.3. Mějme úlohu

$$\min\{f(x, y) \mid x, y \geq 0, 2x + y \geq 1, x + 3y \geq 1\}.$$

Nakreslete množinu přípustných řešení. Pro každou z následujících účelových funkcí najděte úvahou množinu optimálních řešení a optimální hodnotu:

- $f(x, y) = x + y$
- $f(x, y) = x$
- $f(x, y) = \min\{x, y\}$
- $f(x, y) = \max\{x, y\}$
- $f(x, y) = |x + y|$
- $f(x, y) = x^2 + 9y^2$

V kterých případech se jedná o konvexní optimalizační úlohu?

16.4. Uvažujme úlohu, známou jako *lineární lomené programování*:

$$\begin{aligned} \min \quad & (\mathbf{c}^T \mathbf{x} + d) / (\mathbf{e}^T \mathbf{x} + f) \\ \text{za podmíněk} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ & \mathbf{e}^T \mathbf{x} + f > 0 \end{aligned}$$

- Je účelová funkce na množině přípustných řešení konvexní?
- Dokažte, že úloha je ekvivalentní lineárnímu programu (s proměnnými \mathbf{x}, z)

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{y} + dz \\ \text{za podmíněk} \quad & \mathbf{A}\mathbf{y} \geq \mathbf{b}z \\ & \mathbf{e}^T \mathbf{y} + fz = 1 \\ & z \geq 0 \end{aligned}$$

16.5. Najděte explicitní řešení pro následující úlohy QCQP (\mathbf{A}, \mathbf{B} jsou pozitivně definitní):

- $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{A} \mathbf{x} \leq 1\}$
- $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, (\mathbf{x} - \mathbf{b})^T \mathbf{A} (\mathbf{x} - \mathbf{b}) \leq 1\}$
- $\min\{\mathbf{x}^T \mathbf{B} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{A} \mathbf{x} \leq 1\}$

16.6. Formulujte úlohu $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_4$ jako konvexní QCQP.

16.7. Dokažte, že pro libovolný vektor $\mathbf{x} \in \mathbb{R}^n$ a skaláry $y \geq 0, z \geq 0$ platí

$$\mathbf{x}^T \mathbf{x} \leq yz \iff \|(2\mathbf{x}, y - z)\|_2 \leq y + z.$$

Uvažujte úlohu, kdy maximalizujeme harmonický průměr afinních funkcí, tedy funkci

$$f(\mathbf{x}) = \left(\sum_{i=1}^m (\mathbf{a}_i^T \mathbf{x} - b_i)^{-1} \right)^{-1}$$

za podmíněk $\mathbf{a}_i^T \mathbf{x} > b_i$. Je tato úloha (po možné jednoduché transformaci) konvexní? Vyjádřete úlohu jako SOCP pomocí dokázané ekvivalence.

16.8. (*) Máme konvexní funkci jedné proměnné $f: \mathbb{R} \rightarrow \mathbb{R}$. Dáme do grafu funkce žebřík o délce 1 tak, aby oba konce ležely na grafu. Předpokládáme-li, že tření mezi žebříkem a grafem je nulové, zaujme žebřík stav lokálního minima potenciální energie (která je přímo úměrná výšce středu žebříku). Zformulujte jako optimalizační úlohu. Bude tato úloha konvexní? Zkuste najít situaci, kdy potenciální energie bude mít více než jedno lokální minimum.

Nápověda a řešení

16.4.a) Ne.

16.4.b) Uvažujte substituci $\mathbf{y} = \mathbf{x}/(\mathbf{e}^T \mathbf{x} + f), z = 1/(\mathbf{e}^T \mathbf{x} + f)$.

16.5.a) Viz Cvičení 11.5.

16.5.b) Substituuje $\mathbf{y} = \mathbf{x} - \mathbf{b}$.

16.5.c) Optimální hodnota je nula.

16.7. Místo maximalizace funkce $f(\mathbf{x})$ minimalizujme funkci $1/f(\mathbf{x})$, která je konvexní na množině přípustných hodnot. Úloha je ekvivalentní úloze

$$\begin{aligned} \min \quad & t_1 + \dots + t_m \\ \text{za podmíněk} \quad & t_i (\mathbf{a}_i^T \mathbf{x} - b_i) \geq 1, \quad i = 1, \dots, m \\ & t_i \geq 0, \quad i = 1, \dots, m. \end{aligned}$$

Použitím dokázané ekvivalence převedeme na SOCP

$$\begin{aligned} \min \quad & t_1 + \dots + t_m \\ \text{za podmíněk} \quad & \|(2, \mathbf{a}_i^T \mathbf{x} - b_i - t_i)\|_2 \geq \mathbf{a}_i^T \mathbf{x} - b_i + t_i, \quad i = 1, \dots, m \\ & \mathbf{a}_i^T \mathbf{x} \geq b_i, \quad i = 1, \dots, m \\ & t_i \geq 0, \quad i = 1, \dots, m. \end{aligned}$$

Kapitola 17

Lagrangeova dualita

Zatímco v §14 jsme odvodili dualitu pro lineární programování, zde popíšeme základ teorie duality pro obecné optimalizační úlohy. Dualita v LP se pak bude jevit jako speciální případ.

17.1 Minimální nerovnost

Pro libovolné množiny X a Y a libovolnou funkci $L: X \times Y \rightarrow \mathbb{R}$ platí **minimaxní nerovnost**

$$\min_{x \in X} \max_{y \in Y} L(x, y) \geq \max_{y \in Y} \min_{x \in X} L(x, y). \quad (17.1)$$

Zde předpokládáme, že všechna minima a maxima existují¹. V nerovnosti (17.1) nastává rovnost právě tehdy, když existuje bod $(x^*, y^*) \in X \times Y$ takový, že

$$L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*) \quad \forall x \in X, y \in Y. \quad (17.2)$$

Takovému bodu (x^*, y^*) říkáme **sedlový bod** funkce L na $X \times Y$.

Uvedené skutečnosti se snadno dokáží, podrobné důkazy vynecháme. Pro důkaz druhého tvrzení je užitečné si uvědomit, že podmínku (17.2) lze psát jako $F(x^*) = G(y^*)$.

Příklad 17.1. Necht' $X = Y = \{1, 2, 3, 4\}$.

| | 1 | 2 | 3 | 4 | $F(x)$ | | 1 | 2 | 3 | 4 | $F(x)$ |
|--------|----|---|---|----|--------|--------|----|---|---|----|--------|
| 1 | -1 | 4 | 7 | 4 | 7 | 1 | -1 | 4 | 7 | 4 | 7 |
| 2 | 4 | 4 | 6 | -2 | 6 | 2 | 4 | 4 | 6 | -2 | 6 |
| 3 | 1 | 5 | 3 | 3 | 5 | 3 | 1 | 0 | 3 | 3 | 3 |
| 4 | 3 | 5 | 3 | 2 | 5 | 4 | 3 | 3 | 3 | 2 | 3 |
| $G(y)$ | -1 | 4 | 3 | -2 | | $G(y)$ | -1 | 0 | 3 | -2 | |

Funkce L v levé tabulce nemá sedlový bod a máme $\min_{x \in X} F(x) > \max_{y \in Y} G(y)$. Funkce L v pravé tabulce má sedlový bod (dokonce dva, v rámečcích) a $\min_{x \in X} F(x) = \max_{y \in Y} G(y)$. \square

¹Kdyby ne, mohli bychom min/max nahradit inf/sup a nerovnost by stále platila.

17.2 Lagrangeova duální úloha

Ke každé optimalizační úloze (kterou nazýváme primární) lze sestavit jinou optimalizační úlohu (nazývanou duální) tak, že mezi nimi platí více či méně užitečné vztahy. Jedna forma duality se získá následovně: chytře zvolíme množiny X, Y a funkci L tak, aby levá strana minimaxní nerovnosti (17.1) byla primární (tedy původní) úloha. Pravá strana pak bude duální úloha.

Necht' množina $X \subseteq \mathbb{R}^n$, funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a zobrazení $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ jsou dány; jak dále uvidíme, reprezentují primární úlohu. Zvolme $Y = \mathbb{R}_+^m$ a

$$L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + \mathbf{y}^T \mathbf{g}(\mathbf{x}). \quad (17.3)$$

Podívejme se, jak po této volbě vypadá levá strana nerovnosti (17.1), tedy primární úloha. Je

$$F(\mathbf{x}) = \max_{\mathbf{y} \in Y} L(\mathbf{x}, \mathbf{y}) = \max_{\mathbf{y} \geq \mathbf{0}} [f(\mathbf{x}) + \mathbf{y}^T \mathbf{g}(\mathbf{x})] = \begin{cases} f(\mathbf{x}) & \text{když } \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\ \infty & \text{jinak,} \end{cases} \quad (17.4)$$

kde ' ∞ ' značí, že úloha $\max_{\mathbf{y} \in Y} L(\mathbf{x}, \mathbf{y})$ je neomezená. Abyste to uviděli, promyslete si, že pro každý vektor $\mathbf{a} \in \mathbb{R}^m$ platí

$$\max_{\mathbf{y} \geq \mathbf{0}} \mathbf{y}^T \mathbf{a} = \begin{cases} 0 & \text{když } \mathbf{a} \leq \mathbf{0}, \\ \infty & \text{jinak.} \end{cases}$$

Po dosazení (17.4) do levé strany (17.1) tedy

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} L(\mathbf{x}, \mathbf{y}) = \min \{ f(\mathbf{x}) \mid \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leq \mathbf{0} \}. \quad (17.5)$$

Toto je naše primární úloha. Všimněme si, jak jsme omezující podmínky $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ zahrnuli do funkce F za cenu, že funkce F může nabývat i nekonečných hodnot.

Nyní se podívejme na duální úlohu. Ta zní

$$\max_{\mathbf{y} \in Y} \min_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{y}) = \max_{\mathbf{y} \geq \mathbf{0}} G(\mathbf{y}), \quad (17.6)$$

kde

$$G(\mathbf{y}) = \min_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{x} \in X} [f(\mathbf{x}) + \mathbf{y}^T \mathbf{g}(\mathbf{x})]. \quad (17.7)$$

Přesnější tvar duální úlohy nezískáme, dokud nebudeme mít zadánu primární úlohu (tedy X, f, \mathbf{g}) konkrétněji. Nyní můžeme říci jen to, že úloha (17.7) může být pro nějaká \mathbf{y} neomezená a tedy funkce G nabývat hodnot $-\infty$. To bude opět reprezentovat omezení duální úlohy.

Všimněte si, že pro každé \mathbf{x} je funkce L afinní funkcí proměnné \mathbf{y} . Dle Věty 15.6 je tedy duální funkce G konkávní a tedy duální úloha bude maximalizace konkávní funkce, tedy konvexní úloha. To platí vždy, i když primární úloha není konvexní.

Právě sestavená úloha (17.6) se nazývá **Lagrangeova duální úloha** k (primární) úloze (17.5). Funkce (17.3) se říká Lagrangeova funkce – všimněte si, že je to ta samá funkce, kterou jste potkali u metody Lagrangeových multiplikátorů.

17.3 Silná dualita

Z nerovnosti (17.1) plyne, že optimální hodnota primární úlohy není menší než optimální hodnota duální úlohy. Tato skutečnost je známa jako věta o **slabé dualitě**. Rozdílu mezi primární a duální optimální hodnotou se říká **dualitní mezera**. Když v nerovnosti (17.1) nastane rovnost, jsou si optimální hodnoty primární a duální úlohy rovny neboli dualitní mezera je nulová. V tom případě říkáme, že pro naši úlohu platí **silná dualita**.

Silná dualita může platit pro velice různé úlohy. Uvedeme nyní, ve Větě 17.1, jednu postačující (avšak nikoliv nutnou) podmínku, za které platí silná dualita.

Řekneme, že funkce $g_1, \dots, g_m: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině X splňují *Slaterovu podmínku*, když existuje vnitřní bod \mathbf{x} množiny X takový, že $\mathbf{g}(\mathbf{x}) < \mathbf{0}$, neboli

$$g_1(\mathbf{x}) < 0, \dots, g_m(\mathbf{x}) < 0. \quad (17.8)$$

Pokud je prvních $k \leq m$ funkcí g_i afinních, podmínku (17.8) lze změkčit na

$$g_1(\mathbf{x}) \leq 0, \dots, g_k(\mathbf{x}) \leq 0, g_{k+1}(\mathbf{x}) < 0, \dots, g_m(\mathbf{x}) < 0. \quad (17.9)$$

Věta 17.1. *Necht'*

- množina $X \subseteq \mathbb{R}^n$ je konvexní,
- funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a $g_1, \dots, g_m: \mathbb{R}^n \rightarrow \mathbb{R}$ jsou konvexní na X ,
- funkce g_1, \dots, g_m na množině X splňují Slaterovu podmínku.

Pak platí silná dualita, neboli optimální hodnoty úloh (17.5) a (17.6) jsou si rovny.

Dále uveďme obdobu věty o komplementaritě.

Věta 17.2. *Necht' $\mathbf{x} \in X$ je optimum primární úlohy a $\mathbf{y} \in \mathbb{R}_+^m$ je optimum duální úlohy a nastává silná dualita. Pak platí podmínky komplementarity*

$$y_i g_i(x_i) = 0 \quad \forall i = 1, \dots, m. \quad (17.10)$$

Důkaz. Platí

$$F(\mathbf{x}) = f(\mathbf{x}) = G(\mathbf{y}) = \min_{\mathbf{x}' \in X} [f(\mathbf{x}') + \mathbf{y}^T \mathbf{g}(\mathbf{x}')] \leq f(\mathbf{x}) + \mathbf{y}^T \mathbf{g}(\mathbf{x}) \leq f(\mathbf{x}).$$

Druhá rovnost plyne ze silné duality. Třetí rovnost je definice duální úlohy. První nerovnost plyne z definice minima. Druhá nerovnost platí, protože $\mathbf{y} \geq \mathbf{0}$ a $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$.

Ale protože $f(\mathbf{x})$ je na začátku i na konci řetězce nerovností, musí obě nerovnosti být rovnostmi, $f(\mathbf{x}) + \mathbf{y}^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x})$. Z toho plyne

$$\mathbf{y}^T \mathbf{g}(\mathbf{x}) = 0. \quad (17.11)$$

To je ale ekvivalentní podmínkám (17.10), protože $\mathbf{y} \geq \mathbf{0}$ a $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$. \square

17.4 Příklady

Příklad 17.2. Necht' $f(\mathbf{x}) = \mathbf{c}^T \mathbf{x}$, $\mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}$, $X = \mathbb{R}^n$. Primární úloha je

$$\min_{\mathbf{x} \in \mathbb{R}^n} \max_{\mathbf{y} \geq \mathbf{0}} L(\mathbf{x}, \mathbf{y}) = \min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b} \}. \quad (17.12)$$

Odvoďme duální úlohu. Lagrangeova funkce je

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{c}^T \mathbf{x} + \mathbf{y}^T (\mathbf{b} - \mathbf{A}\mathbf{x}) = (\mathbf{c}^T - \mathbf{y}^T \mathbf{A})\mathbf{x} + \mathbf{y}^T \mathbf{b}, \quad (17.13)$$

tedy

$$G(\mathbf{y}) = \min_{\mathbf{x} \in \mathbb{R}^n} L(\mathbf{x}, \mathbf{y}) = \begin{cases} \mathbf{y}^T \mathbf{b} & \text{když } \mathbf{c}^T - \mathbf{y}^T \mathbf{A} = \mathbf{0}, \\ \infty & \text{jinak.} \end{cases}$$

Duální úloha tedy je

$$\max_{\mathbf{y} \geq \mathbf{0}} G(\mathbf{y}) = \max \{ \mathbf{b}^T \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^m, \mathbf{A}^T \mathbf{y} = \mathbf{c}, \mathbf{y} \geq \mathbf{0} \}. \quad (17.14)$$

To je stejný výsledek, jaký bychom dostali podle návodu v §14.1 na konstrukci duálního LP. Konkrétně, primární a duální úloha je dvojice úloh (14.1). \square

Příklad 17.3. Necht' $f(\mathbf{x}) = \mathbf{c}^T \mathbf{x}$, $\mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}$, $X = \mathbb{R}_+^n$. Primární úloha je

$$\min_{\mathbf{x} \in \mathbb{R}_+^n} \max_{\mathbf{y} \geq \mathbf{0}} L(\mathbf{x}, \mathbf{y}) = \min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}_+^n, \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}. \quad (17.15)$$

Lagrangeova funkce je (17.13). Je

$$G(\mathbf{y}) = \min_{\mathbf{x} \in \mathbb{R}_+^n} L(\mathbf{x}, \mathbf{y}) = \begin{cases} \mathbf{y}^T \mathbf{b} & \text{když } \mathbf{c}^T - \mathbf{y}^T \mathbf{A} \geq \mathbf{0}, \\ -\infty & \text{jinak.} \end{cases}$$

Duální úloha tedy je

$$\max_{\mathbf{y} \geq \mathbf{0}} G(\mathbf{y}) = \max \{ \mathbf{b}^T \mathbf{y} \mid \mathbf{A}^T \mathbf{y} \leq \mathbf{c}, \mathbf{y} \geq \mathbf{0} \}. \quad (17.16)$$

To je opět stejný výsledek, jaký bychom dostali podle návodu v §14.1. \square

Zatím jsme v primární úloze uvažovali jen omezení typu nerovnosti. Omezení typu rovnosti $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ lze nahradit dvěma omezeními typu nerovnosti $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ a $-\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$. Pak

$$L(\mathbf{x}, \mathbf{y}_+, \mathbf{y}_-) = f(\mathbf{x}) + \mathbf{y}_+^T \mathbf{g}(\mathbf{x}) - \mathbf{y}_-^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x}) + (\mathbf{y}_+ - \mathbf{y}_-)^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x}) + \mathbf{y}^T \mathbf{g}(\mathbf{x}), \quad (17.17)$$

kde jsme označili $\mathbf{y} = \mathbf{y}_+ - \mathbf{y}_-$. Funkci (17.17) můžeme nyní přejmenovat na $L(\mathbf{x}, \mathbf{y})$, která má stejný tvar jako (17.3). Podmínky $\mathbf{y}_+, \mathbf{y}_- \geq \mathbf{0}$ se v rozdílu zruší, tedy $\mathbf{y} \in \mathbb{R}^m$. Ověříme, že

$$F(\mathbf{x}) = \max_{\mathbf{y} \in \mathbb{R}^m} L(\mathbf{x}, \mathbf{y}) = \begin{cases} f(\mathbf{x}) & \text{když } \mathbf{g}(\mathbf{x}) = \mathbf{0}, \\ \infty & \text{jinak,} \end{cases}$$

tedy primární úloha je

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in \mathbb{R}^m} L(\mathbf{x}, \mathbf{y}) = \min \{ f(\mathbf{x}) \mid \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) = \mathbf{0} \}, \quad (17.18)$$

jak jsme chtěli. Duální úloha je

$$\max_{\mathbf{y} \in \mathbb{R}^m} \min_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{y}) = \max_{\mathbf{y} \in \mathbb{R}^m} G(\mathbf{y}). \quad (17.19)$$

Všimněte si, že Slaterova podmínka nebude platit, když zobrazení \mathbf{g} nebude afinní.

Příklad 17.4. Napišme duální úlohu k úloze

$$\min\{\mathbf{x}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}\}.$$

Máme $X = \mathbb{R}^n$ a

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{x} + \mathbf{y}^T (\mathbf{b} - \mathbf{A}\mathbf{x}) = \mathbf{x}^T \mathbf{x} - \mathbf{y}^T \mathbf{A}\mathbf{x} + \mathbf{y}^T \mathbf{b},$$

duální funkce je tedy

$$G(\mathbf{y}) = \min_{\mathbf{x} \in \mathbb{R}^n} L(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{x} \in \mathbb{R}^n} (\mathbf{x}^T \mathbf{x} - \mathbf{y}^T \mathbf{A}\mathbf{x} + \mathbf{y}^T \mathbf{b}).$$

Řešení musí splňovat $\partial L(\mathbf{x}, \mathbf{y})/\partial \mathbf{x} = \mathbf{0}$, což dá $\mathbf{x} = \mathbf{A}^T \mathbf{y}/2$. Po dosazení dostaneme

$$G(\mathbf{y}) = L(\mathbf{A}^T \mathbf{y}/2, \mathbf{y}) = \mathbf{b}^T \mathbf{y} - \mathbf{y}^T \mathbf{A} \mathbf{A}^T \mathbf{y}/4. \quad \square$$

Příklad 17.5. Řešme lineární program

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}.$$

Duál bychom mohli sestavit podle návodu v kapitole o dualitě v LP. Ale postupujme jinak. Zvolme Lagrangeovu funkci jako (17.12) a $X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}$. Duální funkce bude

$$G(\mathbf{y}) = \min_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{0} \leq \mathbf{x} \leq \mathbf{1}} [(\mathbf{c}^T - \mathbf{y}^T \mathbf{A})\mathbf{x} + \mathbf{y}^T \mathbf{b}] = \mathbf{b}^T \mathbf{y} + \mathbf{1}^T \max\{\mathbf{0}, \mathbf{A}^T \mathbf{y} - \mathbf{c}\}$$

neboť $\min_{\mathbf{0} \leq \mathbf{x} \leq \mathbf{1}} \mathbf{d}^T \mathbf{x} = \min\{\mathbf{0}, \mathbf{d}\} = -\max\{\mathbf{0}, -\mathbf{d}\}$ (kde \min a \max se rozumí po složkách). \square

Příklad 17.6. Řešme celočíselný lineární program

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \{0, 1\}^n, \mathbf{A}\mathbf{x} = \mathbf{b}\}.$$

Necht' Lagrangeova funkce je (17.12) a necht' $X = \{0, 1\}^n$. Duální úloha je stejná jako v minulém příkladě. Zde předpoklady Věty 17.1 neplatí, protože množina X není konvexní. Opravdu, silná dualita u celočíselného programování obecně neplatí. \square

Silná dualita může nastat (i když spíše vyjimečně) i pro nekonvexní úlohu. Jednou takovou třídou úloh je libovolné (tedy ne nutně konvexní) QCQP s nejvýše *jedním* omezením.

Příklad 17.7. Úloha QCQP

$$\min\{\mathbf{x}^T \mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\} \quad (17.20)$$

není konvexní, pokud matice \mathbf{A} není pozitivně semidefinitní. Řešení ale najdeme snadno pomocí spektrálního rozkladu:

$$\min_{\mathbf{x}^T \mathbf{x} = 1} \mathbf{x}^T \mathbf{A}\mathbf{x} = \min_{\mathbf{x}^T \mathbf{x} = 1} \mathbf{x}^T \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T \mathbf{x} = \min_{\mathbf{z}^T \mathbf{z} = 1} \mathbf{z}^T \mathbf{\Lambda}\mathbf{z} = \lambda_{\min}(\mathbf{A}).$$

Ukážeme navíc, že platí silná dualita. Máme

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A}\mathbf{x} + \mathbf{y}(1 - \mathbf{x}^T \mathbf{x}) = \mathbf{x}^T (\mathbf{A} - \mathbf{y}\mathbf{I})\mathbf{x} + \mathbf{y}.$$

Tedy

$$G(\mathbf{y}) = \begin{cases} \mathbf{y} & \text{když } \mathbf{A} - \mathbf{y}\mathbf{I} \text{ je pozitivně semidefinitní,} \\ -\infty & \text{jinak.} \end{cases}$$

Ale matice $\mathbf{A} - \mathbf{y}\mathbf{I}$ je pozitivně semidefinitní právě tehdy, když nejmenší vlastní číslo matice $\mathbf{A} - \mathbf{y}\mathbf{I}$ je nezáporné, neboli (viz Cvičení 6.4) $\lambda_{\min}(\mathbf{A}) \geq \mathbf{y}$. Duální úloha tedy je

$$\max\{\mathbf{y} \in \mathbb{R} \mid \lambda_{\min}(\mathbf{A}) \geq \mathbf{y}\}.$$

Ta má očividně optimální řešení $\mathbf{y} = \lambda_{\min}(\mathbf{A})$. \square

Kapitola 18

Vícekritériální optimalizace

18.1 Uspořádání na množině

Binární relace na množině Y je množina $R \subseteq Y \times Y$. Binární relace je

- **reflexivní**, když $(x, x) \in R$ pro každé $x \in Y$,
- **transitivní**, když $(x, y) \in R$ a $(y, z) \in R$ implikuje $(x, z) \in R$,
- **antisymetrická**, když $(x, y) \in R$ a $(y, x) \in R$ implikuje $x = y$.

Částečné uspořádání (krátce jen **uspořádání**) na množině Y je binární relace na Y , která je reflexivní, transitivní a antisymetrická.

Relace uspořádání se obvykle značí infixově symbolem \preceq , tedy místo $(x, y) \in R$ píšeme $x \preceq y$. Pokud potřebujeme rozlišit více různých kvasi-uspořádání, používáme symboly jako $\leq_1, \leq_2, \leq', \leq'$ atd.

Prvky $x, y \in Y$ jsou **srovnatelné** v uspořádání \preceq , když $x \preceq y$ nebo $y \preceq x$ nebo obojí. (Kvasi-)uspořádání je **úplné** (neboli **totální**), když každé dva prvky z Y jsou srovnatelné.

Příklad 18.1.

- $Y = \mathbb{R}$ a \preceq je přirozené uspořádání reálných čísel. Tato relace je úplné uspořádání.
- $Y \subseteq 2^U$ a \preceq je relace inkluze na množině 2^U , tedy $x \preceq y$ právě když $x \subseteq y$. Zde U je libovolná množina a 2^U značí množinu všech jejích podmnožin. Tato relace je uspořádání, ale není úplné.
- $Y = \mathbb{N}$ a \preceq je relace dělitelnosti, tj. $x \preceq y$ právě když x dělí y . Tato relace je uspořádání, není úplné.
- $\mathbf{x} \preceq \mathbf{y}$ právě když $\sum_{i=1}^n x_i \leq \sum_{i=1}^n y_i$. Tato relace není antisymetrická, tedy není uspořádání. \square

Nás ovšem nejvíce zajímá případ $Y = \mathbb{R}^n$.

Příklad 18.2. Příklady uspořádání na množině \mathbb{R}^n :

- Uspořádání 'po složkách': $\mathbf{x} \preceq \mathbf{y}$ právě když $x_i \leq y_i$ pro všechna $i = 1, \dots, n$. Toto uspořádání není úplné: např. pro $n = 2$ jsou vektory $\mathbf{x} = (0, 1)$ a $\mathbf{y} = (1, 0)$ nesrovnatelné.
- Lexikografické uspořádání: $\mathbf{x} \preceq \mathbf{y}$ právě když

$$(\mathbf{x} = \mathbf{y}) \text{ nebo } (\exists m)(x_m < y_m)(\forall i < m)(x_i = y_i).$$

Toto uspořádání je úplné.

- Definujme $\mathbf{x} \preceq \mathbf{y}$ právě když $\sum_{i=1}^n x_i \leq \sum_{i=1}^n y_i$. Tato relace není uspořádání, protože není antisymetrická. \square

Definice 18.1. Prvek $x \in Y$ se nazývá (vzhledem ke kvasi-uspořádání \preceq)

- **minimální prvek** množiny Y , když $y \preceq x$ implikuje $x \preceq y$, pro všechna $y \in Y$.
- **nejmenší prvek** množiny Y , když $x \preceq y$, pro všechna $y \in Y$.

Pro totální kvasi-uspořádání oba pojmy splývají.

18.2 Úlohy vícekritériální optimalizace

V klasické optimalizaci jsme se zabývali úlohami typu

$$\min_{x \in X} f(x), \quad (18.1)$$

kde X je množina přípustných řešení a $f: X \rightarrow \mathbb{R}$ je účelová (neboli kritériální) funkce. Optimální hodnoty této úlohy jsou minimální prvky množiny $f(X) = \{f(x) \mid x \in X\} \subseteq \mathbb{R}$. Zde pojem 'minimální prvek' se myslí vzhledem k přirozenému uspořádání na \mathbb{R} .

Zobecníme tuto úlohu. Necht' $f: X \rightarrow Y$ a necht' \preceq je (kvasi-)uspořádání na množině Y . Pak úlohou (18.1) budeme rozumět nalezení minimálních prvků množiny $f(X) \subseteq Y$ vzhledem k uspořádání \preceq . Případně pro každý minimální prvek y množiny $f(X)$ chceme najít argument $x \in X$, ve kterém se nabývá, tedy splňujícím $y = f(x)$.

Nejčastěji v aplikacích potkáme případ $Y = \mathbb{R}^n$. Pak mluvíme o *vícekritériální optimalizaci*¹, neboť vlastně chceme minimalizovat více skalárních kritérií (složek zobrazení $\mathbf{f}: X \rightarrow \mathbb{R}^n$, hodnoty zobrazení jsou vektory a tedy ho píšeme tučně) najednou. Dále se omezíme pouze na tento případ.

Příklad 18.3. V obchodě nabízejí čtyři druhy aut s těmito vlastnostmi:

| | | VW Golf | Opel Astra | Ford Focus | Toyota Corolla |
|----------|-------------|---------|------------|------------|----------------|
| cena | [tis. euro] | 16 | 15 | 14 | 15 |
| spotřeba | [l/100km] | 7.2 | 7.0 | 7.5 | 8.2 |

Chceme levné auto s malou spotřebou. Která auta je dobré si koupit a která naopak nekoupit?

Máme $X = \{\text{VW Golf, Opel Astra, Ford Focus, Toyota Corolla}\}$ a $Y = \mathbb{R}^2$. Tabulka definuje zobrazení \mathbf{f} . Není ovšem jasné, jaké (kvasi-)uspořádání na množině \mathbb{R}^2 použít pro rozhodování.

Rozhodujme se vzhledem k uspořádání 'po složkách' \leq^2 . Vzhledem k tomuto uspořádání nemá množina $\mathbf{f}(X)$ nejmenší prvek, neboli kritéria jsou v konfliktu. Její minimální prvky jsou $\mathbf{f}(\text{Opel Astra}) = (15, 7.0)$ a $\mathbf{f}(\text{Ford Focus}) = (14, 7.5)$.

Rozhodujme se vzhledem k lexikografickému uspořádání, přesněji nejprve se rozhodujme dle ceny a pak dle spotřeby. Nyní minimální prvek množiny $\mathbf{f}(X)$ je $\mathbf{f}(\text{Ford Focus}) = (14, 7.5)$. \square

¹ Angl. *multicriteria optimization*. Názvosloví ovšem není jednotné, jindy se používají názvy *multiobjective optimization* nebo *vector optimization* (neboť hodnoty zobrazení f jsou vektory).

Příklad 18.4. Chceme řešit (přesně či přibližně) nehomogenní lineární soustavu $\mathbf{Ax} \approx \mathbf{b}$, ale zároveň chceme, aby velikost vektoru \mathbf{x} byla malá. Řešíme tedy úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^n} (\|\mathbf{Ax} - \mathbf{b}\|_2, \|\mathbf{x}\|_2).$$

Jaké jsou minimální prvky množiny $\mathbf{f}(\mathbb{R}^2)$ vzhledem k uspořádání po složkách? Je jich nekonečně mnoho, obrázek. Vezmeme-li lexikografické uspořádání a je-li soustava přeuročená, dostaneme metodu nejmenší normy. \square

Příklad 18.5. V okrese je n vesnic se souřadnicemi $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^2$. Do jakého místa $\mathbf{x} \in \mathbb{R}^2$ máme umístit heliport, aby byl ke všem vesnicím blízko?

Máme $X = \mathbb{R}^2$, $Y = \mathbb{R}^n$, a $f_i(\mathbf{x}) = \|\mathbf{a}_i - \mathbf{x}\|_2$. Řešíme úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^2} (\|\mathbf{a}_1 - \mathbf{x}\|_2, \dots, \|\mathbf{a}_n - \mathbf{x}\|_2).$$

Součástí úlohy není (kvasi-)uspořádání na množině \mathbb{R}^n . Jaká (kvasi-)uspořádání jsou vhodná?

Uspořádání po složkách: množina minimálních prvků množiny $\mathbf{f}(\mathbb{R}^2)$ je konvexní obal bodů $\mathbf{a}_1, \dots, \mathbf{a}_n$. To je intuitivně jasné (i když přesný důkaz vynecháme): neleží-li \mathbf{x} v tomto konvexním obalu, můžeme změnit \mathbf{x} tak, že se vzdálenost k některým bodům \mathbf{a}_i zmenší a k ostatním se nezmění. Tedy je hloupost umístit heliport mimo tento konvexní obal.

Max-uspořádání: vede na úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^2} \max_{i=1}^n \|\mathbf{a}_i - \mathbf{x}\|_2.$$

Toto je úloha klasické (skalární) optimalizace. Této formulaci se někdy říká *minimaxní* formulace. Minimalizujeme vzdálenost heliportu od nejbvzdálenější vesnice. \square