

Optimalizace

Elektronická skripta předmětu A4B33OPT.

Text může být v průběhu semestru měněn, toto je verze z **12. ledna 2018**.

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	7
1.1	Značení	7
1.1.1	Množiny	7
1.1.2	Zobrazení	8
1.1.3	Funkce a zobrazení více reálných proměnných	9
1.1.4	Extrémy funkce na množině	9
1.2	Úloha spojité optimalizace	10
1.3	Cvičení	12
I	Použití lineární algebry v optimalizaci	13
2	Maticová algebra	14
2.1	Binární operace s maticemi	14
2.2	Transpozice a symetrie	15
2.3	Hodnost	15
2.4	Inverze	16
2.5	Determinant	16
2.6	Stopa	17
2.7	Matice s jedním sloupcem nebo jedním řádkem	17
2.8	Matice sestavené z bloků	18
2.9	Zločiny při práci s maticemi	19
2.10	Cvičení	20
3	Linearita	23
3.1	Lineární podprostory	23
3.2	Lineární zobrazení	24
3.2.1	Prostor obrazů	26
3.2.2	Nulový prostor	27
3.3	Afinní podprostor a zobrazení	28
3.4	Cvičení	30
4	Ortogonalita	34
4.1	Standardní skalární součin	34
4.2	Ortogonální podprostory	34
4.3	Ortonormální množina vektorů	36
4.4	Matice s ortonormálními sloupci	37

4.5	Ortogonalní projekce	38
4.5.1	Projekce na ortogonální doplněk	39
4.5.2	Vzdálenost bodu od podprostoru	40
4.6	Gram-Schmidtova ortonormalizace	40
4.6.1	QR rozklad	41
4.7	Skalární součin a norma matic	41
4.8	Cvičení	42
5	Nehomogenní lineární soustavy	46
5.1	Přibližné řešení ve smyslu nejmenších čtverců	46
5.1.1	Ortogonalní projekce na podprostor daný obecnou bází	48
5.1.2	Řešení pomocí QR rozkladu	48
5.1.3	Lineární regrese	49
5.1.4	Statistické odůvodnění kritéria nejmenších čtverců	50
5.1.5	Vícekritériální nejmenší čtverce, regularizace	50
5.2	Řešení s nejmenší normou	51
5.3	(*) Pseudoinverze obecné matice	52
5.4	Cvičení	53
6	Spektrální rozklad a kvadratické funkce	57
6.1	Vlastní čísla a vektory	57
6.1.1	Spektrální rozklad	58
6.2	Kvadratická forma	59
6.3	Kvadratická funkce	62
6.3.1	Kvadratika	63
6.4	Cvičení	64
7	Proložení bodů podprostorem	68
7.1	Problém nejmenší stopy	68
7.2	Proložení bodů podprostorem	69
7.2.1	Nejbližší matice nižší hodnosti	71
7.2.2	Co když prokládáme afinním podprostorem?	71
7.3	Přeurčené homogenní lineární soustavy	72
7.4	Singulární rozklad	73
7.4.1	Nejbližší matice nižší hodnosti	74
7.5	Cvičení	75
II	Nelineární optimalizace	78
8	Nelineární funkce a zobrazení	79
8.1	Spojitosť	80
8.2	Derivace funkce jedné proměnné	81
8.3	Parciální derivace	81
8.4	Derivace zobrazení	82
8.4.1	Derivace složeného zobrazení	83
8.4.2	Derivace maticových výrazů	84
8.5	Směrová derivace	85

8.6	Gradient	87
8.7	Parciální derivace druhého řádu	87
8.8	Taylorův polynom	88
8.9	Vnitřek a hranice množiny	90
8.10	Lokální extrémy funkce na množině	91
8.11	Cvičení	92
9	Volné lokální extrémy	96
9.1	Analytické podmínky	96
9.1.1	Pro jednu proměnnou	96
9.1.2	Pro více proměnných	97
9.2	Sestupné metody	98
9.3	Gradientní metoda	98
9.3.1	(★) Závislost na lineární transformaci souřadnic	99
9.4	Newtonova metoda	99
9.4.1	Použití na soustavy nelineárních rovnic	100
9.4.2	Použití na minimalizaci funkce	101
9.5	Nelineární metoda nejmenších čtverců	102
9.5.1	Gauss-Newtonova metoda	102
9.5.2	Rozdíl oproti Newtonově metodě	103
9.5.3	Levenberg-Marquardtova metoda	104
9.6	Cvičení	104
10	Lokální extrémy vázané rovnostmi	107
10.1	Tečný a ortogonální prostor k povrchu	108
10.2	Podmínky prvního řádu	109
10.3	(★) Podmínky druhého řádu	112
10.4	Cvičení	114
III	Lineární programování	117
11	Lineární programování	118
11.1	Transformace úloh LP	119
11.1.1	Po částech afinní funkce	120
11.2	Optimální výrobní program	121
11.3	Směšovací (výživová) úloha	122
11.4	Dopravní úloha	123
11.5	Přeurčené lineární soustavy	124
11.5.1	Vektorové normy	124
11.5.2	Přibližné řešení přeurčených lineárních soustav	125
11.5.3	Lineární regrese	125
11.6	Celočíselné lineární programování, LP relaxace	127
11.6.1	Nejlepší přiřazení	128
11.6.2	Nejmenší vrcholové pokrytí	129
11.6.3	Největší nezávislá množina	130
11.7	Cvičení	130

12 Konvexní množiny a mnohostěny	135
12.1 Konvexní množiny	135
12.2 Čtyři druhy kombinace vektorů	136
12.3 Konvexní mnohostěny	136
12.3.1 (★) Extrémní body	138
12.3.2 (★) Extrémy lineární funkce na mnohostěnu	139
12.4 Cvičení	141
13 Simplexová metoda	144
13.1 Stavební kameny algoritmu	145
13.1.1 Přechod k sousední standardní bázi	145
13.1.2 Kdy je nové báze řešení přípustné?	146
13.1.3 Co když je celý sloupec nekladný?	147
13.1.4 Ekvivalentní úpravy účelového řádku	147
13.1.5 Co udělá přechod k sousední bázi s účelovou funkcí?	148
13.2 Základní algoritmus	148
13.3 Inicializace algoritmu	150
13.3.1 Dvoufázová simplexová metoda	151
13.4 Cvičení	153
14 Dualita v lineárním programování	156
14.1 Konstrukce duální úlohy	156
14.2 Věty o dualitě	157
14.3 Příklady na konstrukci a interpretaci duálních úloh	160
14.3.1 Mechanické modely	163
14.4 Cvičení	166
IV Konvexní optimalizace	168
15 Konvexní funkce	169
15.1 Vztah konvexní funkce a konvexní množiny	171
15.2 Konvexita diferencovatelných funkcí	172
15.3 Operace zachovávající konvexitu funkcí	173
15.3.1 Nezáporná lineární kombinace	173
15.3.2 Skládání funkcí	174
15.3.3 Maximum	174
15.4 Cvičení	176
16 Konvexní optimalizační úlohy	178
16.1 Příklady nekonvexních úloh	178
16.2 Konvexní optimalizační úloha ve standardním tvaru	179
16.3 Ekvivalentní transformace úlohy	180
16.4 Třídy konvexních optimalizačních úloh	181
16.4.1 Lineární programování (LP)	181
16.4.2 Kvadratické programování (QP)	181
16.4.3 Kvadratické programování s kvadratickými omezeními (QCQP)	182
16.4.4 Programování na kuželu druhého řádu (SOCP)	182

16.4.5 Semidefinitní programování (SDP)	183
16.5 Konvexní relaxace nekonvexních úloh	184
16.6 Cvičení	184
17 (★) Lagrangeova dualita	187
17.1 Minimaxní nerovnost	187
17.2 Lagrangeova duální úloha	188
17.3 Silná dualita	189
17.4 Příklady	190
Rejstřík	191

Kapitola 1

Značení a základní pojmy

1.1 Značení

Pokud potkáte ve skriptech slovo vysázené **tučně**, jde o nově zavedený pojem, který máte pochopit a zapamatovat si. Ve skriptech tedy nepotkáte slovo ‘definice’, protože všechna tučná slova jsou definice. Slova vysázená *kurzívou* znamenají buď zdůraznění, nebo nový avšak všeobecně známý pojem. Odstavce, věty, důkazy, příklady a cvičení označené hvězdičkou (★) jsou rozšiřující (a často zajímavé) a není nezbytné umět je ke zkoušce.

Zopakujme 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
$A_1 \times \dots \times A_n$	kartézský součin množin, množina $\{(a_1, \dots, a_n) \mid a_1 \in A_1, \dots, a_n \in A_n\}$
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 (angl. *mapping*, *map*) 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 $b = f(a) \in B$. *Funkce* (angl. *function*) znamená přesně totéž jako zobrazení, slovo *funkce* se však obvykle používá pouze pro zobrazení do číselných množin (tedy $B = \mathbb{R}, \mathbb{Z}, \mathbb{C}$ apod.). *Transformace* je zobrazení nějaké množiny do sebe, tedy $f: A \rightarrow A$.

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

$$f(A') = \{f(a) \mid a \in A'\} = \{b \in B \mid b = f(a), a \in A'\} \quad (1.2)$$

(čárka v definici pravé množiny značí konjunkci, tedy jde o množinu všech $b \in B$ které splňují $b = f(a)$ a zároveň $a \in A'$). 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 $f(A') = \{a^2 \mid a \in A'\} = \{1, 9, 16\}$. 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 jen $\{f(a) \mid \varphi(a)\}$, pokud je A jasné z kontextu. Např. $\{x^2 \mid x \in \mathbb{R}, -1 < x < 1\} = [0, 1)$, kde $[0, 1)$ značí polouzavřený interval $[0, 1)$.

Zobrazení se nazývá:

- *injektivní* (neboli prosté), jestliže každý vzor má jiný obraz, tj. $f(a) = f(a') \Rightarrow a = a'$.
- *surjektivní* (neboli A na B), jestliže každý obraz má aspoň jeden vzor, tj. $f(A) = B$.
- *bijektivní* (neboli vzájemně jednoznačné), jestliže je zároveň injektivní a surjektivní.

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 $a \in A$.

¹Přesná 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 $b = b'$.

1.1.3 Funkce a zobrazení více reálných proměnných

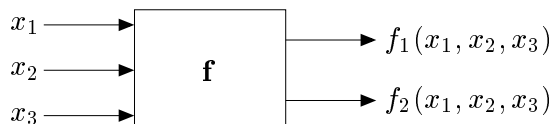
Uspořádané n -tici $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ reálných čísel říkáme (n -rozměrný) **vektor**. Zápis

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

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 . Takový prvek x také nazýváme *minimální argument* funkce na množině. 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 minimálních argumentů funkce f na množině X 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 minimu 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 z 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 $\max_{i=1}^5 a_i = 3$, $\operatorname{argmax}_{i=1}^5 a_i = \{3, 5\}$. □

1.2 Úloha spojitě optimalizace

Matematickou *optimalizací* v obecném smyslu rozumíme hledání minima nějaké funkce $f: X \rightarrow \mathbb{R}$ na nějaké množině X . Tato formulace je velmi obecná, neboť množina X může být zcela libovolná. Tento kurs se zabývá *spojitou optimalizací*, ve které máme $X \subseteq \mathbb{R}^n$ a množina X má nespočetný počet prvků a je popsána jako množina řešení soustavy rovnic a nerovnic. Tedy X je množina všech vektorů $(x_1, \dots, x_n) \in \mathbb{R}^n$ splňujících

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

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

pro dané funkce $g_1, \dots, g_m, h_1, \dots, h_l: \mathbb{R}^n \rightarrow \mathbb{R}$. Funkce f, g_i, h_i musejí být navíc 'hezké' (spojité), abychom mohli použít aparát matematické analýzy. 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:

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

Toto je tedy *úloha spojitě optimalizace v obecném tvaru*.

Příklad 1.2. Pastevce 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).

Označme strany obdélníka x, y . Řešíme úlohu

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

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

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

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

\square

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

- Funkce f se nazývá *účelová* (také pokutová, cenová, kritériální) 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* nebo *optimální argumenty* úlohy. Číslo $\min_{\mathbf{x} \in X} f(\mathbf{x})$ se říká *optimální hodnota* úlohy.
- 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*.

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.

- a) $\min\{x^2 + y^2 \mid x > 0, y > 0, xy \geq 1\}$
- b) $\min\{(x-2)^2 + (y-1)^2 \mid x^2 \leq 1, y^2 \leq 1\}$
- c) 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.
- d) Jaké má rozměry válec s jednotkovým objemem a nejmenším povrchem?
- e) Najděte rozměry püllitru, na jehož výrobu je třeba co nejméně skla. Tloušťka stěn je zanedbatelná.
- f) Najděte obsah největšího obdélníka vepsaného do kružnice s poloměrem 1.
- g) 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?
- h) Najděte bod v rovině na parabole s rovnicí $y = x^2$ nejbližší bodu $(3, 0)$.
- i) 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?
- j) x, y jsou čísla v intervalu $[1, 5]$ taková, že jejich součet je 6. Najděte tato čísla tak, aby xy^2 bylo (a) co nejmenší a (b) co největší.
- k) 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 .
- l) *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. Necht' X je libovolná množina a $f: X \rightarrow \mathbb{R}$. Necht' $g: \mathbb{R} \rightarrow \mathbb{R}$ je rostoucí funkce. Dokažte, že $\operatorname{argmin}_{x \in X} f(x) = \operatorname{argmin}_{x \in X} g(f(x))$.

Část I

Použití lineární algebry v optimalizaci

Kapitola 2

Maticová algebra

Cílem této kapitoly je zopakovat si základní maticové pojmy a naučit se manipulovat s maticovými výrazy a rovnicemi, aniž byste toho zatím museli mnoho vědět o lineární algebře.

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 matici 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 Binární operace s maticemi

V algebře reálných matic se reálná čísla nazývají také **skaláry**¹. Na maticích jsou definovány následující operace:

- Součin skaláru $\alpha \in \mathbb{R}$ a matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ je matice $\alpha \mathbf{A} = \mathbf{A} \alpha = [\alpha a_{ij}] \in \mathbb{R}^{m \times n}$.

Součin $\frac{1}{\alpha} \mathbf{A}$ píšeme krátce jako $\frac{\mathbf{A}}{\alpha}$ nebo \mathbf{A}/α . Součin $(-1)\mathbf{A}$ píšeme krátce jako $-\mathbf{A}$.

¹Přesněji, pohlížíme-li na množinu všech matic rozměru $m \times n$ jako na lineární prostor, jedná se o skaláry tohoto lineárního prostoru.

- Součet matic $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ je matice $\mathbf{A} + \mathbf{B} = [a_{ij} + b_{ij}] \in \mathbb{R}^{m \times n}$. Rozdíl matic je $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$.
- **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

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

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

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 násobení matice skalárem je jiná operace, než maticový součin.

Pro čtvercovou matici \mathbf{A} značí \mathbf{A}^k maticový součin k matic \mathbf{A} .

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 Hodnost

Hodnost (angl. *rank*) matice je dimenze lineárního obalu jejích sloupců. Značíme ji $\text{rank } \mathbf{A}$. Platí

$$\text{rank } \mathbf{A} = \text{rank}(\mathbf{A}^T), \quad (2.2)$$

tedy hodnost je také rovna dimenzi lineárního obalu řádků. Důkaz této rovnosti není jednoduchý, najdete ho v každé učebnici lineární algebry.

Z (2.2) plyne, že pro matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ je

$$\text{rank } \mathbf{A} \leq \min\{m, n\}. \quad (2.3)$$

Když $\text{rank } \mathbf{A} = \min\{m, n\}$, říkáme, že matice má **plnou hodnost**. 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 s plnou hodnotí se nazývá **regulární**. Čtvercová matice, která nemá plnou hodnotu, se nazývá **singulární**.

Je užitečné si pamatovat nerovnost (důkaz viz Cvičení 3.13)

$$\text{rank}(\mathbf{AB}) \leq \min\{\text{rank } \mathbf{A}, \text{rank } \mathbf{B}\}. \quad (2.4)$$

2.4 Inverze

Když platí

$$\mathbf{AB} = \mathbf{I}, \quad (2.5)$$

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 3.5). Levá inverze matice existuje právě tehdy, má-li matice lineárně nezávislé sloupce (což plyne z předchozího, když si (2.5) napíšeme jako $\mathbf{B}^T \mathbf{A}^T = \mathbf{I}$; viz také Věta 3.7).

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} . Matice má inverzi, právě když je regulární. 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

$$\det \mathbf{A} = \sum_{\sigma} \text{sgn } \sigma \prod_{i=1}^n a_{i\sigma(i)}, \quad (2.6)$$

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

- $\det \mathbf{I} = 1$
- $\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})$
- $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1}$ (plyne z předchozího pro $\mathbf{B} = \mathbf{A}^{-1}$)
- $\det \mathbf{A}^T = \det \mathbf{A}$
- $\det \mathbf{A} = 0$ právě tehdy, když \mathbf{A} je singulární
- 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 změní znaménko determinantu.

Dá se dokázat, že první dvě vlastnosti determinant jednoznačně určují: existuje právě jedna funkce $f: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ taková, že $f(\mathbf{I}) = 1$ a $f(\mathbf{AB}) = f(\mathbf{A})f(\mathbf{B})$ pro libovolné \mathbf{A}, \mathbf{B} .

2.6 Stopa

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

$$\text{tr } \mathbf{A} = a_{11} + \dots + a_{nn}. \quad (2.7)$$

Vlastnosti (dokažte!):

- $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr } \mathbf{A} + \text{tr } \mathbf{B}$
- $\text{tr}(\alpha \mathbf{A}) = \alpha \text{tr } \mathbf{A}$
- $\text{tr}(\mathbf{A}^T) = \text{tr } \mathbf{A}$
- $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ pro každé $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{B} \in \mathbb{R}^{n \times m}$ (tzv. *cykličnost stopy*)

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. $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{CBA})$.

2.7 Matice s jedním sloupcem nebo jedním řádkem

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

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

$$\mathbf{x} = \underbrace{(x_1, \dots, x_m)}_{\text{uspořádaná } m\text{-tice}} = \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}}_{\text{matice } m \times 1} \in \mathbb{R}^m$$

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).

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

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

- Pro $\mathbf{x} \in \mathbb{R}^m$ a $\mathbf{y} \in \mathbb{R}^n$ 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, $[\mathbf{e}_1 \ \dots \ \mathbf{e}_n] = \mathbf{I}_n$.

2.8 Matice sestavené z bloků

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

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \quad [\mathbf{A} \ \mathbf{B}], \quad \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}. \quad (2.8)$$

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

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

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

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

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

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

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

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} [\mathbf{b}_1 \ \dots \ \mathbf{b}_n] = [\mathbf{Ab}_1 \ \dots \ \mathbf{Ab}_n] = \begin{bmatrix} \mathbf{a}_1^T \mathbf{B} \\ \vdots \\ \mathbf{a}_m^T \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T \mathbf{b}_1 & \dots & \mathbf{a}_1^T \mathbf{b}_n \\ \vdots & \ddots & \vdots \\ \mathbf{a}_m^T \mathbf{b}_1 & \dots & \mathbf{a}_m^T \mathbf{b}_n \end{bmatrix}. \quad (2.9)$$

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

$$\mathbf{AB} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_p] \begin{bmatrix} \mathbf{b}_1^T \\ \vdots \\ \mathbf{b}_p^T \end{bmatrix} = \mathbf{a}_1 \mathbf{b}_1^T + \dots + \mathbf{a}_p \mathbf{b}_p^T. \quad (2.10)$$

2.9 Zločiny při práci s maticemi

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

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 je použití zlomku pro matice, např. $\frac{\mathbf{A}}{\mathbf{B}}$. ‘Zlomková čára’ není pro matice definována. Nebyla by totiž jednoznačná, protože může znamenat buď $\mathbf{A}\mathbf{B}^{-1}$ nebo $\mathbf{B}^{-1}\mathbf{A}$.
- Inverze čtvercové, ale evidentně singulární matice. Např. $(\mathbf{w}\mathbf{w}^T)^{-1}$, kde $\mathbf{w} \in \mathbb{R}^3$. Singulárिता matice plyne např. z (2.4).
- Předpoklad existence levé inverze široké matice. Ale široká matice má lineárně závislé sloupce, proto nemá levou inverzi. Podobně úzká 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 zločin.

Příklad 2.2. Vidíme-li výraz $(\mathbf{A}^T \mathbf{B})^{-1}$, musí nám okamžitě hlavou proběhnout tyto úvahy o rozměrech matic $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{B} \in \mathbb{R}^{k \times p}$:

- Aby byly kompatibilní velikosti matic v násobení, musí být $m = k$. (Výjimkou je případ, kdy \mathbf{A} je skalár nebo \mathbf{B} je skalár – pak by $\mathbf{A}^T \mathbf{B}$ byl sice nezvyklý ale korektní zápis.)
- Jelikož součin $\mathbf{A}^T \mathbf{B}$ je rozměru $n \times p$, musí být $n = p$, protože invertovat můžeme jen čtvercovou matici. Ted’ tedy víme, že obě matice musí mít stejný rozměr.
- Z nerovnosti (2.4) je jasné, že pokud by \mathbf{A}^T byla úzká nebo \mathbf{B} široká, $\mathbf{A}^T \mathbf{B}$ by byla singulární a tedy by neměla inverzi. Abychom se tomu vyhnuli, musí být obě matice buď čtvercové nebo úzké, $m \geq n$.

Závěr: Aby výraz $(\mathbf{A}^T \mathbf{B})^{-1}$ měl smysl, je nutné, aby matice \mathbf{A}, \mathbf{B} měly stejný rozměr a byly čtvercové nebo úzké. \square

Použití neexistujících maticových identit

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

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

Neekvivalentní úpravy (ne)rovníc

Zde pachatel 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 ne $a^2 = b^2 \Rightarrow a = b$. Příklady:

- Student si myslí, že $\mathbf{a}^T \mathbf{x} = \mathbf{a}^T \mathbf{y} \Rightarrow \mathbf{x} = \mathbf{y}$ (není pravda, ani když vektor \mathbf{a} je nemulový).
- Student si myslí, že pokud $\mathbf{A} \in \mathbb{R}^{3 \times 5}$ a $\mathbf{AX} = \mathbf{AY}$, pak $\mathbf{X} = \mathbf{Y}$ (není pravda, protože \mathbf{A} nemá lineárně nezávislé sloupce, tedy nemá levou inverzi).
- Student si myslí, že $\mathbf{A}^T \mathbf{A} = \mathbf{B}^T \mathbf{B} \Rightarrow \mathbf{A} = \mathbf{B}$ (není pravda dokonce ani pro skaláry).
- Student řeší soustavu rovnic $\{\mathbf{a}^T \mathbf{x} = 0, \mathbf{x}^T \mathbf{x} = 1\}$ pro neznámou \mathbf{x} (tedy $\mathbf{a} \in \mathbb{R}^n$ je dáno a $\mathbf{x} \in \mathbb{R}^n$ je neznámá). Dělá to tak, že ‘vyjádří’ \mathbf{x} z první rovnice a dosadí ho do druhé rovnice. To je těžký zločin, protože rovnice $\mathbf{a}^T \mathbf{x} = 0$ má nekonečně mnoho řešení a tedy z ní neplyne, že \mathbf{x} je rovno jakémukoliv jednomu vektoru.

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. `A=randn(5,3); B=randn(3,6); (A*B)'\B'*A'`. Samozřejmě to není důkaz.

2.10 Cvičení

2.1. Vyřešte tyto rovnice a soustavy rovnic 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 $\mathbf{b}_i = \mathbf{X}\mathbf{a}_i$ (kde $i = 1, \dots, k$) pro neznámou matici $\mathbf{X} \in \mathbb{R}^{m \times n}$. Napište soustavu jako jedinou maticovou rovnici. 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 `A=randn(m,n); b=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 $\mathbf{P}\mathbf{u} = \mathbf{q}$.
 b) Je-li $\mathbf{a}, \mathbf{x} \in \mathbb{R}^m$, $\mathbf{b}, \mathbf{y} \in \mathbb{R}^n$, ukažte, že $\mathbf{x} = (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}(\mathbf{a} - \mathbf{B}\mathbf{D}^{-1}\mathbf{b})$. Jakou to má výpočetní výhodu oproti počítání \mathbf{u} přímo ze soustavy $\mathbf{P}\mathbf{u} = \mathbf{q}$?

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?

- a) $\mathbf{A}\mathbf{x} = \mathbf{b}$, neznámá \mathbf{x} .
 b) $\mathbf{x}^T \mathbf{A} \mathbf{x} = 1$, neznámá \mathbf{x} .
 c) $\mathbf{a}^T \mathbf{X} \mathbf{b} = 0$, neznámá \mathbf{X} .
 d) $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^T = \mathbf{C}$, neznámá \mathbf{X} .
 e) $\{\mathbf{X}^T \mathbf{Y} = \mathbf{A}, \mathbf{X}\mathbf{Y}^T = \mathbf{B}\}$, 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

$$\text{vec}(\mathbf{A}\mathbf{B}\mathbf{C}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec } \mathbf{B}. \quad (2.11)$$

Použijte tohoto vzorce pro převedení následujících soustav rovnic s neznámou maticí \mathbf{X} do tvaru $\mathbf{P}\mathbf{u} = \mathbf{q}$ 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.

- a) $\{\mathbf{b}_i^T \mathbf{X} \mathbf{a}_i = 0, i = 1, \dots, k\}$
 b) $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^T = \mathbf{C}$

2.7. Dokažte, že rovnice $\mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A} = \mathbf{I}$ nemá řešení pro žádné \mathbf{A}, \mathbf{B} .

2.8. *Komutátorem* dvou matic rozumíme matici $[\mathbf{A}, \mathbf{B}] = \mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A}$. Dokažte:

- a) Komutátor symetrických matic je antisymetrická matice.
 b) $[\mathbf{A}, [\mathbf{B}, \mathbf{C}]] + [\mathbf{B}, [\mathbf{C}, \mathbf{A}]] + [\mathbf{C}, [\mathbf{A}, \mathbf{B}]] = \mathbf{0}$ (*Jacobiho identita*)
 c) $[\mathbf{A}, \mathbf{B}\mathbf{C}] = [\mathbf{A}, \mathbf{B}]\mathbf{C} + \mathbf{B}[\mathbf{A}, \mathbf{C}]$

2.9. Dokažte *Sherman-Morrisonův vzorec* (kde \mathbf{A} je regulární a $\mathbf{v}^T \mathbf{A}^{-1} \mathbf{u} + 1 \neq 0$)

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

a *Sherman-Morrison-Woodburyho vzorec* (kde $\mathbf{I} + \mathbf{V}^T \mathbf{A}^{-1} \mathbf{U}$ je regulární)

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

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

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

- a) $\mathbf{A} + \mathbf{A}^T$ je symetrická,
- b) $\mathbf{A} - \mathbf{A}^T$ je antisymetrická,
- c) existuje právě jedna symetrická \mathbf{B} a právě jedna antisymetrická \mathbf{C} tak, že $\mathbf{A} = \mathbf{B} + \mathbf{C}$,
- d) $\mathbf{A}^T \mathbf{A}$ 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 *involuce*.

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. Dokažte, že inverze regulární symetrické matice je symetrická matice.

2.18. (*) 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. Lze napsat jako $\mathbf{B} = \mathbf{XA}$, kde $\mathbf{a}_i \in \mathbb{R}^n$ jsou sloupce $\mathbf{A} \in \mathbb{R}^{n \times k}$ a $\mathbf{b}_i \in \mathbb{R}^m$ jsou sloupce $\mathbf{B} \in \mathbb{R}^{m \times k}$. 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.

2.7. Použijte cykličnost stopy.

2.17. Necht' \mathbf{A} je regulární a symetrická (tedy čtvercová). Necht' $\mathbf{BA} = \mathbf{I}$, tedy \mathbf{B} je levá inverze \mathbf{A} . Z toho $\mathbf{A}^T \mathbf{B}^T = \mathbf{AB}^T = \mathbf{I}$, tedy \mathbf{B}^T je pravá inverze \mathbf{A} . Ale pro regulární matici jsou si levá a pravá inverze rovny (viz §2.4), tedy $\mathbf{B}^T = \mathbf{B}$.

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

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

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**.

Věta 3.1. Jsou-li vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$ lineárně nezávislé, koeficienty $\alpha_1, \dots, \alpha_k$ jsou vektorem (3.1) určeny jednoznačně (tj. soustava (3.1) s neznámými $\alpha_1, \dots, \alpha_k$ má právě jedno řešení).

Důkaz. Necht' kromě rovnice (3.1) platí také $\mathbf{x} = \beta_1 \mathbf{x}_1 + \dots + \beta_k \mathbf{x}_k$. Odečtením obou rovnic máme $\mathbf{0} = (\alpha_1 - \beta_1) \mathbf{x}_1 + \dots + (\alpha_k - \beta_k) \mathbf{x}_k$. Ale z (3.2) plyne $\alpha_i - \beta_i = 0$, tedy $\alpha_i = \beta_i$. \square

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 , jestliže každá lineární kombinace každé (konečné) množiny vektorů z X leží v X (neboli že množina X je uzavřená vůči lineárním kombinacím):

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

Snadno se ukáže, že 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 . Platí následující netriviální tvrzení (důkazy najdete v učebnicích lineární algebry):

Věta 3.2.

- Z každé konečné množiny vektorů lze vybrat bázi jejich lineárního obalu.

- Každou lineárně nezávislou množinu vektorů z lineárního podprostoru lze doplnit na jeho bázi.
- Každý lineární podprostor má (alespoň jednu) bázi.
- Každá báze každého lineárního podprostoru má stejný počet vektorů.

Počet vektorů báze lineárního podprostoru X se nazývá jeho **dimenze**, značíme ji $\dim X$. Je-li $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ báze podprostoru X 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.

Věta 3.3. Pro každé podprostory $X, Y \subseteq \mathbb{R}^n$ platí:

- $X \subseteq Y$ implikuje $\dim X \leq \dim Y$.
- $X \subseteq Y$ a $\dim X = \dim Y$ implikuje $X = Y$.

Důkaz. Protože $X \subseteq Y$, každá báze podprostoru X patří do Y . Dle Věty 3.2 lze tuto bázi doplnit na bázi podprostoru Y , odtud první tvrzení. Jestliže navíc $\dim X = \dim Y$, každá báze X je už bází Y (tedy doplnění nepřidá žádný vektor), odtud druhé tvrzení. \square

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 3.4. 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

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

pro všechna $\mathbf{x} \in \mathbb{R}^n$. Matice \mathbf{A} je navíc zobrazením \mathbf{f} určena jednoznačně.

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

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

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

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 $\mathbf{x} = x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n$. Z (3.4) plyne

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

Nyní označíme $[\mathbf{f}(\mathbf{e}_1) \ \cdots \ \mathbf{f}(\mathbf{e}_n)] = \mathbf{A}$. 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 platí $\mathbf{A} = \mathbf{B}$, protože 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 $\mathbf{A} \mathbf{x}$ je maticový součin matice $m \times n$ maticí $n \times 1$ (viz §2.7). Označíme-li $\mathbf{y} = \mathbf{A} \mathbf{x}$, 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

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

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

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

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 (2.9) a (2.10).

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

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

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

3.2.1 Prostor obrazů

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.
- Dle (3.8) je to lineární obal sloupců matice \mathbf{A} . Tedy je to lineární podprostor \mathbb{R}^m .

Dimenze lineárního obalu sloupců se nazývá hodnota matice (už jsme ji potkali v §2.3),

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

Věta 3.5. *Následující tvrzení jsou ekvivalentní:*

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

Důkaz.

- $1 \Leftrightarrow 2 \Leftrightarrow 3 \Leftrightarrow 4$ plyne přímo z definic.
- $3 \Leftrightarrow 5$ plyne z definice hodnosti a z (2.2).
- $2 \Rightarrow 6$ platí, neboť soustava $\mathbf{Ab}_i = \mathbf{e}_i$ 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 $6 \Rightarrow 2$ položíme $\mathbf{x} = \mathbf{By}$.
- Tvrzení $1 \Leftrightarrow 7$ plyne z rovnosti (5.4a), uvedené později. □

Věta 3.6. 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 matice \mathbf{B} jsou lineárně nezávislé.

Důkaz. První tvrzení říká, že jestliže soustava $\mathbf{z} = \mathbf{ABx}$ má řešení, pak i soustava $\mathbf{z} = \mathbf{Ay}$ má řešení. To je ale jasné, protože můžeme položit $\mathbf{y} = \mathbf{Bx}$. Druhé tvrzení navíc říká, že jestliže \mathbf{B} má lineárně nezávislé řádky a $\mathbf{z} = \mathbf{Ay}$ má řešení, pak i $\mathbf{z} = \mathbf{ABx}$ má řešení. To platí, neboť dle implikace $5 \Rightarrow 2$ ve Větě 3.5 má soustava $\mathbf{Bx} = \mathbf{y}$ řešení pro každé \mathbf{y} . \square

3.2.2 Nulový prostor

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

$$\text{null } \mathbf{A} = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{0} \} \subseteq \mathbb{R}^n. \quad (3.12)$$

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 3.7. Následující tvrzení jsou ekvivalentní:

1. $\text{null } \mathbf{A} = \{\mathbf{0}\}$ (tj. nulový prostor je triviální).
2. Soustava $\mathbf{Ax} = \mathbf{0}$ má jediné řešení $\mathbf{x} = \mathbf{0}$.
3. $\text{rank } \mathbf{A} = n$.
4. Zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ je injektivní (viz §1.1.2).
5. Sloupce matice \mathbf{A} jsou lineárně nezávislé.
6. Matice \mathbf{A} má levou inverzi, tj. existuje \mathbf{B} tak, že $\mathbf{BA} = \mathbf{I}$.
7. Matice $\mathbf{A}^T \mathbf{A} \in \mathbb{R}^{n \times n}$ je regulární.

Důkaz.

- Ekvivalence $1 \Leftrightarrow 2$ plyne z definice nulového prostoru (3.12).
- Ekvivalence $2 \Leftrightarrow 5$ plyne z definice lineární nezávislosti (3.2), tj. $\mathbf{Ax} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0}$.
- Ekvivalence $3 \Leftrightarrow 5$ plyne z definice hodnosti (3.11).
- Tvrzení 4 říká, že pro každé \mathbf{x}, \mathbf{y} je $\mathbf{Ax} = \mathbf{Ay} \Rightarrow \mathbf{x} = \mathbf{y}$, tj. $\mathbf{A}(\mathbf{x} - \mathbf{y}) = \mathbf{0} \Rightarrow \mathbf{x} - \mathbf{y} = \mathbf{0}$. Ale to je definice (3.2) lineární nezávislosti sloupců \mathbf{A} . Tedy platí $2 \Leftrightarrow 4$.
- Tvrzení 6 je ekvivalentní tomu, že matice \mathbf{A}^T má pravou inverzi, tj. $\mathbf{A}^T \mathbf{B}^T = \mathbf{I}$. Tedy $3 \Leftrightarrow 6$ plyne z Věty 3.5.
- Tvrzení $1 \Leftrightarrow 7$ je plyne z rovnosti (5.4b), uvedené později. \square

Věta 3.8. Pro libovolné matice \mathbf{A}, \mathbf{B} platí:

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

Důkaz. První tvrzení říká, že $\mathbf{B}\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{A}\mathbf{B}\mathbf{x} = \mathbf{0}$, což platí vynásobením maticí \mathbf{A} zleva. Druhé tvrzení říká, že když \mathbf{A} má lineárně nezávislé sloupce, pak $\mathbf{A}\mathbf{B}\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{B}\mathbf{x} = \mathbf{0}$. To platí, neboť dle implikace $5 \Rightarrow 2$ ve Větě 3.7 máme $\mathbf{A}\mathbf{y} = \mathbf{0} \Rightarrow \mathbf{y} = \mathbf{0}$. \square

Dimenze prostoru obrazů a nulového prostoru jsou svázány následujícím vztahem:

Věta 3.9. Pro každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ platí

$$\dim \text{rng } \mathbf{A} + \dim \text{null } \mathbf{A} = n. \quad (3.13)$$

Věta 3.9 se říká *rank-plus-nullity theorem*, kde *nullity* označuje dimenzi nulového prostoru matice. Její důkaz je netriviální, najdete ho v každé učebnici lineární algebry. Interpretace věty:

- Každá dimenze na vstupu zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ se buď ‘smáčkne’ do nulového vektoru nebo se objeví na výstupu.
- Počet lineárně nezávislých řešení lineární homogenní soustavy $\mathbf{A}\mathbf{x} = \mathbf{0}$ je $n - \text{rank } \mathbf{A}$.
- Protože $\text{rank } \mathbf{A} = \text{rank}(\mathbf{A}^T)$, rovnost (3.13) lze psát také jako

$$\dim \text{rng}(\mathbf{A}^T) + \dim \text{null } \mathbf{A} = n. \quad (3.14)$$

Protože $\text{rng}(\mathbf{A}^T)$ je lineární obal řádků matice \mathbf{A} a $\text{null } \mathbf{A}$ je množina všech vektorů kolmých na řádky \mathbf{A} , rovnost (3.14) říká, že součet dimenzí těchto dvou podprostorů prostoru \mathbb{R}^n je n . O tom si více řekneme v části o ortogonálním doplňku (§4.2).

3.3 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. Množina $A \subseteq \mathbb{R}^n$ je **afinní podprostor**¹ lineárního prostoru \mathbb{R}^n , jestliže každá afinní kombinace každé (konečné) množiny vektorů z A leží v A (neboli množina A je uzavřená vůči afinním kombinacím):

$$\mathbf{x}_1, \dots, \mathbf{x}_k \in A, \quad \alpha_1 + \dots + \alpha_k = 1 \quad \implies \quad \alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k \in A. \quad (3.15)$$

Afinní kombinace *nezávisí na počátku*. To znamená, že afinní kombinace vektorů posunutých o vektor \mathbf{x}_0 je rovna afinní kombinaci neposunutých vektorů posunutých o \mathbf{x}_0 . To snadno dokážeme:

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

Na rozdíl od toho, součet vektorů (a tedy i obecná lineární kombinace) na počátku závisí.

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 $\mathbf{x}_1, \mathbf{x}_2$. Tato přímka je afinní podprostor \mathbb{R}^n . Na obrázku dole vlevo je několik bodů na přímce a příslušné koeficienty (α_1, α_2) .

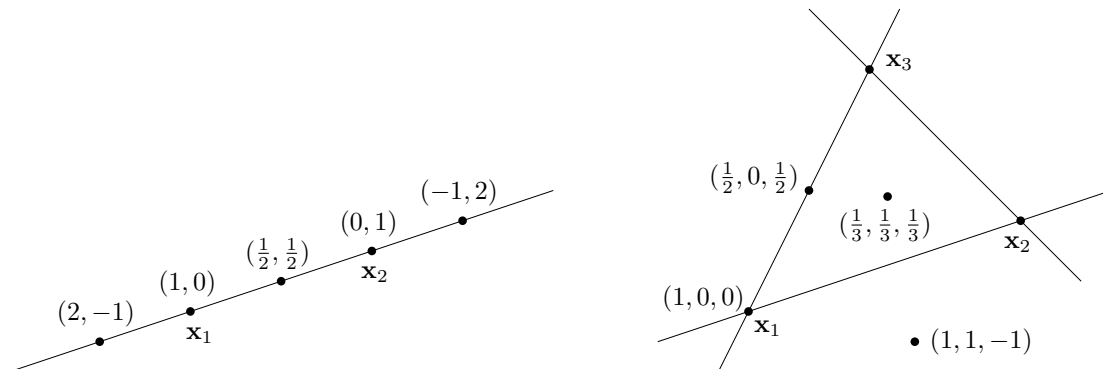
¹Všimněte si, že definujeme afinní *podprostor* lineárního prostoru, ale už ne afinní *prostor* sám o sobě. Definice afinního prostoru bez odkazu k nějakému lineárnímu prostoru (tj. pomocí axiomů) existuje, ale neuvádíme ji.

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. Tato rovina je afinní podprostor \mathbb{R}^n . Na obrázku vpravo je několik bodů v této rovině a příslušné koeficienty $(\alpha_1, \alpha_2, \alpha_3)$.



Protože afinní kombinace nezávisí na poloze počátku, do obrázků jsme počátek ani nekreslili a prvky \mathbb{R}^n jsme kreslili jako body a ne jako vektory (tj. šipky spojující počátek s koncovým bodem), což by bylo třeba u lineárních kombinací². \square

Body $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ jsou **afinně nezávislé**, jestliže žádný není afinní kombinací ostatních.

Příklad 3.7. Dva body $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ jsou afinně nezávislé, právě když $\mathbf{x}_1 \neq \mathbf{x}_2$ (neboli nejsou identické). 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 *kolineární*). 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*). \square

Věta 3.10. Pro množinu $A \subseteq \mathbb{R}^n$ jsou následující tvrzení ekvivalentní:

1. A je afinní podprostor \mathbb{R}^n .
2. Existuje lineární podprostor $X \subseteq \mathbb{R}^n$ a vektor $\mathbf{x}_0 \in \mathbb{R}^n$ tak, že $A = \{ \mathbf{x} + \mathbf{x}_0 \mid \mathbf{x} \in X \}$.
3. Existuje matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ a vektor $\mathbf{b} \in \mathbb{R}^m$ tak, že $A = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b} \}$.

Důkaz. Důkaz $1 \Rightarrow 2$: Necht' $\mathbf{x}_0 \in A$. Ukážeme, že množina $X = \{ \mathbf{x} - \mathbf{x}_0 \mid \mathbf{x} \in A \}$ je lineární podprostor, tedy že lineární kombinace vektorů z X leží v X . Tzn. že pro každé $\mathbf{x}_1, \dots, \mathbf{x}_k \in A$ a $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ musí existovat $\mathbf{x} \in A$ tak, že $\alpha_1(\mathbf{x}_1 - \mathbf{x}_0) + \dots + \alpha_k(\mathbf{x}_k - \mathbf{x}_0) = \mathbf{x} - \mathbf{x}_0$. To platí pro $\mathbf{x} = \alpha_0 \mathbf{x}_0 + \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$ kde $\alpha_0 = 1 - \alpha_1 - \dots - \alpha_k$ (ověřte dosazením!).

Důkaz $2 \Rightarrow 3$: Necht' \mathbf{A} je matice taková, že $X = \text{null } \mathbf{A}$. Necht' \mathbf{x}_0 je libovolné takové, že $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$ (tzv. *partikulární řešení soustavy*). Necht' $\mathbf{b} = \mathbf{A}\mathbf{x}_0$. Pak

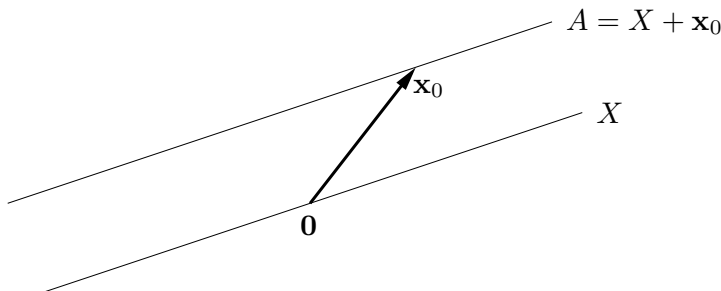
$$\{ \mathbf{x} + \mathbf{x}_0 \mid \mathbf{x} \in X \} = \{ \mathbf{x} + \mathbf{x}_0 \mid \mathbf{A}\mathbf{x} = \mathbf{0} \} = \{ \mathbf{x} \mid \mathbf{A}(\mathbf{x} - \mathbf{x}_0) = \mathbf{0} \} = \{ \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b} \}.$$

Důkaz $3 \Rightarrow 1$: Máme dokázat, že množina $A = \{ \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b} \}$ je afinní podprostor. Necht' $\mathbf{x}_1, \dots, \mathbf{x}_k \in A$ a $\alpha_1 + \dots + \alpha_k = 1$. Pak

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

²To je jeden z důvodů, proč se prvkům afinního (pod)prostoru říká *body* zatímco prvkům vektorového (= lineárního) (pod)prostoru *vektory*.

Ekvivalence $1 \Leftrightarrow 2$ ve větě říká, že afinní podprostor není nic jiného než ‘posunutý’ lineární podprostor (tedy nemusí procházet počátkem, na rozdíl od lineárního podprostoru). Viz obrázek (ve kterém jsme použili zkrácené označení $\{\mathbf{x} + \mathbf{x}_0 \mid \mathbf{x} \in X\} = X + \mathbf{x}_0$):



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**.

Ekvivalence $1 \Leftrightarrow 3$ ve větě říká, že afinní podprostor není nic jiného než množina řešení nějaké (ne nutně homogenní) lineární soustavy³.

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 dokázat (proved'te!), že zobrazení \mathbf{f} je afinní, právě když existuje matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ a vektor $\mathbf{b} \in \mathbb{R}^m$ tak, že

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}. \quad (3.16)$$

Všimněte si, že dle Věty 3.10 je množina $\mathbf{f}(\mathbb{R}^n) \subseteq \mathbb{R}^m$ afinní podprostor.

Příklad 3.8. Zobrazení $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definované jako $\mathbf{f}(x_1, x_2) = (x_1 + x_2 + 1, x_1 - x_2 - 2, 2x_1)$ 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 (3.16):

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

Pro $m = 1$ se zobrazení (3.16) nazývá také **afinní funkce**⁴ $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a má tvar

$$f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b = a_1 x_1 + \dots + a_n x_n + b, \quad (3.17)$$

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

3.4 Cvičení

3.1. Rozhodněte, zda následující množiny jsou lineární nebo afinní podprostory \mathbb{R}^n a když ano, určete jejich dimenze:

³Proto se afinnímu podprostoru také říká *lineární varieta*. Je to speciální případ *algebraické variety*, což je množina všech řešení soustavy polynomiálních rovnic.

⁴V 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.

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

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 (3.16). 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 (napište jeho bázi) a $\mathbf{x}_0 \in \mathbb{R}^3$.

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

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

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.

- a) Pokud \mathbf{AB} má plnou hodnotu, pak \mathbf{A} a \mathbf{B} mají plnou hodnotu.
- b) Pokud \mathbf{A} a \mathbf{B} mají plnou hodnotu, pak \mathbf{AB} má plnou hodnotu.
- c) Pokud \mathbf{A} a \mathbf{B} mají triviální nulový prostor, pak \mathbf{AB} má triviální nulový prostor.
- d) (\star) 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.

- e) (\star) 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.

Obecněji, máme k dvojic bodů $(\mathbf{p}_1, \mathbf{q}_1), \dots, (\mathbf{p}_k, \mathbf{q}_k)$, 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:

- a) $f(1, 2) = 2, f(3, 4) = 3$.

- b) $f(1, 2) = 2$, $f(3, 4) = 3$, $f(5, 6) = 4$.
 c) $f(1, 0, 1) = -1$, $f(0, 1, 2) = 1$, $f(1, 1, 3) = 2$.
- 3.11. Máme podprostor $X \subseteq \mathbb{R}^3$ s bází tvořenou vektory $(1, 2, 3)$ a $(-1, 0, 1)$. Necht' $\mathbf{x} = (2, 2, 2)$. Platí $\mathbf{x} \in X$? Pokud ano, najděte souřadnice vektoru \mathbf{x} v této bázi.
- 3.12. Necht' \mathbf{A} má lineárně nezávislé sloupce. Ukažte, že každá báze podprostoru $\text{rng } \mathbf{A}$ je tvořena sloupci matice \mathbf{AB} pro nějakou regulární matici \mathbf{B} .
- 3.13. Dokažte, že pro libovolné dvě matice platí $\text{rank}(\mathbf{AB}) \leq \min\{\text{rank } \mathbf{A}, \text{rank } \mathbf{B}\}$.
- 3.14. Dokažte, že pro každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ hodnosti r existují matice $\mathbf{B} \in \mathbb{R}^{m \times r}$ a $\mathbf{C} \in \mathbb{R}^{r \times n}$ tak, že $\mathbf{A} = \mathbf{BC}$. Poznamenejme, že tomuto rozkladu se říká **rozklad matice podle hodnosti** (*rank factorization*). Jestliže $r \ll \min\{m, n\}$, můžeme tím v některých situacích ušetřit paměť nebo čas. Porovnejte velikost paměti potřebné pro uložení matice \mathbf{A} a matic \mathbf{B}, \mathbf{C} . Jak byste efektivně spočítali součin \mathbf{Ax} ?
- 3.15. (★) 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 $\text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$. Všimněte si, že vlastně dokazujete třetí tvrzení Věty 3.2.
- 3.16. (★) Máme lineárně nezávislé vektory $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ a hledáme vektory $\mathbf{a}_{m+1}, \dots, \mathbf{a}_n \in \mathbb{R}^n$ tak, aby vektory $\mathbf{a}_1, \dots, \mathbf{a}_n$ byly lineárně nezávislé. Dokažte, že to jde udělat následovně. Necht' $\mathbf{A} \in \mathbb{R}^{m \times n}$ je matice s řádky $\mathbf{a}_1^T, \dots, \mathbf{a}_m^T$. Vybereme množinu $J \subseteq \{1, \dots, n\}$ lineárně nezávislých sloupců matice \mathbf{A} (viz předchozí cvičení). Pak zvolíme vektory $\mathbf{a}_{m+1}, \dots, \mathbf{a}_n$ jako $n - m$ vektorů standardní báze $\mathbf{e}_j \in \mathbb{R}^n$ kde $j \in \{1, \dots, n\} \setminus J$. Všimněte si, že vlastně dokazujete čtvrté tvrzení Věty 3.2.

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 je prázdná (tedy není afinní podprostor), protože soustava $\mathbf{ax}^T = \mathbf{I}$ nemá řešení pro žádné \mathbf{a}, \mathbf{x} (možný důkaz: je $\text{rank } \mathbf{I} = n$, ale $\text{rank}(\mathbf{ax}^T) \leq 1$).
- 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 $\mathbf{f}(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$, 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 $\mathbf{q}_i = \mathbf{Ap}_i + \mathbf{b}$, $i = 1, \dots, k$, pro \mathbf{A}, \mathbf{b} .

- 3.10.a) Je $f(x_1, x_2) = a_1x_1 + a_2x_2$. 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.
- 3.13. Z Věty 3.6 je $\text{rank}(\mathbf{AB}) \leq \text{rank} \mathbf{A}$. Dle (2.2) je $\text{rank}(\mathbf{AB}) = \text{rank}((\mathbf{AB})^T) = \text{rank}(\mathbf{B}^T \mathbf{A}^T) \leq \text{rank}(\mathbf{B}^T) = \text{rank} \mathbf{B}$.
- 3.14. 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} .

Kapitola 4

Ortogonalita

4.1 Standardní skalární součin

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

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

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**

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

Norma splňuje **trojúhelníkovou nerovnost** $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$. Ta 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 bázevský vektor podprostoru X , neboť (dokažte!)

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

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

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

nebot' jediný vektor kolmý sám na sebe je $\mathbf{0}$ (rozmyslete!).

Ortogonální 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!).

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. Pozor, stěna místnosti není kolmá na podlahu (proč?). \square

S ortogonálním doplňkem jsme se vlastně už setkali v §3.2.2: nulový prostor je tvořen všemi vektory kolmými na řádky matice. Tedy je-li $X = \text{rng}(\mathbf{A}^T)$, pak $X^\perp = \text{null } \mathbf{A}$. Přesněji by se toto tvrzení dokázalo z (4.4).

Příklad 4.3. Ortogonální doplněk podprostoru $X = \text{span}\{(1, 2, 3), (0, 1, -1)\}$ je množina všech vektorů (x_1, x_2, x_3) splňujících $x_1 + 2x_2 + 3x_3 = x_2 - x_3 = 0$, neboli $X^\perp = \text{null} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{bmatrix}$. \square

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

1. $\dim X + \dim X^\perp = n$,
2. $X \perp Y$ a $\dim X + \dim Y = n$ implikuje $Y = X^\perp$,
3. $(X^\perp)^\perp = X$.

Důkaz.

1. Tvrzení 1 je rovnost (3.14), kde $X = \text{rng}(\mathbf{A}^T)$ a $X^\perp = \text{null } \mathbf{A}$.
2. Platí $X \perp Y \Leftrightarrow Y \subseteq X^\perp$ (nebot' $Y \subseteq X^\perp$ znamená, že každý vektor z Y je kolmý na X). Dle tvrzení 1 tedy $\dim(X^\perp) = n - \dim X = \dim Y$. Dle Věty 3.3 proto $Y = X^\perp$.
3. Necht' $Y = X^\perp$. Z toho plyne $X \perp Y$ a $\dim X + \dim Y = n$. Z toho ale dle tvrzení 2 plyne $X = Y^\perp = (X^\perp)^\perp$. \square

Dle tvrzení 3 je $Y = X^\perp \Leftrightarrow Y = X^\perp$. V tom případě proto říkáme, že podprostory X a Y jsou *ortogonálním doplňkem jeden druhého*.

Zformulujme podrobněji souvislost ortogonálního doplňku a nulového prostoru:

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

$$(\text{rng } \mathbf{A})^\perp = \text{null}(\mathbf{A}^T). \quad (4.7a)$$

$$(\text{null } \mathbf{A})^\perp = \text{rng}(\mathbf{A}^T), \quad (4.7b)$$

Důkaz. Stačí si uvědomit, pro každou $\mathbf{A} \in \mathbb{R}^{m \times n}$, že:

- $\text{rng } \mathbf{A} = \{ \mathbf{Ax} \mid \mathbf{x} \in \mathbb{R}^n \} \subseteq \mathbb{R}^m$ 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 prostor 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 prostor všech vektorů kolmých na sloupce \mathbf{A} .

Rovnost (4.7a) plyne přímo z toho (pro přesný důkaz bychom užili (4.4)). Rovnost (4.7b) se získá použitím rovnosti (4.7a) na matici \mathbf{A}^T a rovnosti $(X^\perp)^\perp = X$. \square

4.3 Ortonormální množina vektorů

Vektor \mathbf{x} nazveme **normalizovaný**, pokud má jednotkovou délku ($\|\mathbf{x}\| = 1 = \mathbf{x}^T \mathbf{x}$). Množinu vektorů $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ nazveme **ortonormální** (krátce říkáme, že *vektory* $\mathbf{x}_1, \dots, \mathbf{x}_k$ *jsou ortonormální*¹), jestliže každý vektor z této množiny je normalizovaný a každá dvojice vektorů z této množiny je ortogonální, tedy

$$\mathbf{x}_i^T \mathbf{x}_j = \begin{cases} 0 & \text{když } i \neq j, \\ 1 & \text{když } i = j. \end{cases} \quad (4.8)$$

Věta 4.3. Ortonormální množina vektorů je lineárně nezávislá.

Důkaz. Vynásobme levou stranu implikace (3.2) skalárně vektorem \mathbf{x}_i , což dá

$$0 = \mathbf{x}_i^T \mathbf{0} = \alpha_1 \mathbf{x}_i^T \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_i^T \mathbf{x}_k = \alpha_i \mathbf{x}_i^T \mathbf{x}_i = \alpha_i.$$

Když to uděláme pro každé i , máme $\alpha_1 = \dots = \alpha_k = 0$. \square

Věta 4.4. Necht' jsou vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$ ortonormální. Necht'

$$\mathbf{x} = \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k. \quad (4.9)$$

Pak $\alpha_i = \mathbf{x}_i^T \mathbf{x}$.

Důkaz. Vynásobením rovnice (4.9) zleva skalárně vektorem \mathbf{x}_i máme $\mathbf{x}_i^T \mathbf{x} = \mathbf{x}_i^T \mathbf{x}_i \alpha_i = \alpha_i$. \square

Připomeňme (§3.1), že skaláry $\alpha_1, \dots, \alpha_k$ jsou souřadnice vektoru \mathbf{x} v ortonormální bázi $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ podprostoru $\text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$. Ze střední školy víte (zopakujte o odvod'te!), že $\mathbf{x}_i^T \mathbf{x}$ je délka ortogonální projekce vektoru \mathbf{x} do (normalizovaného) vektoru \mathbf{x}_i . Uvědomte si rozdíl oproti situaci, kdy vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$ jsou lineárně nezávislé, ale ne ortonormální (Věta 3.1): v tom případě bychom koeficienty α_i museli počítat řešením lineární soustavy (4.9).

Dosazením do (4.9) dostaneme, že pro každé $\mathbf{x} \in \text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ platí

$$\mathbf{x} = (\mathbf{x}_1^T \mathbf{x}) \mathbf{x}_1 + \dots + (\mathbf{x}_k^T \mathbf{x}) \mathbf{x}_k. \quad (4.10)$$

Podotkneme, že ortonormální báze lineárního (pod)prostoru odpovídá tomu, co ze základní školy znáte pod pojmem 'kartézský souřadnicový systém'. Souřadnice bodu (pod)prostoru vůči jeho ortonormální bázi (viz §3.1) pak známe jako jeho 'kartézské souřadnice'.

¹To je přísně řečeno nesprávně, protože ortonormalita není vlastnost jednoho vektoru, ale množiny vektorů.

4.4 Matice s ortonormálními sloupci

Necht' sloupce matice $\mathbf{U} \in \mathbb{R}^{m \times n}$ tvoří ortonormální množinu vektorů². Dle Věty 4.3 jsou sloupce \mathbf{U} lineárně nezávislé, tedy nutně $m \geq n$ (tj. \mathbf{U} je čtvercová nebo úzká). Podmínku ortonormality (4.8) sloupců matice \mathbf{U} lze psát stručně jako

$$\mathbf{U}^T \mathbf{U} = \mathbf{I}. \quad (4.11)$$

To také znamená, že \mathbf{U}^T je levá inverze matice \mathbf{U} a \mathbf{U} je pravá inverze matice \mathbf{U}^T (viz §2.4). Dle Věty 4.4 pro každé $\mathbf{x} \in \text{rng } \mathbf{U}$ platí $\mathbf{U}\mathbf{U}^T \mathbf{x} = \mathbf{x}$, což je rovnost (4.10) v maticovém tvaru.

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

$$\mathbf{f}(\mathbf{x})^T \mathbf{f}(\mathbf{y}) = (\mathbf{U}\mathbf{x})^T (\mathbf{U}\mathbf{y}) = \mathbf{x}^T \mathbf{U}^T \mathbf{U} \mathbf{y} = \mathbf{x}^T \mathbf{y}. \quad (4.12)$$

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**.

Věta 4.5. Pro každou čtvercovou matici \mathbf{U} platí

$$\mathbf{U}^T \mathbf{U} = \mathbf{I} \iff \mathbf{U}^T = \mathbf{U}^{-1} \iff \mathbf{U}\mathbf{U}^T = \mathbf{I}. \quad (4.13)$$

Důkaz. Necht' platí levá rovnost. Pak jsou sloupce \mathbf{U} ortonormální a tedy lineárně nezávislé. To ale znamená, že \mathbf{U} je regulární, protože je čtvercová. Vynásobením levé rovnice maticí \mathbf{U}^{-1} zprava získáme prostřední rovnici. Vynásobením prostřední rovnice maticí \mathbf{U} zleva získáme pravou rovnici. Zbýlé implikace dokážeme podobně. \square

Věta říká, že má-li čtvercová matice ortonormální sloupce, má ortonormální i řádky, a inverze takové matice se spočítá jednoduše transpozicí. Čtvercové matice splňující podmínky (4.13) se říká **ortogonální matice**.

Zdůrazněme, že pokud \mathbf{U} je obdélníková s ortonormálními sloupci, neplatí $\mathbf{U}\mathbf{U}^T = \mathbf{I}$ (platí pouze $\mathbf{U}\mathbf{U}^T \mathbf{x} = \mathbf{x}$ pro každé $\mathbf{x} \in \text{rng } \mathbf{U}$). Dále, pokud má \mathbf{U} ortogonální (ne však ortonormální) sloupce, nemusí mít ortogonální řádky³.

Necht' \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 $\det \mathbf{U} \in \{-1, 1\}$.

- Pokud $\det \mathbf{U} = 1$, matici se říká **speciální ortogonální** nebo také **rotační**, protože transformace $\mathbf{f}(\mathbf{x}) = \mathbf{U}\mathbf{x}$ (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 $\det \mathbf{U} = -1$, transformace \mathbf{f} je složením otočení a *zrcadlení* (reflexe) kolem nadroviny procházející počátkem.

Příklad 4.4. 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

²Takové matice se častěji značí písmeny $\mathbf{U}, \mathbf{V}, \mathbf{Q}$ a ne \mathbf{A}, \mathbf{B} jak jsme zvyklí.

³To je možná důvod, proč se čtvercové matice 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 4.5. Zrcadlení (neboli reflexe) v \mathbb{R}^2 kolem přímky procházející počátkem a směrnici $\tan(\varphi/2)$ je reprezentováno ortogonální maticí

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

□

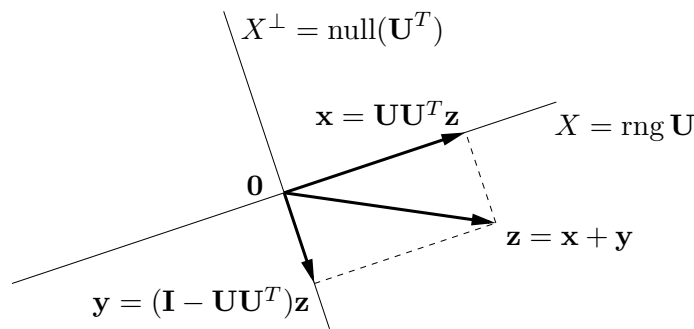
Příklad 4.6. Permutační matice je čtvercová matice, jejíž sloupce jsou permutované vektory standardní báze. Např.

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

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

4.5 Ortogonální projekce

Ortogonální projekce vektoru $\mathbf{z} \in \mathbb{R}^m$ na podprostor $X \subseteq \mathbb{R}^m$ je vektor $\mathbf{x} \in X$ takový, že $(\mathbf{z} - \mathbf{x}) \perp X$. Viz obrázek:



Ukážeme, jak ortogonální projekci spočítat, je-li dána ortonormální báze podprostoru X , tedy je-li dána matice $\mathbf{U} \in \mathbb{R}^{m \times n}$ taková, že $X = \text{rng } \mathbf{U}$ a $\mathbf{U}^T \mathbf{U} = \mathbf{I}$. Protože (jak ukážeme v §4.6) každý podprostor má ortonormální bázi, dokážeme tím zároveň i existenci a jednoznačnost ortogonální projekce. Případem, kdy podprostor X je zadán libovolnou (ne nutně ortonormální) bází, se budeme zabývat později v §5.1.1.

Věta 4.6. Necht' $\mathbf{U}^T \mathbf{U} = \mathbf{I}$. Ortogonální projekce vektoru \mathbf{z} na podprostor $\text{rng } \mathbf{U}$ je vektor

$$\mathbf{x} = \mathbf{U} \mathbf{U}^T \mathbf{z}. \quad (4.14)$$

Důkaz. Hledáme vektor $\mathbf{x} \in \text{rng } \mathbf{U}$ splňující $(\mathbf{z} - \mathbf{x}) \perp \text{rng } \mathbf{U}$. Dle (3.10) je $\mathbf{x} \in \text{rng } \mathbf{U}$, právě když $\mathbf{x} = \mathbf{U} \boldsymbol{\alpha}$ pro nějaké $\boldsymbol{\alpha}$. Podmínka $(\mathbf{z} - \mathbf{x}) \perp \text{rng } \mathbf{U}$ říká, že vektor $\mathbf{z} - \mathbf{x} = \mathbf{z} - \mathbf{U} \boldsymbol{\alpha}$ je kolmý na sloupce matice \mathbf{U} . To lze psát jako $\mathbf{U}^T (\mathbf{z} - \mathbf{U} \boldsymbol{\alpha}) = \mathbf{0}$ neboli $\mathbf{U}^T \mathbf{z} = \mathbf{U}^T \mathbf{U} \boldsymbol{\alpha}$. Protože $\mathbf{U}^T \mathbf{U} = \mathbf{I}$, tato rovnice má právě jedno řešení $\boldsymbol{\alpha} = \mathbf{U}^T \mathbf{z}$. Po dosazení $\mathbf{x} = \mathbf{U} \boldsymbol{\alpha} = \mathbf{U} \mathbf{U}^T \mathbf{z}$. □

Označíme-li $\mathbf{u}_1, \dots, \mathbf{u}_n$ sloupce matice \mathbf{U} , dle (2.10) lze vzorec (4.14) psát jako⁴

$$\mathbf{x} = \mathbf{U} \mathbf{U}^T \mathbf{z} = \mathbf{u}_1 \mathbf{u}_1^T \mathbf{z} + \dots + \mathbf{u}_n \mathbf{u}_n^T \mathbf{z} = (\mathbf{u}_1^T \mathbf{z}) \mathbf{u}_1 + \dots + (\mathbf{u}_n^T \mathbf{z}) \mathbf{u}_n, \quad (4.15)$$

⁴Protože $\mathbf{u} \mathbf{u}^T \mathbf{z} = (\mathbf{u}^T \mathbf{z}) \mathbf{u}$. Pozor ale: závorka 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).

kde skaláry $\alpha_i = \mathbf{u}_i^T \mathbf{z}$ jsou souřadnice vektoru \mathbf{x} v ortonormální bázi $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ podprostoru X (srov. Věta 4.4). Na vzorec (4.14) se lze tedy dívat tak, že $\boldsymbol{\alpha} = \mathbf{U}^T \mathbf{z} \in \mathbb{R}^n$ jsou souřadnice vektoru $\mathbf{x} = \mathbf{U}\boldsymbol{\alpha}$ v ortonormální bázi tvořené sloupci \mathbf{U} .

Má-li matice \mathbf{U} jediný sloupec (označme ho $\mathbf{U} = \mathbf{u} \in \mathbb{R}^m$), pak

$$\mathbf{x} = \mathbf{u}\mathbf{u}^T \mathbf{z} = (\mathbf{u}^T \mathbf{z})\mathbf{u}. \quad (4.16)$$

To je tedy vzoreček pro projekci vektoru \mathbf{z} na přímku $X = \text{rng } \mathbf{u} = \text{span}\{\mathbf{u}\}$ procházející počátkem (kde $\|\mathbf{u}\| = 1$), který máte znát ze střední školy. Středoškolské odvození: skalární součin $\mathbf{u}^T \mathbf{z}$ je délka (se znaménkem) průmětu a (4.16) je vektor této délky ve směru \mathbf{u} .

Zobrazení dané vzorcem (4.14) (za předpokladu $\mathbf{U}^T \mathbf{U} = \mathbf{I}$) je lineární zobrazení reprezentované maticí

$$\mathbf{P} = \mathbf{U}\mathbf{U}^T, \quad (4.17)$$

které se proto říká **ortogonální projektor** (na podprostor $X = \text{rng } \mathbf{U}$). Z (4.17) ihned plynou (ověřte!) dvě vlastnosti ortogonálního projektoru:

$$\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^T \quad (4.18)$$

(kde \mathbf{P}^2 je zkratka pro $\mathbf{P}\mathbf{P}$). Vlastnost $\mathbf{P}^2 = \mathbf{P}$ (říká se jí *idempotence* matice \mathbf{P}) říká očividnou věc: když vektor jednou promítneme na nějaký podprostor, tak opětovné promítnutí na tentýž podprostor ho už nezmění. Dá se ukázat (důkaz neuvádíme), že každá čtvercová matice s těmito vlastnostmi je ortogonální projektor (na nějaký podprostor).

4.5.1 Projekce na ortogonální doplněk

Pohled na ortogonální projekci se stane úplnější, zahrneme-li do něj kromě podprostoru X i jeho ortogonální doplněk X^\perp (viz obrázek výše). Vlastně jsme ukázali, že pro každý vektor $\mathbf{z} \in \mathbb{R}^m$ existuje právě jedna dvojice vektorů $\mathbf{x} \in X$ a $\mathbf{y} \in X^\perp$ tak, že $\mathbf{x} + \mathbf{y} = \mathbf{z}$. Opravdu, podmínku $\mathbf{y} \in X^\perp$ lze psát jako $\mathbf{y} \perp X$ neboli $(\mathbf{z} - \mathbf{x}) \perp X$ (domyslete do konce!).

Co je prostorem obrazů a nulovým prostorem ortogonálního projektoru? Úvahou (kam se promítne každý vektor? které vektory se promítnou do počátku?) snadno vidíme, že

$$\text{rng } \mathbf{P} = X, \quad (4.19a)$$

$$\text{null } \mathbf{P} = X^\perp. \quad (4.19b)$$

Algebraicky to také plyne z Vět 3.6 a 3.8.

Dále z obrázku vidíme, že je-li \mathbf{P} projektor na X , tak $\mathbf{I} - \mathbf{P}$ je projektor na X^\perp . Vskutku: $\mathbf{y} = \mathbf{z} - \mathbf{x} = \mathbf{z} - \mathbf{P}\mathbf{z} = (\mathbf{I} - \mathbf{P})\mathbf{z}$.

Rozvedme tuto myšlenku: zkoumejme (čtvercovou) ortogonální matici $[\mathbf{U} \ \mathbf{V}] \in \mathbb{R}^{m \times m}$, která je složená ze dvou bloků $\mathbf{U} \in \mathbb{R}^{m \times n}$ a $\mathbf{V} \in \mathbb{R}^{m \times (m-n)}$. Z ortonormality sloupců je jasné, že $\mathbf{U}^T \mathbf{U} = \mathbf{I}$ a $\mathbf{V}^T \mathbf{V} = \mathbf{I}$, a také $\mathbf{U}^T \mathbf{V} = \mathbf{0}$ neboli $\text{rng } \mathbf{U} \perp \text{rng } \mathbf{V}$. Platí ovšem dokonce

$$(\text{rng } \mathbf{U})^\perp = \text{rng } \mathbf{V}, \quad (4.20)$$

což plyne z Věty 4.1, neboť $\dim \text{rng } \mathbf{U} + \dim \text{rng } \mathbf{V} = n + (m - n) = m$. Vidíme, že rozdělení sloupců ortogonální matice do dvou bloků generuje rozklad prostoru na podprostor a jeho ortogonální doplněk. Navíc z Věty 4.5 máme

$$[\mathbf{U} \ \mathbf{V}] [\mathbf{U} \ \mathbf{V}]^T = \mathbf{U}\mathbf{U}^T + \mathbf{V}\mathbf{V}^T = \mathbf{I}, \quad (4.21)$$

jak ověříme roznásobením blokových matic. To souhlasí, protože projektor na podprostor X je $\mathbf{P} = \mathbf{U}\mathbf{U}^T$ a tedy projektor na podprostor X^\perp je $\mathbf{V}\mathbf{V}^T = \mathbf{I} - \mathbf{U}\mathbf{U}^T = \mathbf{I} - \mathbf{P}$.

4.5.2 Vzdálenost bodu od podprostoru

Z obrázku je intuitivně jasné, že projekce $\mathbf{x} = \mathbf{P}\mathbf{z}$ je bod podprostoru X nejbližší bodu \mathbf{z} . Důkaz tohoto tvrzení vynecháme (lze ho dokázat pomocí Pythagorovy věty). Tato vlastnost ortogonální je důležitá pro optimalizaci a využijeme ji např. v příští kapitole.

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 $(\text{rng } \mathbf{U})^\perp = \text{null}(\mathbf{U}^T)$.

4.6 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

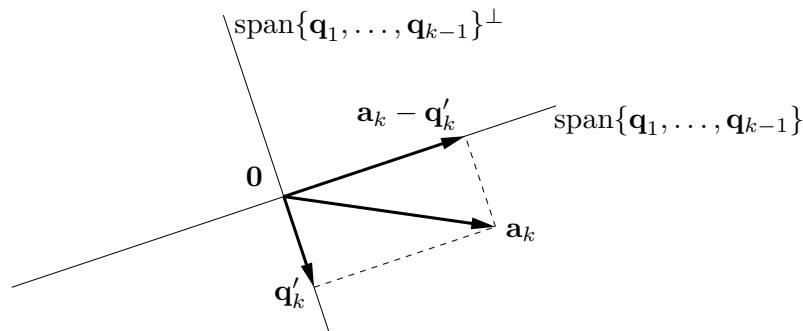
- $\mathbf{q}_1, \dots, \mathbf{q}_n$ jsou ortonormální,
- pro každé $k = 1, \dots, n$ platí $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_k\} = \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$.

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

Myšlenka algoritmu je jednoduchá. Předpokládejme, že již máme vektory $\mathbf{q}_1, \dots, \mathbf{q}_{k-1}$ jejichž lineární obal je $\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{k-1}\}$. 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 (s užitím vzorce (4.15))

$$\mathbf{q}'_k = \mathbf{a}_k - \sum_{i=1}^{k-1} (\mathbf{q}_i^T \mathbf{a}_k) \mathbf{q}_i, \quad \mathbf{q}_k = \frac{\mathbf{q}'_k}{\|\mathbf{q}'_k\|}. \quad (4.22)$$

Protože vektor \mathbf{q}'_k je projekce na $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}^\perp$, je kolmý na vektory $\mathbf{q}_1, \dots, \mathbf{q}_{k-1}$. Podle 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 $\mathbf{q}'_k \neq \mathbf{0}$. Viz obrázek:



Celý algoritmus provede iteraci (4.22) 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, to lze ale poměrně snadnou změnou (kterou neuvádíme) napravit.

4.6.1 QR rozklad

Věta 4.7. Každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ kde $m \geq n$ lze rozložit na součin

$$\mathbf{A} = \mathbf{Q}\mathbf{R}, \quad (4.23)$$

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.

Rozkladu matice (4.23) se říká **QR rozklad** matice. QR rozklad je užitečný v mnoha aplikacích, např. na řešení lineárních soustav. K tomu se dostaneme později.

Gram-Schmidtova ortogonalizace umožňuje dokázat Větu 4.7 pro případ, kdy matice \mathbf{A} má lineárně nezávislé sloupce:

Důkaz. Rovnosti (4.22) lze napsat jako

$$\mathbf{a}_k = \sum_{i=1}^{k-1} r_{ik} \mathbf{q}_i + r_{kk} \mathbf{q}_k = \sum_{i=1}^k r_{ik} \mathbf{q}_i, \quad (4.24)$$

kde $r_{ik} = \mathbf{q}_i^T \mathbf{a}_k$ pro $i < k$ a $r_{kk} = \|\mathbf{q}'_k\|$. Soustava rovnic (4.24) pro $k = 1, \dots, n$ se dá napsat v maticovém tvaru 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 matice $\mathbf{R} = [r_{ik}]$ je horní trojúhelníková (rozmyslete!). \square

Modifikací Gramm-Schmidtova algoritmu lze dosáhnout toho, že počítá QR rozklad i pro matici \mathbf{A} s lineárně závislými sloupci. Jiné způsoby počítání QR rozkladu jsou založené na *Householderových reflexích* nebo *Givensových rotacích* (neuvádíme). V Matlabu ho spočítáte příkazem `[Q,R]=qr(A,0)`. Podotkneme, že Věta 4.7 popisuje tzv. *redukovanou* verzi QR rozkladu. Příkaz `[Q,R]=qr(A)` počítá *plnou* verzi, která doplní matici \mathbf{Q} přidáním sloupců na ortogonální matici $b \times m$ a matici \mathbf{R} nulovými řádky na rozměr $m \times n$.

4.7 Skalární součin a norma matic

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

$$\mathbf{A} \cdot \mathbf{B} = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij}, \quad (4.25)$$

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!)

$$\mathbf{A} \cdot \mathbf{B} = \text{tr}(\mathbf{A}^T \mathbf{B}), \quad (4.26)$$

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 (4.25) indukuje **normu matice** $\mathbf{A} \in \mathbb{R}^{m \times n}$ jako

$$\|\mathbf{A}\| = \sqrt{\mathbf{A} \cdot \mathbf{A}} = \sqrt{\text{tr}(\mathbf{A}^T \mathbf{A})} = \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \right)^{1/2}. \quad (4.27)$$

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 jsme použili cykličnost stopy (viz §2.6). Tedy i norma (4.27) je invariantní vůči isometrii,

$$\|\mathbf{UAV}^T\| = \|\mathbf{A}\|. \quad (4.28)$$

Existují i jiné maticové normy. Maticová norma (4.27) se nazývá **Frobeniova norma** a pokud ji chceme odlišit od jiných norem, značíme ji $\|\cdot\|_F$.

4.8 Cvičení

- 4.1. Máme vektory $\mathbf{x} = (1, 2, 3)$ a $\mathbf{y} = (-1, 0, 1)$. Spočítejte (a) délku vektoru \mathbf{x} , (b) vzdálenost bodů \mathbf{x} a \mathbf{y} , (c) úhel mezi vektory \mathbf{x} a \mathbf{y} .
- 4.2. Najděte bázi ortogonálního doplnku prostoru $\text{span}\{(0, 1, 1), (1, 2, 3)\}$.
- 4.3. Najděte ortonormální bázi podprostoru $\text{span}\{(1, 1, 1, -1), (2, -1, -1, 1), (-1, 2, 2, 1)\}$ pomocí QR rozkladu (použijte Matlab).
- 4.4. Dokažte, že součin ortogonálních matic je ortogonální matice.
- 4.5. Pro jaké n je matice $\text{diag}(-\mathbf{1}_n)$ (tedy diagonální matice se samými mínus jedničkami na diagonále) rotační?
- 4.6. 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 ‘kroutit’ při rotaci v n -rozměrném prostoru. Jaké je toto číslo pro $n = 2, 3, 4$? Najděte vzorec pro obecné n .
- 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. Máme vektory $\mathbf{x}_1 = (1, -1, 0, 2)$, $\mathbf{x}_2 = (1, 1, 1, 0)$, $\mathbf{x}_3 = (-1, -1, 2, 0)$.
 - a) Ověřte, že vektory $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ jsou po dvojicích ortogonální.
 - b) Najděte libovolnou bázi podprostoru $\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}^\perp$.
 - c) Najděte ortogonální projektory na podprostory $\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ a $\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}^\perp$.
- 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. Spočtete co nejjednodušším způsobem inverzi matice $\frac{1}{3} \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$. Je tato matice ortogonální projektor?
- 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 $\mathbf{U}^T \mathbf{U} = \mathbf{I}$ reprezentuje matice $\mathbf{H} = \mathbf{I} - 2\mathbf{U}\mathbf{U}^T$ zrcadlení (reflexi) vzhledem k podprostoru $X = (\text{rng } \mathbf{U})^\perp$, proto této matici budeme říkat *reflektor*.

- a) Odvod'te 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{U} má jediný sloupec, tedy $\mathbf{H} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^T$ kde $\mathbf{u}^T \mathbf{u} = 1$. Tato matice reprezentuje zrcadlení kolem nadroviny s normálovým vektorem \mathbf{u} a je známa jako *elementární reflektor* nebo *Householderova matice*.

- 4.14. (\star) *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. Existuje isometrie $\mathbf{f}: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ taková, že $\mathbf{f}(1, -1, 2) = (1, 2, -1, 1)$ a $\mathbf{f}(1, 1, 0) = (0, 1, -1, 0)$?
- 4.16. Dokažte: Jestliže sloupce matice \mathbf{U} tvoří ortonormální bázi nějakého podprostoru, pak sloupce matice $\mathbf{V} = \mathbf{U}\mathbf{C}$ tvoří ortonormální bázi téhož podprostoru pro každou ortogonální matici \mathbf{C} .
- 4.17. 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 $\{\mathbf{x} \mid \mathbf{U}^T \mathbf{x} = \mathbf{b}\}$.

- 4.18. Necht' $X, Y \subseteq \mathbb{R}^n$ jsou podprostory takové, že $X \subseteq Y$. Tvrdíme, že když vektor ortogonálně promítneme nejdříve na Y a potom na X , dostaneme stejný výsledek, jako když ho ortogonálně promítneme rovnou na X . Dokažte toto tvrzení algebraicky nebo ho vyvrát'te.
- 4.19. 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 $\|\mathbf{u}\| \leq 1$.

4.20. Necht' \mathbf{P} je ortogonální projektor (na nějaký podprostor). Dokažte, že

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

4.21. Obecnou *projekci* se v lineární algebře rozumí každé lineární zobrazení $\mathbf{f}(\mathbf{z}) = \mathbf{P}\mathbf{z}$, které je idempotentní, tedy $\mathbf{f}(\mathbf{f}(\mathbf{z})) = \mathbf{f}(\mathbf{z})$ neboli $\mathbf{P}^2 = \mathbf{P}$. Projekce nemusí být ortogonální, může být šikmá – pak promítáme ve směru podprostoru $\text{null } \mathbf{P}$ na podprostor $\text{rng } \mathbf{P}$.

- a) Dokažte, že pro každý vektor \mathbf{z} platí $\mathbf{z} - \mathbf{P}\mathbf{z} \in \text{null } \mathbf{P}$.
- b) Projekce je ortogonální, právě když $\mathbf{P}^T = \mathbf{P}$. Dokažte, že potom $(\text{rng } \mathbf{P})^\perp = \text{null } \mathbf{P}$.

- 4.22. Necht' \mathbf{A}, \mathbf{B} jsou libovolné matice splňující $\mathbf{AB} = \mathbf{I}$ (tedy \mathbf{A} je levá inverze \mathbf{B} a \mathbf{B} je pravá inverze \mathbf{A} , viz §2.4). Dokažte, že:
- $\mathbf{P} = \mathbf{BA}$ je (obecný) projektor dle Cvičení 4.21, tedy platí $\mathbf{P}^2 = \mathbf{P}$.
 - $\text{null } \mathbf{P} = \text{null } \mathbf{A}$ a $\text{rng } \mathbf{P} = \text{rng } \mathbf{B}$.
 - Co jsou matice \mathbf{A}, \mathbf{B} v případě ortogonální projekce?
- 4.23. Pro jaké vektory \mathbf{x}, \mathbf{y} platí trojúhelníková nerovnost $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ s rovností?
- 4.24. Dokažte Pythagorovu větu: $\mathbf{x} \perp \mathbf{y} \Rightarrow \|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$. Dokažte její zobecnění: jsou-li vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$ po dvojicích ortogonální, pak $\|\mathbf{x}_1 + \dots + \mathbf{x}_k\|^2 = \|\mathbf{x}_1\|^2 + \dots + \|\mathbf{x}_k\|^2$.
- 4.25. 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:
- Pro každou \mathbf{A} je $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A}) = (\mathbf{I} + \mathbf{A})(\mathbf{I} - \mathbf{A})$.
 - Pro každou \mathbf{A} je $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A})^{-1} = (\mathbf{I} + \mathbf{A})^{-1}(\mathbf{I} - \mathbf{A})$.
 - Pro každou antisymetrickou \mathbf{A} je $\mathbf{F}(\mathbf{A})$ ortogonální.
 - Pro každou ortogonální \mathbf{A} je $\mathbf{F}(\mathbf{A})$ antisymetrická.
 - 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.

Nápověda a řešení

- 4.2. Např. $(1, 1, -1)$
- 4.3. Báze je $\{(1, 1, 1, -1)/2, (3, -1, -1, 1)/\sqrt{12}, (0, 1, 1, 2)/\sqrt{6}\}$
- 4.5. Musí být $\det \text{diag}(-\mathbf{1}_n) = (-1)^n > 0$, tedy pro sudá n .
- 4.6. $\binom{n}{2} = n(n-1)/2$
- 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? Ort. projektor to není, protože nesplňuje $\mathbf{P}^2 = \mathbf{P}$.
- 4.11.c) Plyne z (b).
- 4.11.d) Plyne z (c) s použitím $(X^\perp)^\perp = X$.
- 4.13.c) Je $\mathbf{x} \in X = (\text{rng } \mathbf{U})^\perp = \text{null}(\mathbf{U}^T)$, tedy $\mathbf{U}^T \mathbf{x} = \mathbf{0}$, tedy $\mathbf{H}\mathbf{x} = \mathbf{x}$. Geometricky to znamená, že když vektor leží v rovině zrcadlení, tak ho zrcadlení nezmění.
- 4.13.d) $\mathbf{x} \perp X$ je to samé jako $\mathbf{x} \in X^\perp = ((\text{rng } \mathbf{U})^\perp)^\perp = \text{rng } \mathbf{U}$, tedy $\mathbf{x} = \mathbf{U}\boldsymbol{\alpha}$ pro nějaké $\boldsymbol{\alpha}$. Tedy $\mathbf{U}\mathbf{U}^T \mathbf{x} = \mathbf{U}\mathbf{U}^T \mathbf{U}\boldsymbol{\alpha} = \mathbf{U}\boldsymbol{\alpha} = \mathbf{x}$ (taky to lze vidět z (4.19a), protože $\mathbf{U}\mathbf{U}^T$ je projektor na X). Tedy $\mathbf{H}\mathbf{x} = -\mathbf{x}$. Geometricky: zrcadlení vektoru kolmého k rovině zrcadlení vektoru obrátí orientaci.
- 4.15. Ne, protože isometrie zachovává eukleidovskou normu, ale $\|(1, -1, 2)\| \neq \|(1, 2, -1)\|$.
- 4.16. Víme, že $\mathbf{U} \in \mathbb{R}^{m \times n}$ a $\mathbf{C} \in \mathbb{R}^{n \times n}$, kde $m \geq n$, a $\mathbf{U}^T \mathbf{U} = \mathbf{I}$, $\mathbf{C}^T \mathbf{C} = \mathbf{I} = \mathbf{C}\mathbf{C}^T$. Sloupce matic \mathbf{U} a \mathbf{V} tvoří báze stejného podprostoru, neboť dle Věty 3.6 je $\text{rng } \mathbf{U} = \text{rng } \mathbf{V}$ (\mathbf{C} je matice přechodu k jiné bázi). Sloupce matice \mathbf{V} jsou ortonormální, neboť $\mathbf{V}^T \mathbf{V} = \mathbf{C}^T \mathbf{U}^T \mathbf{U} \mathbf{C} = \mathbf{C}^T \mathbf{C} = \mathbf{I}$.
- 4.17.a) $\|\mathbf{U}^T \mathbf{x}\|$
- 4.17.b) $\|\mathbf{b}\|$
- 4.17.c) $\|\mathbf{U}^T \mathbf{x} - \mathbf{b}\|$, což plyne buď z 4.17.a posunutím afinního podprostoru do počátku, nebo z 4.17.b posunutím bodu \mathbf{x} do počátku.

- 4.18. Je pravdivé. Důkaz: Necht' $\begin{bmatrix} \mathbf{U} & \mathbf{V} \end{bmatrix}$ je matice s ortonormálními sloupci taková, že $X = \text{rng } \mathbf{U}$ a $Y = \text{rng } \begin{bmatrix} \mathbf{U} & \mathbf{V} \end{bmatrix}$. Z toho plyne $X \subseteq Y$. Všimněte si (použijeme později), že $\mathbf{U}^T \mathbf{U} = \mathbf{I}$ a $\mathbf{U}^T \mathbf{V} = \mathbf{0}$. Označme $\mathbf{P} = \mathbf{U} \mathbf{U}^T$ projektor na X a $\mathbf{Q} = \begin{bmatrix} \mathbf{U} & \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{U} & \mathbf{V} \end{bmatrix}^T$ projektor na Y . Máme dokázat, že pro každé \mathbf{z} platí $\mathbf{P} \mathbf{Q} \mathbf{z} = \mathbf{P} \mathbf{z}$, neboli $\mathbf{P} \mathbf{Q} = \mathbf{P}$. Platí $\mathbf{P} \mathbf{Q} = \mathbf{U} \mathbf{U}^T \begin{bmatrix} \mathbf{U} & \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{U} & \mathbf{V} \end{bmatrix}^T = \mathbf{U} \begin{bmatrix} \mathbf{U}^T \mathbf{U} & \mathbf{U}^T \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{U} & \mathbf{V} \end{bmatrix}^T = \mathbf{U} \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U} & \mathbf{V} \end{bmatrix}^T = \mathbf{U} \mathbf{U}^T = \mathbf{P}$.
- 4.19.b) Doplníme matici \mathbf{U} na ortogonální matici $\mathbf{W} = \begin{bmatrix} \mathbf{U} & \mathbf{V} \end{bmatrix}$. Řádek této matice je $\mathbf{w}^T = \begin{bmatrix} \mathbf{u}^T & \mathbf{v}^T \end{bmatrix}$. Z ortogonality \mathbf{W} je však $\mathbf{w}^T \mathbf{w} = \mathbf{u}^T \mathbf{u} + \mathbf{v}^T \mathbf{v} = 1$. Z toho $\mathbf{u}^T \mathbf{u} \leq 1$.
- 4.20.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 $|\mathbf{u}_i^T \mathbf{u}_j| \leq 1$.
- 4.20.b) Máme $p_{ii} = \mathbf{u}_i^T \mathbf{u}_i \geq 0$
- 4.21.a) $\mathbf{z} - \mathbf{P} \mathbf{z} \in \text{null } \mathbf{P}$ znamená $\mathbf{P}(\mathbf{z} - \mathbf{P} \mathbf{z}) = \mathbf{P}(\mathbf{I} - \mathbf{P}) \mathbf{z} = \mathbf{0}$. Ale $\mathbf{P}(\mathbf{I} - \mathbf{P}) = \mathbf{P} - \mathbf{P}^2 = \mathbf{P} - \mathbf{P} = \mathbf{0}$.
- 4.21.b) Dle (4.7a) je $(\text{rng } \mathbf{P})^\perp = \text{null}(\mathbf{P}^T) = \text{null } \mathbf{P}$. Srov. s Cvičením 5.16.
- 4.22.a) $\mathbf{P}^2 = \mathbf{P} \mathbf{P} = \mathbf{B} \mathbf{A} \mathbf{B} \mathbf{A} = \mathbf{B} \mathbf{A} = \mathbf{P}$
- 4.22.b) \mathbf{B} má levou inverzi a tedy dle Věty 3.7 má l.n. sloupce. Dle Věty 3.8 tedy $\text{null } \mathbf{P} = \text{null}(\mathbf{B} \mathbf{A}) = \text{null } \mathbf{A}$. $\text{rng } \mathbf{P} = \text{rng } \mathbf{B}$ dokážeme podobně z Vět 3.5 a 3.6.
- 4.22.c) $\mathbf{A} = \mathbf{U}$ a $\mathbf{B} = \mathbf{U}^T$, kde \mathbf{U} má ortonormální sloupce.

Kapitola 5

Nehomogenní lineární soustavy

Mějme soustavu m lineárních rovnic o n neznámých

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

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 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 (dle Věty 3.10).

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. Rozlišme tři případy:

- Soustava nemá řešení. To nastane právě tehdy, když $\mathbf{b} \notin \text{rng } \mathbf{A}$. Taková soustava se nazývá **přeurčená**. 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, čímž se budeme zabývat v §5.2.

5.1 Přibližné řešení ve smyslu nejmenších čtverců

Pokud soustava (5.1) nemá řešení, řešme ji přibližně (což můžeme značit $\mathbf{Ax} \approx \mathbf{b}$). Hledejme takové \mathbf{x} , aby eukleidovská norma vektoru $\mathbf{r} = \mathbf{b} - \mathbf{Ax}$ 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

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

Protože minimalizujeme součet čtverců reziduí, mluvíme o přibližném řešení soustavy **ve smyslu nejmenších čtverců** (*least squares solution*)¹.

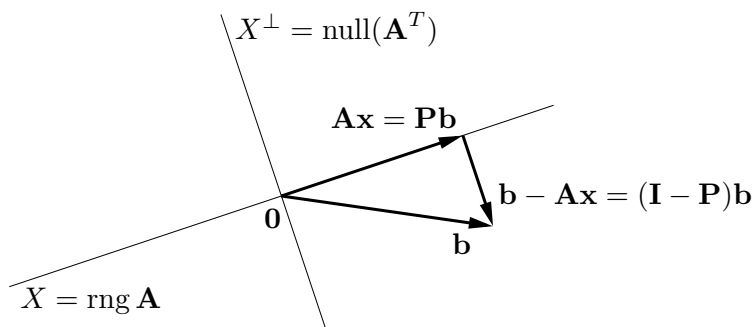
Příklad 5.1. Soustava třech rovnic o dvou neznámých

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

¹Přesně by se mělo říkat *least sum of squares*, protože je i metoda založená na *least median of squares*.

je přeurčená. 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 její řešitelnost, uvedeme následující větu.

Věta 5.1. Pro každou matici \mathbf{A} platí²

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

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

Důkaz. Rovnost (5.4b) 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:

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

nebot' pro libovolný vektor \mathbf{y} platí $\mathbf{y}^T \mathbf{y} = 0 \Rightarrow \mathbf{y} = \mathbf{0}$ (proč?).

Nyní použijeme Větu 3.9. Z (5.4b) a (3.13) je $\dim \text{rng}(\mathbf{A}^T \mathbf{A}) = \dim \text{rng}(\mathbf{A}^T)$. Z Věty 3.6 je $\text{rng}(\mathbf{A}^T \mathbf{A}) \subseteq \text{rng}(\mathbf{A}^T)$. Dle Věty 3.3 platí (5.4a). \square

Soustava (5.3) má řešení, právě když $\mathbf{A}^T \mathbf{b} \in \text{rng}(\mathbf{A}^T \mathbf{A})$. Ale z (5.4a) je $\text{rng}(\mathbf{A}^T) = \text{rng}(\mathbf{A}^T \mathbf{A})$. Protože $\mathbf{A}^T \mathbf{b} \in \text{rng}(\mathbf{A}^T)$ pro libovolné \mathbf{A}, \mathbf{b} , vidíme, že soustava má *vždy* řešení. Je zajímavé, že pro důkaz tohoto tvrzení jsme potřebovali rovnost (5.4a), jejíž důkaz není snadný (používá Větu 3.9). Zdá-li se vám, že musí existovat jednodušší důkaz, schválně ho zkuste najít!

Zkombinujeme-li (5.4a) (použité jednou na matici \mathbf{A} a jednou na matici \mathbf{A}^T) a (2.2), máme

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

²Matice tvaru $\mathbf{A}^T \mathbf{A}$ či \mathbf{AA}^T se objevují v různých situacích. Označme jako $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ sloupce matice \mathbf{A} . Matici $\mathbf{A}^T \mathbf{A} \in \mathbb{R}^{n \times n}$ se říká *Gramova matice* vektorů $\mathbf{a}_1, \dots, \mathbf{a}_n$ a její prvky jsou skalární součiny $\mathbf{a}_i^T \mathbf{a}_j$ (viz (2.9)). Matici $\mathbf{AA}^T \in \mathbb{R}^{m \times m}$ lze zase vidět (až na skalární násobek) jako empirickou kovarianční matici n pozorování $\mathbf{a}_1, \dots, \mathbf{a}_n$ m -tice náhodných proměnných.

Dle (5.5) je matice $\mathbf{A}^T \mathbf{A}$ regulární, právě 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!).

5.1.1 Ortogonální projekce na podprostor daný obecnou bází

Pokud \mathbf{x} je řešení normální rovnice, vektor $\mathbf{A}\mathbf{x}$ 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 $\mathbf{A}\mathbf{x} = \mathbf{P}\mathbf{b}$, 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 s bází (ne nutně ortonormální) tvořenou sloupci matice \mathbf{A} . Zdůrazněme, že projektor (5.7) vyjde stejný pro libovolnou bázi podprostoru X (viz Cvičení 5.11). Pokud je báze ortonormální, je $\mathbf{A}^T \mathbf{A} = \mathbf{I}$ a (5.7) se redukuje na (4.17).

Projekce na X^\perp má přirozenou úlohu v problému (5.2): hodnota jeho minima je $\|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 = \|\mathbf{b} - \mathbf{P}\mathbf{b}\|^2 = \|(\mathbf{I} - \mathbf{P})\mathbf{b}\|^2$.

5.1.2 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 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 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 $\mathbf{A}^T \mathbf{A}\mathbf{x} = \mathbf{A}^T \mathbf{b}$, 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 $\mathbf{A}^T \mathbf{A}\mathbf{x} = \mathbf{A}^T \mathbf{b}$ 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 $\mathbf{A} = \mathbf{Q}\mathbf{R}$. 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

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

Všimněte si, že soustava (5.8) není ekvivalentní původní soustavě $\mathbf{A}\mathbf{x} = \mathbf{b}$, 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 přeuročena, výsledkem je přibližné řešení ve smyslu nejmenších čtverců, přičemž použitý algoritmus používá QR rozklad. Pochopte všechny funkce operátorů *lomítko* a *zpětné lomítko* pomocí studia příkazů `help mrdivide` a `help mldivide`!

5.1.3 Lineární regrese

Regrese je modelování funkční závislosti nějaké proměnné na jiné proměnné. Modelujme závislost proměnné $y \in \mathbb{R}$ na proměnné $x \in X$ (kde X je nějaká množina, nejčastěji ale ne nutně $X = \mathbb{R}^d$) regresní funkcí

$$y = f(x, \boldsymbol{\theta}),$$

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

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

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

$$f(x, \boldsymbol{\theta}) = \theta_1 \varphi_1(x) + \dots + \theta_n \varphi_n(x) = \boldsymbol{\varphi}(x)^T \boldsymbol{\theta} \quad (5.10)$$

nějakých daných funkcí³ $\varphi_j: X \rightarrow \mathbb{R}$. Pak

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

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

Příklad 5.3. Nejjednodušší případ je pro $n = 1$ a konstantní funkci $\varphi_i(x) = 1$. Funkce (5.10) je tedy $f(x, \theta) = \theta$. Úloha (5.9) zní $\min_{\theta \in \mathbb{R}} \sum_i (y_i - \theta)^2$. Snadno spočítáme (udělejte!), že řešením je aritmetický průměr $\theta = \frac{1}{m} \sum_{i=1}^m y_i$ čísel y_1, \dots, y_m . \square

Příklad 5.4. *Polynomiální regrese*⁴. Necht' $X = \mathbb{R}$ a $\varphi_j(x) = x^{j-1}$. Pak regresní funkce

$$f(x, \boldsymbol{\theta}) = \theta_1 + \theta_2 x + \theta_3 x^2 + \dots + \theta_n x^{n-1}$$

³Funkce φ_j se často nazývají *bázové funkce* (pokud jsou ovšem lineárně nezávislé).

⁴Nedejte 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 $\boldsymbol{\theta}$.

je polynom stupně $n - 1$. Matice

$$\mathbf{A} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^{n-1} \end{bmatrix}$$

je známá jako *Vandermondova matice*. □

5.1.4 Statistické odůvodnění kritéria nejmenších čtverců

Možná se ptáte, proč se má nalezení přibližného řešení přeuročené soustavy formulovat zrovna jako (5.2). Uvedeme statistický důvod, odkud se kritérium nejmenšího součtu čtverců vzalo.

Odhadujme skryté parametry \mathbf{x} nějakého systému z měření \mathbf{b} na systému. Budiž vázány známou lineární závislostí $\mathbf{b} = \mathbf{A}\mathbf{x}$. Měření jsou zatížena chybami, které jsou způsobeny šumem senzorů, nepřesnostmi měření, nedokonalou znalostí modelu, apod. Tedy

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{r}, \quad (5.11)$$

kde $\mathbf{r} = (r_1, \dots, r_m)$ jsou náhodné proměnné modelující chyby měření $\mathbf{b} = (b_1, \dots, b_m)$. Metoda nejmenších čtverců říká, že máme minimalizovat $\|\mathbf{r}\|_2^2 = \sum_{i=1}^m r_i^2$, 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

$$p(\mathbf{r}) = p(r_1, \dots, r_m) = \prod_{i=1}^m p(r_i) = \prod_{i=1}^m c e^{-r_i^2/(2\sigma^2)}. \quad (5.12)$$

Dále použijeme *princip maxima věrohodnosti*. Ten říká, že parametry \mathbf{x} se mají najít tak, aby $p(\mathbf{r}) = p(\mathbf{b} - \mathbf{A}\mathbf{x})$ 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 $\sum_i r_i^2$.

5.1.5 Vícekriteriální nejmenší čtverce, regularizace

V některých úlohách se hodí minimalizovat více kritérií tvaru $\|\mathbf{A}_i\mathbf{x} - \mathbf{b}_i\|$ ‘současně’. K tomu se dá přistoupit tak, že minimalizujeme (nezáporně) vážený součet kritérií, tedy funkci⁵

$$\mu_1 \|\mathbf{A}_1\mathbf{x} - \mathbf{b}_1\|^2 + \cdots + \mu_k \|\mathbf{A}_k\mathbf{x} - \mathbf{b}_k\|^2 \quad (5.13)$$

⁵Matematicky elegantnější by samozřejmě bylo ‘schovat’ skaláry μ_i do matic \mathbf{A}_i a vektorů \mathbf{b}_i a tedy je tam nepsat. Odvození minima funkce (5.13) by pak bylo kratší.

kde $\mu_i \geq 0$, $\mathbf{A}_i \in \mathbb{R}^{m_i \times n}$ a $\mathbf{b}_i \in \mathbb{R}^{m_i}$. Minimalizace této funkce není nic nového pod sluncem, protože se dá převést na tvar (5.2). Opravdu, výraz (5.13) je roven (viz Cvičení 5.18)

$$\left\| \begin{bmatrix} \sqrt{\mu_1}(\mathbf{A}_1 \mathbf{x} - \mathbf{b}_1) \\ \vdots \\ \sqrt{\mu_k}(\mathbf{A}_k \mathbf{x} - \mathbf{b}_k) \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} \sqrt{\mu_1} \mathbf{A}_1 \\ \vdots \\ \sqrt{\mu_k} \mathbf{A}_k \end{bmatrix} \mathbf{x} - \begin{bmatrix} \sqrt{\mu_1} \mathbf{b}_1 \\ \vdots \\ \sqrt{\mu_k} \mathbf{b}_k \end{bmatrix} \right\|^2 = \|\mathbf{A}' \mathbf{x} - \mathbf{b}'\|^2,$$

kde $\mathbf{A}' \in \mathbb{R}^{m' \times n}$ a $\mathbf{b}' \in \mathbb{R}^{m'}$ kde $m' = m_1 + \dots + m_k$. Jestliže jsou sloupce matice \mathbf{A}' lineárně nezávislé, optimální \mathbf{x} je rovno (ověřte roznásobením blokových matic!)

$$\mathbf{x} = (\mathbf{A}'^T \mathbf{A}')^{-1} \mathbf{A}'^T \mathbf{b}' = (\mu_1 \mathbf{A}_1^T \mathbf{A}_1 + \dots + \mu_k \mathbf{A}_k^T \mathbf{A}_k)^{-1} (\mu_1 \mathbf{A}_1^T \mathbf{b}_1 + \dots + \mu_k \mathbf{A}_k^T \mathbf{b}_k). \quad (5.14)$$

Speciálně, někdy chceme přibližně řešit soustavu $\mathbf{A} \mathbf{x} = \mathbf{b}$ a zároveň chceme, aby norma řešení \mathbf{x} nebyla moc velká. To lze formulovat jako

$$\min_{\mathbf{x} \in \mathbb{R}^n} (\|\mathbf{A} \mathbf{x} - \mathbf{b}\|^2 + \mu \|\mathbf{x}\|^2). \quad (5.15)$$

pro zvolenou váhu $\mu > 0$. Přidání členu $\mu \|\mathbf{x}\|^2$ se říká (Tichonovova) **regularizace** úlohy (5.2). Dosazením do vzorečku (5.14) je optimální řešení rovno $\mathbf{x} = \mathbf{A}_\mu^+ \mathbf{b}$ kde

$$\mathbf{A}_\mu^+ = (\mathbf{A}^T \mathbf{A} + \mu \mathbf{I})^{-1} \mathbf{A}^T \quad (5.16)$$

je ‘regularizovaná pseudoinverze’ matice \mathbf{A} . Důležité je, že matice $\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}$ regulární pro každé $\mu > 0$ a $\mathbf{A} \in \mathbb{R}^{m \times n}$ (viz Cvičení 5.19), tedy \mathbf{A}_μ^+ je vždy definována. Viz také Cvičení 5.20.

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. Přírodním kritériem je minimalizovat euklidovskou normu (tedy vzdálenost od počátku) řešení, což vede na úlohu

$$\min \{ \|\mathbf{x}\|^2 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A} \mathbf{x} = \mathbf{b} \}. \quad (5.17)$$

Místo normy $\|\mathbf{x}\|$ opět minimalizujeme její čtverec. Tato úloha je známa jako řešení nehomogenní lineární soustavy **s nejmenší normou** (*least norm solution*). Podotkneme, že někdy je vhodné použít jiná kritéria než nejmenší eukleidovskou normu, viz např. Cvičení 10.18.

Příklad 5.5. Soustava dvou rovnic o třech neznámých

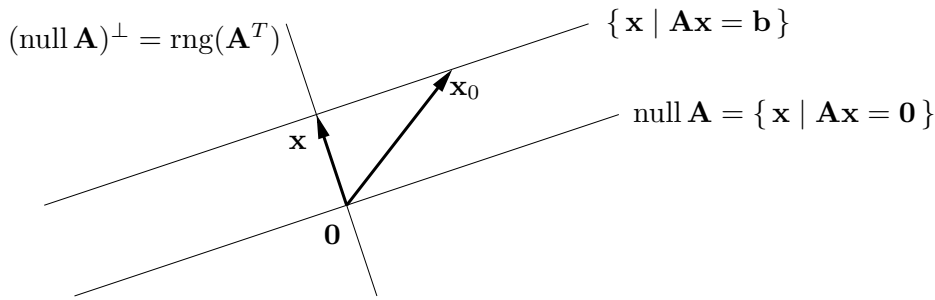
$$\begin{aligned} x + 2y + z &= 1 \\ -x + y + 2z &= 2 \end{aligned}$$

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 důkaz Věty 3.10) jako

$$\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} = \mathbf{b} \} = \text{null } \mathbf{A} + \mathbf{x}_0, \quad (5.18)$$

kde \mathbf{x}_0 je libovolné (partikulární) řešení soustavy, tedy $\mathbf{A} \mathbf{x}_0 = \mathbf{b}$. Množina (5.18) je afinní podprostor \mathbb{R}^n , je to lineární podprostor $\text{null } \mathbf{A}$ posunutý o \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 (4.7b). Neboli musí existovat vektor $\mathbf{y} \in \mathbb{R}^m$ tak, že $\mathbf{x} = \mathbf{A}^T \mathbf{y}$. Pro vyřešení úlohy (5.17) tedy musíme vyřešit soustavu rovnic

$$\mathbf{A}^T \mathbf{y} = \mathbf{x}, \quad (5.19a)$$

$$\mathbf{A} \mathbf{x} = \mathbf{b}. \quad (5.19b)$$

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

Vyřešme 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

$$\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} \quad (5.20)$$

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

$$\mathbf{x} = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{A} \mathbf{x}_0 = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{b} = \mathbf{A}^+ \mathbf{b}. \quad (5.21)$$

Všimněte si, že vzorce (5.6) a (5.20) dohromady definují pseudoinverzi libovolné matice (čtvercové, úzké nebo široké) s plnou hodnotou (tedy $\text{rank } \mathbf{A} = \min\{m, n\}$).

5.3 (★) Pseudoinverze obecné matice

Zatím jsme odděleně diskutovali případy, kdy soustava má žádné, jedno, nebo nekonečně mnoho řešení. Překvapivě, tyto tři případy lze spojit do jediné formulace. Zopakujme, že optimální řešení úlohy (5.2) jsou právě řešení soustavy normálních rovnic (5.3). Co když je ale sama soustava (5.3) nedourčená? Pak můžeme hledat její řešení s nejmenší normou, tj. řešit úlohu

$$\min\{\|\mathbf{x}\|^2 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}\}. \quad (5.22)$$

Protože úloha (5.17) má pro každou matici \mathbf{A} právě jedno optimální řešení, má i úloha (5.22) pro každou \mathbf{A} právě jedno optimální řešení. Toto řešení, \mathbf{x}^* , má tedy následující vlastnosti:

- Pokud soustava $\mathbf{A} \mathbf{x} = \mathbf{b}$ 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 (5.17).

Lze ukázat, že řešení úlohy (5.22) lze opět psát jako $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$ kde matici \mathbf{A}^+ nazýváme **pseudoinverze** (přesněji *Moore-Penroseova pseudoinverze*) matice \mathbf{A} . Když \mathbf{A} má lineárně nezávislé sloupce, pseudoinverze je rovna (5.6). Když \mathbf{A} má lineárně nezávislé řádky, pseudoinverze je rovna (5.20). Když ovšem \mathbf{A} nemá plnou hodnotu, pseudoinverzi je nutno počítat jinak. Elegantně se to udělá pomocí SVD, což neuvádíme.

Je několik jiných způsobů, jak definovat pseudoinverzi obecné matice. Zmíníme ještě jeden zajímavý. Všimněte si, že úloha (5.15) je v ‘něco mezi’ úlohami (5.2) a (5.17). Neformálně, (5.17) je minimalizace $\|\mathbf{x}\|^2 + \mu \|\mathbf{Ax} - \mathbf{b}\|^2$ pro velmi velké μ . To je ale totéž jako (5.15) pro velmi malé (kladné) μ . Lze ukázat, že pseudoinverze obecné matice je rovna

$$\mathbf{A}^+ = \lim_{\mu \rightarrow 0^+} \mathbf{A}_\mu^+. \quad (5.23)$$

5.4 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.

- e) Máme m nadrovin v prostoru \mathbb{R}^n , kde i -tá nadrovina 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.
- f) 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:
1. Kolik řešení má soustava $\mathbf{P}\mathbf{u} = \mathbf{q}$? Odpověď dokažte a interpretejte.
 2. Jsou sloupce \mathbf{P} lineárně nezávislé?

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

- g) Závislost výkonu P kotle na průtoku G plynu a průřezu S díry na přívod vzduchu je modelována funkcí $\hat{P}(G, S) = G(a_1 + a_2 S + a_3 10^{G+S} + a_4 10^{-S})$. Odhadujeme koeficienty a_1, \dots, a_4 z naměřených trojic $(G_1, S_1, P_1), \dots, (G_n, S_n, P_n)$.
- h) Známý průběh ceny akcie jisté firmy po dnech je daný posloupností p_1, \dots, p_k . Chceme předpovídat cenu akcie den dopředu. Tuto cenu modelujeme *autoregresní funkcí* $\hat{p}_{t+1} = \beta_1 + \beta_2 p_t + \beta_3 p_{t-1}$. Odhadněte koeficienty β_i tak, aby celková chyba predikce $\sum_{t=3}^k (p_t - \hat{p}_t)^2$ byla na onom známém průběhu ceny minimální.

- 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 zaveďte 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' $X = \text{span}\left\{ \left(-\frac{3}{5}, 0, \frac{4}{5}, 0\right), (0, 0, 0, 1), \left(\frac{4}{5}, 0, \frac{3}{5}, 0\right) \right\}$. 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 singulární 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 (5.20) pro libovolné (úzké, široké nebo čtvercové) matice plné hodnosti:

- a) $\mathbf{A}^+ = \mathbf{A}^{-1}$ když \mathbf{A} je čtvercová
- b) $(\mathbf{A}^+)^+ = \mathbf{A}$
- c) $(\mathbf{A}^T)^+ = (\mathbf{A}^+)^T$
- d) $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$, $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$
- e) $(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+$, $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}$

$$\text{f) } \mathbf{A}^T = \mathbf{A}^T \mathbf{A} \mathbf{A}^+ = \mathbf{A}^+ \mathbf{A} \mathbf{A}^T$$

$$\text{g) } (\mathbf{A}^T \mathbf{A})^+ = \mathbf{A}^+ (\mathbf{A}^T)^+, (\mathbf{A} \mathbf{A}^T)^+ = (\mathbf{A}^T)^+ \mathbf{A}^+$$

- 5.10. Spočítejte pseudoinverzi nenulového skaláru (tj. matice s jedním řádkem a jedním sloupcem), nenulového sloupcového vektoru (tj. matice s jedním sloupcem) a nenulového řádkového vektoru (tj. 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 $\mathbf{A}\mathbf{C}$ pro různé regulární matice $\mathbf{C} \in \mathbb{R}^{n \times n}$ (tedy \mathbf{C} je matice přechodu k jiné bázi), viz Cvičení 3.12.
- 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. Dokažte, že pokud $\mathbf{A}^T \mathbf{A} = \mathbf{B}^T \mathbf{B}$, pak existuje \mathbf{C} tak, že $\mathbf{A} = \mathbf{C}\mathbf{B}$.
- 5.14. Najděte co nejjednodušší vzorec pro
- vzdálenost počátku $\mathbf{0}$ od nadroviny $\{\mathbf{x} \mid \mathbf{a}^T \mathbf{x} = b\}$,
 - vzdálenost počátku $\mathbf{0}$ od afinního podprostoru $\{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}\}$ (kde \mathbf{A} má lineárně nezávislé řádky),
 - vzdálenost bodu \mathbf{x} od nadroviny $\{\mathbf{x} \mid \mathbf{a}^T \mathbf{x} = b\}$.
- 5.15. Neformálně je jasné, že nedourčená soustava bude mít vždy jen jedno řešení s nejmenší normou. Dokažte toto formálně, tj. dokažte, že jestliže soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ má alespoň jedno řešení pak soustava (5.19) má právě jedno řešení.
- 5.16. Čtvercová matice \mathbf{A} se nazývá *normální*, když $\mathbf{A}^T \mathbf{A} = \mathbf{A} \mathbf{A}^T$. Příkladem je symetrická nebo antisymetrická matice. Dokažte, že pro normální matice platí $(\text{rng } \mathbf{A})^\perp = \text{null } \mathbf{A}$.
- 5.17. Ukázali jsme (viz §4.5), že ortogonální projektor (4.17) na podprostor reprezentovaný ortonormální bází splňuje $\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^T$. Dokažte tyto rovnosti i pro ortogonální projektor (5.7) na podprostor reprezentovaný obecnou bází.
- 5.18. Pro vektory $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ ukažte, že $\left\| \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \right\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2$.
- 5.19. Dokažte, že matice $\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}$ je regulární pro každou matici \mathbf{A} a každé $\mu > 0$.
- 5.20. Uvažujme ‘regularizovanou pseudoinverzi’ (5.23). Dokažte, že pro každou matici \mathbf{A} a každé $\mu > 0$ platí $(\mathbf{A}^T \mathbf{A} + \mu \mathbf{I})^{-1} \mathbf{A}^T = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T + \mu \mathbf{I})^{-1}$. Dumejte nad významem této rovnosti.
- 5.21. (★) Dokažte, že pro každé matice \mathbf{A}, \mathbf{B} takové, že matice $\begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix}$ je čtvercová regulární a $\mathbf{A}^T \mathbf{B} = \mathbf{0}$, platí $\mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T + \mathbf{B}(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T = \mathbf{I}$. Vysvětlete rozdíl oproti (4.21).

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. Příklad: $m = 2, n = 1, \mathbf{A} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
- 5.1.c) Platí. Matice \mathbf{A} má hodnost m , tedy lineárně nezávislé řádky, tedy má netriviální nulový prostor.
- 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 přeúřčenou 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 $\mathbf{P} = [\mathbf{s}_1 \quad -\mathbf{s}_2]$, $\mathbf{u} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$, $\mathbf{q} = \mathbf{a}_2 - \mathbf{a}_1$.
- 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. $f(\mathbf{x}) = (\mathbf{A}\mathbf{x} - \mathbf{b})^T \mathbf{W}(\mathbf{A}\mathbf{x} - \mathbf{b}) = \|\mathbf{W}^{1/2} \mathbf{A}\mathbf{x} - \mathbf{W}^{1/2} \mathbf{b}\|^2$.
- 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 hodnotu.
- 5.8. \mathbf{A} je regulární, tedy $X = \mathbb{R}^m$. Projektor je identita.
- 5.9.b) Když \mathbf{A} má l.n. sloupce, dle (5.6) je $\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$. Protože $(\mathbf{A}^T \mathbf{A})^{-1}$ je regulární, \mathbf{A}^+ má l.n. řádky. Dle (5.20) tedy $\mathbf{A}^{++} = \mathbf{A}^{+T} (\mathbf{A}^+ \mathbf{A}^{+T})^{-1} = \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-T} [(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-T}]^{-1} = \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-T} [(\mathbf{A}^T \mathbf{A})^{-T}]^{-1} = \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-T} (\mathbf{A}^T \mathbf{A})^T = \mathbf{A}$. Když \mathbf{A} má l.n. řádky, udělá se to podobně.
- 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. $\mathbf{x} = \mathbf{0}$
- 5.13. Dle (5.4a) je $\text{rng}(\mathbf{A}^T) = \text{rng}(\mathbf{B}^T)$. Tedy každý řádek \mathbf{A} je lineární kombinací řádků \mathbf{B} . Dle Cvičení 3.14 to jde napsat jako $\mathbf{A} = \mathbf{C}\mathbf{B}$ pro nějaké \mathbf{C} .
- 5.14.a) $|\mathbf{b}|/\|\mathbf{a}\|$
- 5.14.b) Čtverec vzdálenosti je roven optimální hodnotě úlohy (5.17), 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) $|\mathbf{a}^T \mathbf{x} - b|/\|\mathbf{a}\|$
- 5.16. Dle (4.7a) a (5.4b) je $(\text{rng } \mathbf{A})^\perp = \text{null}(\mathbf{A}^T) = \text{null}(\mathbf{A} \mathbf{A}^T) = \text{null}(\mathbf{A}^T \mathbf{A}) = \text{null } \mathbf{A}$.
- 5.17. Ihned plyne z toho, že každý podprostor má ortonormální bázi (což jsme dokázali v §4.6), tedy stačí rovnosti dokázat pro (4.17). Ale můžeme také dokázat přímým dosazením:
 $\mathbf{P}^2 = \mathbf{P}\mathbf{P} = \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T = \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$,
 $\mathbf{P}^T = [\mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T]^T = \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-T} \mathbf{A}^T = \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$.
- 5.19. Je $\mathbf{A}^T \mathbf{A} + \mu \mathbf{I} = \mathbf{B}^T \mathbf{B}$, kde $\mathbf{B} = \begin{bmatrix} \mathbf{A} \\ \mu^{1/2} \mathbf{I} \end{bmatrix}$. Matice \mathbf{B} má l.n. sloupce, protože už matice $\mu^{1/2} \mathbf{I}$ je má. Tedy dle (5.5) má $\mathbf{B}^T \mathbf{B}$ plnou hodnotu, tedy je regulární.
- 5.20. Rovnici vynásobte zleva maticí $\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}$ a zprava maticí $\mathbf{A} \mathbf{A}^T + \mu \mathbf{I}$ a pak roznásobte závorky.
- 5.21. $\mathbf{I} = [\mathbf{A} \quad \mathbf{B}] [\mathbf{A} \quad \mathbf{B}]^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix} = [\mathbf{A} \quad \mathbf{B}] \left(\begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix} [\mathbf{A} \quad \mathbf{B}] \right)^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix}$
 $= [\mathbf{A} \quad \mathbf{B}] \begin{bmatrix} \mathbf{A}^T \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^T \mathbf{B} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix} = [\mathbf{A} \quad \mathbf{B}] \begin{bmatrix} (\mathbf{A}^T \mathbf{A})^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{B}^T \mathbf{B})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix}$
 $= \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T + \mathbf{B} (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T$.
- Význam: $[\mathbf{A} \quad \mathbf{B}]$ regulární a $\mathbf{A}^T \mathbf{B} = \mathbf{0}$ implikuje $\text{rng } \mathbf{B} = (\text{rng } \mathbf{A})^\perp$. To souhlasí s tím, že $\mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ je ort. projektor na $\text{rng } \mathbf{A}$ a $\mathbf{B} (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T$ je ort. projektor na $\text{rng } \mathbf{B} = (\text{rng } \mathbf{A})^\perp$.

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

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

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

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

je polynom dvou proměnných třetího stupně, kde např. x^2y 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. $f(x, y) = x^2y - 5y^3$ je homogenní stupně tři.

Vidíme, že afinní funkce (3.17) 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

Necht' 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 splňují

$$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = 0, \quad (6.4)$$

¹Tohle samozřejmě platí jen pro funkce na vektorovém prostoru \mathbb{R}^n . Pokud uvažujeme funkce na abstraktním (tedy definovaným axiomy) 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.

kde funkce $p_A: \mathbb{R} \rightarrow \mathbb{R}$ 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 $\mathbf{0}$). Speciálně, velikost vlastních vektorů nehraje roli a je proto zvykem je normalizovat, $\|\mathbf{v}\| = 1$.

Příklad 6.1. Vlastní čísla matice $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ jsou řešeními rovnice

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

Tato kvadratická rovnice má dva kořeny $\lambda = (5 \pm \sqrt{33})/2$, což jsou tedy 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 (2.6) plyne (promyslete!), že charakteristický polynom má stupeň n . Podle *základní věty algebry* má tedy právě n komplexních kořenů, počítáme-li k -násobný kořen k -krát. 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 kvůli násobnosti. Tomuto seznamu vlastních čísel se říká **spektrum** matice.

Necht' $\lambda_1, \dots, \lambda_n$ 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. Otázka, kdy se tak stane a co to znamená, není jednoduchá a podrobně ji zde diskutovat nepotřebujeme. Řekneme jen, že existuje dobrý důvod vlastní vektory vybrat tak, aby hodnota matice \mathbf{V} byla co možná největší.

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

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

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

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

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.*

Důkaz této věty neuvádíme. Větě 6.1 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 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 §2.7) (přesvědčte se, že druhá rovnost v (6.7) platí!).

Všimněme si, že dle Věty 3.6 (použité dvakrát na výraz $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$) máme

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

Když $\text{rank } \mathbf{A} < n$, některá vlastní čísla jsou nulová a můžeme tedy vynechat sloupce+řádky matice $\mathbf{\Lambda}$ a sloupce matice \mathbf{V} odpovídající nulovým vlastním číslům. To odpovídá vynechání nulových sčítanců v sumě dyád v (6.7).

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

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

Vlastní čísla a vektory jsou rozsáhlé téma, které jsme zde zdaleka nevyčerpali. 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 čísel je komutativní), máme

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} x_i x_j = \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n (a_{ij} + a_{ji}) x_i x_j = \frac{1}{2} \mathbf{x}^T (\mathbf{A} + \mathbf{A}^T) \mathbf{x}. \quad (6.10)$$

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 ($\mathbf{A}^T = \mathbf{A}$). V tom případě tedy $\frac{1}{2}(\mathbf{A} + \mathbf{A}^T) = \mathbf{A}$.

Je poučné dokázat rovnost (6.10) i jinak. Každou čtvercovou matici lze jednoznačně napsat jako součet symetrické a antisymetrické části (viz Cvičení 2.11):

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

Ale pro každé \mathbf{x} máme

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

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.

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. □

Symetrickou matici \mathbf{A} nazýváme

- **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 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.

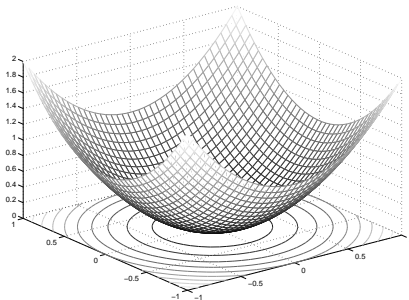
Důkaz. Je-li \mathbf{A} pozitivně semidefinitní, kvadratická forma nemůže být záporná a zároveň pro $\mathbf{x} = \mathbf{0}$ je nulová, proto v bodě $\mathbf{x} = \mathbf{0}$ (i když možná i jinde) nabývá svého minima. Je-li \mathbf{A} indefinitní a např. $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$, bod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

Ze spektrálního rozkladu (6.7) máme

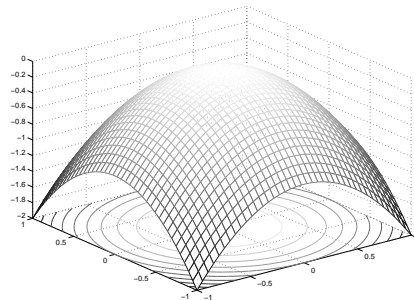
$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \mathbf{x} = \mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 = g(\mathbf{y}) = g(\mathbf{V}^T \mathbf{x}). \quad (6.11)$$

Vidíme, že substituce $\mathbf{x} = \mathbf{V} \mathbf{y}$ (tj. $\mathbf{y} = \mathbf{V}^T \mathbf{x}$) diagonalizovala matici kvadratické formy. Protože matice \mathbf{V} je ortogonální (a proto regulární), transformace $\mathbf{x} = \mathbf{V} \mathbf{y}$ je isometrie a navíc vzájemně jednoznačná (neboli bijekce, viz §1.1.2). Tedy funkce f a g se liší jen otočením příp. zrcadlením.

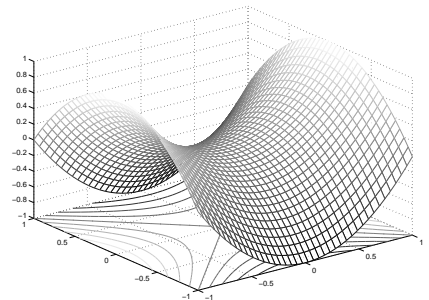
Výhoda je, že vlastnosti kvadratické formy jsou mnohem lépe patrné z jejího diagonálního tvaru g než z původního f . Např. pro případ dvou proměnných ($n = 2$) si grafy diagonální formy g snadno představíme. Jsou-li obě vlastní čísla kladná, funkce g vypadá jako ‘dolík’. Jsou-li obě vlastní čísla záporná, funkce g vypadá jako ‘kopec’. Mají-li vlastní čísla opačná znaménka, tvarem je ‘sedlo’:



$$g(y_1, y_2) = y_1^2 + y_2^2$$

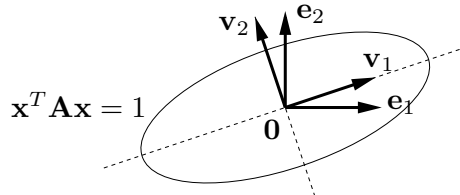


$$g(y_1, y_2) = -y_1^2 - y_2^2$$



$$g(y_1, y_2) = y_1^2 - y_2^2$$

Vrstevnice kvadratické formy s oběma vlastními čísly kladnými je elipsa, směry jejích hlavních os jsou vlastní vektory:



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. To je snadný důsledek diagonalizace (6.11). Protože transformace $\mathbf{x} = \mathbf{V} \mathbf{y}$ je bijekce, platí např. (promyslete!)

$$\forall \mathbf{x} \in \mathbb{R}^n: \mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0 \quad \Longleftrightarrow \quad \forall \mathbf{y} \in \mathbb{R}^n: \mathbf{y}^T \mathbf{\Lambda} \mathbf{y} \geq 0.$$

Tedy definitnost matice \mathbf{A} je stejná jako definitnost matice $\mathbf{\Lambda}$. Ale definitnost diagonální matice $\mathbf{\Lambda}$ je okamžitě patrná ze znamének čísel λ_i . Např. výraz (6.11) je nezáporný pro každé \mathbf{y} , právě když všechna λ_i jsou nezáporná. \square

6.3 Kvadratická funkce

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

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c, \quad (6.12)$$

kde $\mathbf{A}^T = \mathbf{A}$. Oproti⁴ kvadratické formě tedy přibyly lineární a konstantní členy. Všimněte si, že pro $n = 1$ je (6.12) 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

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

Výraz na pravé straně je kvadratická forma s počátkem posunutým do bodu \mathbf{x}_0 , plus konstanta. Této úpravě se říká **doplňení na čtverec**. Znáte ji pro případ $n = 1$, neboť tak se na základní škole odvozuje vzorec pro kořeny kvadratické rovnice jedné proměnné. Spočtěme \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

$$\mathbf{b} = -2\mathbf{A} \mathbf{x}_0, \quad (6.14a)$$

$$c = \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + y_0, \quad (6.14b)$$

z čehož spočítáme \mathbf{x}_0 a y_0 . Pokud soustava (6.14a) nemá řešení, doplňení na čtverec není možné.

Pokud je doplňení na čtverec možné, vyšetření extrémů kvadratické funkce se neliší od vyšetření extrémů kvadratické formy, protože rozdíl je jen v posunutí \mathbf{x}_0 . Pokud doplňení na čtverec možné není, kvadratická funkce extrém nemá (toto tvrzení uvádíme bez důkazu). Extrémy kvadratických funkcí lze také hledat pomocí derivací, ale to si ukážeme až později.

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 doplňení 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

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

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). Účelová funkce této úlohy je kvadratická, je totiž

$$\begin{aligned} \|\mathbf{Ax} - \mathbf{b}\|^2 &= (\mathbf{Ax} - \mathbf{b})^T (\mathbf{Ax} - \mathbf{b}) \\ &= (\mathbf{x}^T \mathbf{A}^T - \mathbf{b}^T) (\mathbf{Ax} - \mathbf{b}) \\ &= \mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - \mathbf{x}^T \mathbf{A}^T \mathbf{b} - \mathbf{b}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{b} \\ &= \mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - 2\mathbf{b}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{b}, \end{aligned} \quad (6.15)$$

kde jsme použili rovnost $\mathbf{b}^T \mathbf{Ax} = (\mathbf{b}^T \mathbf{Ax})^T = \mathbf{x}^T \mathbf{A}^T \mathbf{b}$ (neboť skalár je roven své transpozici). Extrém této kvadratické funkce můžeme najít doplněním na čtverec (viz §6.3). Soustava (6.14a) bude mít tvar $\mathbf{A}^T \mathbf{Ax}_0 = \mathbf{A}^T \mathbf{b}$ (\mathbf{A}, \mathbf{b} zde samozřejmě označuje něco jiného než v (6.14a)), 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

$$\mathbf{x}^T \mathbf{A}^T \mathbf{Ax} = (\mathbf{Ax})^T \mathbf{Ax} = \|\mathbf{Ax}\|^2 \geq 0. \quad (6.16)$$

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

6.3.1 Kvadrika

Vrstevnice kvadratické funkce se nazývá **kvadrika**. Tedy kvadrika je množina⁵

$$\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{x} + c = 0\} \quad (6.17)$$

Když $\mathbf{b} = \mathbf{0}$, množina (6.17) je vrstevnice kvadratické formy a kvadrika má tedy střed v počátku. Když \mathbf{A} je diagonální, kvadrika má osy rovnoběžné se souřadnicovými osami. Když $\mathbf{A} = \mathbf{0}$, množina (6.17) je pouhá nadrovina. Když \mathbf{A} nemá plnou hodnotu, kvadrika je degenerovaná. Množina (6.17) může být i prázdná.

Jestliže kvadratická funkce dovoluje doplnění na čtverec, můžeme množinu (6.17) psát jako

$$\{\mathbf{x} \in \mathbb{R}^n \mid (\mathbf{x} - \mathbf{x}_0)^T \mathbf{A}(\mathbf{x} - \mathbf{x}_0) + y_0 = 0\}, \quad (6.18)$$

což je 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 povrch **elipsoidu**⁶. 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*).

⁵Bez ztráty obecnosti uvažujeme vrstevnici výšky nula (tj. množina kořenů) kvadratické funkce (6.12), neboť c můžeme zvolit libovolně.

⁶Někteří autoři myslí elipsoidem množinu i s vnitřkem, někteří jen její hranici. Rozdíl je stejný jako mezi sférou a koulí.

6.4 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. Jaká jsou vlastní čísla a vlastní vektory (a) nulové, (b) jednotkové, (c) diagonální matice? Jaká jsou vlastní čísla trojúhelníkové matice?
- 6.4. Známe vlastní čísla a vektory matice \mathbf{A} . Jaká jsou vlastní čísla a vektory matice $\mathbf{A} + \alpha \mathbf{I}$?
- 6.5. Necht' $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{B} \in \mathbb{R}^{n \times m}$. Dokažte, že nenulová vlastní čísla matic \mathbf{AB} a \mathbf{BA} jsou stejná. Jaký je vztah jejich vlastních vektorů?
- 6.6. (★) Ukažte, že dvě čtvercové matice komutují, právě když mají stejné vlastní vektory.
- 6.7. Pro dané matice určete, zda jsou pozitivně/negativně (semi)definitní nebo indefinitní:

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

- 6.8. Mějme matici $\mathbf{A} = \begin{bmatrix} 1 & -3 \\ 2 & -4 \end{bmatrix}$. Která z následujících tvrzení jsou pravdivá?
- a) Výraz $\mathbf{x}^T \mathbf{A} \mathbf{x}$ je nezáporný pro každé $\mathbf{x} \in \mathbb{R}^2$.
 - b) Výraz $\mathbf{x}^T \mathbf{A} \mathbf{x}$ je nekladný pro každé $\mathbf{x} \in \mathbb{R}^2$.
 - c) Funkce $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ má v bodě $\mathbf{x} = \mathbf{0}$ extrém.
- 6.9. Máme kvadratickou formu dvou proměnných $f(x, y) = 3x^2 + 2xy + 3y^2$.
- a) 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} .
 - b) Najděte $a, b \in \mathbb{R}$ a ortogonální \mathbf{U} tak, že $f(x, y) = au^2 + bv^2$, kde $\begin{bmatrix} u \\ v \end{bmatrix} = \mathbf{U} \begin{bmatrix} x \\ y \end{bmatrix}$.
 - c) Nakreslete množinu bodů (u, v) splňujících $au^2 + bv^2 = 1$.
 - d) Transformujte tuto množinu do souřadnic (x, y) a nakreslete.
- 6.10. Je množina $\{(x, y) \in \mathbb{R}^2 \mid x^2 - 3xy + y^2 = 1\}$ elipsa nebo hyperbola? Odůvodněte.
- 6.11. (★) Napište v Matlabu funkci `ellipse(A)`, 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).
- 6.12. Ukažte, že je-li $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ spektrální rozklad symetrické matice \mathbf{A} , platí $\mathbf{A}^n = \mathbf{V} \mathbf{\Lambda}^n \mathbf{V}^T$.
- 6.13. Následující kvadratické funkce napište ve tvaru (6.12) se symetrickou \mathbf{A} . Pak najděte jejich extrémy a určete typ každého extrému (použijte doplnění na čtverec).
- a) $f(x, y) = x^2 + 4xy - 2y^2 + 3x - 6y + 5$
 - b) $f(x, y) = x^2 + 2y^2 - 2xy + 2x - y$

- 6.14. Máme neorientovaný graf (V, E) s množinou vrcholů $V = \{1, \dots, n\}$ a množinou hran $E \subseteq \binom{V}{2} = \{\{i, j\} \mid i, j \in V, i \neq j\}$. Každému vrcholu $i \in V$ je přiřazeno číslo $x_i \in \mathbb{R}$, tato čísla tvoří vektor $\mathbf{x} \in \mathbb{R}^n$. Je dána funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ jako $f(\mathbf{x}) = \frac{1}{2} \sum_{\{i, j\} \in E} (x_i - x_j)^2$.
- Ukažte, že f je kvadratická forma.
 - Jaká je definitnost této kvadratické formy?
 - Pro jaká \mathbf{x} platí $f(\mathbf{x}) = 0$?
 - Necht' je graf zadán maticí sousednosti $\mathbf{A} \in \{0, 1\}^{n \times n}$ tak, že $a_{ij} = 1$ právě když $\{i, j\} \in E$. Najděte matici $\mathbf{L} \in \mathbb{R}^{n \times n}$ tak, že $f(\mathbf{x}) = \mathbf{x}^T \mathbf{L} \mathbf{x}$. Hledejte co nejjednodušší vztah pro \mathbf{L} . Použijte přitom kromě matice \mathbf{A} také diagonální matici $\mathbf{D} = \text{diag}(\mathbf{A}\mathbf{1})$; jaký je význam matice \mathbf{D} ?

Poznamenejme, že funkci f (příp. matici \mathbf{L}) se říká *Laplacián* grafu (V, E) .

- 6.15. Musí mít pozitivně semidefinitní matice na diagonále nezáporné prvky? Odpověď dokažte.
- 6.16. Dokažte, že matice $\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}$ je pozitivně definitní pro každou matici \mathbf{A} a každé $\mu > 0$.
- 6.17. Dokažte, že symetrická matice je pozitivně definitní, právě když je pozitivně semidefinitní a invertovatelná.
- 6.18. Dokažte, že každá symetrická pozitivně definitní matice je invertovatelná a její inverze je také pozitivně definitní. Dokažte
- s použitím spektrálního rozkladu,
 - bez použití spektrálního rozkladu.
- 6.19. Dokažte:
- Pro každou matici \mathbf{A} je matice $\mathbf{A}^T \mathbf{A}$ pozitivně semidefinitní.
 - Matice $\mathbf{A}^T \mathbf{A}$ je pozitivně definitní, právě když matice \mathbf{A} má lineárně nezávislé sloupce,
 - Pro každou pozitivně semidefinitní matici \mathbf{B} existuje matice \mathbf{A} tak, že $\mathbf{B} = \mathbf{A}^T \mathbf{A}$.
 - Pro každou pozitivně definitní matici \mathbf{B} existuje regulární matice \mathbf{A} tak, že $\mathbf{B} = \mathbf{A}^T \mathbf{A}$.
- 6.20. (★) Pozitivně semidefinitní symetrické matice lze vnímat jako zobecnění nezáporných čísel. Proto se někdy pozitivní semidefinitnost značí $\mathbf{A} \succeq \mathbf{0}$. Zápis $\mathbf{A} \preceq \mathbf{B}$ je pak zkratkou $\mathbf{B} - \mathbf{A} \succeq \mathbf{0}$. Dokažte, že relace \preceq je částečné uspořádání (tj. reflexivní, tranzitivní a antisymetrická) na množině symetrických matic $n \times n$.
- 6.21. (★) Na základě podobnosti relace \preceq z předchozího cvičení a relace \leq na množině \mathbb{R} bychom očekávali, že:
- Pokud $\mathbf{A} \preceq \mathbf{B}$ a $\mathbf{C} \preceq \mathbf{D}$, potom $\mathbf{A} + \mathbf{C} \preceq \mathbf{B} + \mathbf{D}$.
 - Pokud $\mathbf{A} \succeq \mathbf{0}$ a $\alpha \geq 0$, potom $\alpha \mathbf{A} \succeq \mathbf{0}$.
 - Pokud $\mathbf{A} \succeq \mathbf{0}$, potom $\mathbf{A}^2 \succeq \mathbf{0}$.
 - Pokud $\mathbf{A} \succeq \mathbf{0}$ a $\mathbf{B} \succeq \mathbf{0}$, potom $\mathbf{AB} \succeq \mathbf{0}$.
 - Pokud $\mathbf{A} \succeq \mathbf{0}$ a $\mathbf{B} \succeq \mathbf{0}$, potom $\mathbf{ABA} \succeq \mathbf{0}$.

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

- 6.22. Najděte všechna vlastní čísla projektoru. Odpověď naleznete algebraicky z idempotence projektoru a poté odůvoděte geometrickou úvahou.

- 6.23. Geometrickou úvahou najděte aspoň dva vlastní vektory a příslušná vlastní čísla Householderovy matice ze Cvičení 4.13.
- 6.24. Je známo, že libovolnou rotaci ve třírozměrném prostoru lze realizovat jako rotaci kolem jisté přímky (jdoucí počátkem) o jistý úhel. Geometrickou úvahou zjistěte co nejvíce o vlastních číslech a vektorech rotační matice rozměru 3×3 .
- 6.25. Dokažte, že ortogonální projektor (tj. matice reprezentující ortogonální projekci) je pozitivně semidefinitní. Jaký to má geometrický význam?

Nápověda a řešení

- 6.4. Vlastní čísla se zvětší o α . Vlastní vektory jsou stejné.
- 6.5. Necht' $\mathbf{ABv} = \lambda \mathbf{v} \neq \mathbf{0}$. Z toho plyne $\mathbf{BABv} = \lambda \mathbf{Bv}$, tedy $\mathbf{BAu} = \lambda \mathbf{u}$ kde $\mathbf{u} = \mathbf{Bv}$. Zároveň $\mathbf{u} \neq \mathbf{0}$ protože jinak by bylo $\lambda \mathbf{v} = \mathbf{0}$.
- 6.7. indefinitní, pozitivně definitní, indefinitní, pozitivně semidefinitní
- 6.8. Žádné tvrzení neplatí, protože matice není symetrická. Musíme ji nejdříve symetrizovat, tj. vzít matici $\frac{1}{2}(\mathbf{A} + \mathbf{A}^T)$, a pak teprv počítat vlastní čísla.
- 6.9.a) $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$
- 6.9.b) $a = 2, b = 4, \mathbf{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$
- 6.10. Hyperbola, neboť \mathbf{A} má nenulová vlastní čísla opačných znamének.
- 6.13.a) Převod na tvar (6.12): $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ -6 \end{bmatrix}, c = 5$. Doplnění na čtverec existuje (už proto, že \mathbf{A} je regulární). Funkce nemá extrém (protože \mathbf{A} je indefinitní), má sedlo v bodě $(\frac{1}{2}, -1)$.
- 6.13.b) Má minimum v bodě $-(3, 1)/2$.
- 6.14.a) Je $(x_i - x_j) = x_i^2 - 2x_i x_j + x_j^2$. Tedy f je homogenní polynom stupně dva, což je kvadratická forma.
- 6.14.b) Funkce je součet čtverců (druhých mocnin), tedy $f(\mathbf{x}) \geq 0$ pro všechna $\mathbf{x} \in \mathbb{R}^n$, tedy f je pozitivně semidefinitní. Zároveň $f(\mathbf{x}) = 0$ když všechny složky x_i jsou stejné (tedy $\mathbf{x} = \alpha \mathbf{1}$ pro libovolné $\alpha \in \mathbb{R}$), tedy není pozitivně definitní.
- 6.14.c) Zjevně $f(\mathbf{x}) = 0$ právě když $x_i = x_j$ pro všechna $\{i, j\} \in E$. Když graf je souvislý, nastane to právě když všechna x_i jsou stejná. Jinak to nastane právě když jsou x_i stejná v každé komponentě grafu.
- 6.14.d) Prvek d_{ii} matice \mathbf{D} je stupeň (tedy počet incidentních hran) vrcholu i . Je $\mathbf{L} = \mathbf{D} - \mathbf{A}$, neboť

$$\frac{1}{2} \sum_{\{i,j\}} (x_i - x_j)^2 = \frac{1}{2} \sum_{\{i,j\}} x_i^2 - \sum_{\{i,j\}} x_i x_j + \frac{1}{2} \sum_{\{i,j\}} x_j^2 = \sum_{\{i,j\}} x_i^2 - \sum_{\{i,j\}} x_i x_j = \mathbf{x}^T \mathbf{D} \mathbf{x} - \mathbf{x}^T \mathbf{A} \mathbf{x}.$$

- 6.15. Musí. Stačí vzít $\mathbf{x} = \mathbf{e}_i$ vektory standardní báze.
- 6.16. $\mathbf{x}^T (\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}) \mathbf{x} = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} + \mu \mathbf{x}^T \mathbf{I} \mathbf{x} = \|\mathbf{Ax}\|^2 + \mu \|\mathbf{x}\|^2 > 0$ pro každé $\mathbf{x} \neq \mathbf{0}$.
- 6.17. Necht' $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$. Matice \mathbf{A} je pozitivně semidefinitní, právě když jsou všechny diagonální prvky matice $\mathbf{\Lambda}$ (což jsou vlastní čísla matice \mathbf{A}) nezáporné. Dle (6.8) je \mathbf{A} invertovatelná, právě když $\mathbf{\Lambda}$ je invertovatelná, neboli $\mathbf{\Lambda}$ má na diagonále nenulové prvky. Ale \mathbf{A} je pozitivně definitní, právě když její vl. čísla jsou kladná, tj. nezáporná a nenulová.

- 6.18.a) Je $\mathbf{A}^{-1} = (\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T)^{-1} = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^T$. Ale definitnost matice \mathbf{A} je stejná jako definitnost matice $\mathbf{\Lambda}$. Je jasné, že pokud diagonální prvky λ_i matice $\mathbf{\Lambda}$ jsou kladné, pak jsou kladné i diagonální prvky $1/\lambda_i$ matice $\mathbf{\Lambda}^{-1}$.
- 6.18.b) Necht' \mathbf{A} je pozitivně definitní, tedy $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ pro každé $\mathbf{x} \neq \mathbf{0}$. Kdyby \mathbf{A} nebyla invertovatelná (tj. regulární), dle Věty 3.7 by měla netriviální nulový prostor, tedy $\mathbf{A}\mathbf{x} = \mathbf{0}$ pro nějaké $\mathbf{x} \neq \mathbf{0}$. Ale to nelze, protože pak by také $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$.
- Necht' \mathbf{A} je invertovatelná. Položme $\mathbf{y} = \mathbf{A}\mathbf{x}$ a $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$. Víme, že $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ pro každé $\mathbf{x} \neq \mathbf{0}$. Z toho $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{y}^T\mathbf{A}^{-T}\mathbf{A}\mathbf{A}^{-1}\mathbf{y} = \mathbf{y}^T\mathbf{A}^{-1}\mathbf{y} > 0$. Protože zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ je bijekce, je $\mathbf{y}^T\mathbf{A}^{-1}\mathbf{y} > 0$ pro každé $\mathbf{y} \neq \mathbf{0}$.
- 6.19.a) $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})^T\mathbf{A}\mathbf{x} = \|\mathbf{A}\mathbf{x}\|^2 \geq 0$
- 6.19.b) Víme, že $\mathbf{A}^T\mathbf{A}$ je p.s.d. pro každou \mathbf{A} . Tedy stačí dokázat, že $\mathbf{A}^T\mathbf{A}$ je invertovatelná, právě když \mathbf{A} má l.n. sloupce. To jsme dokázali v (5.5).
- 6.19.c) Ve spektrálním rozkladu $\mathbf{B} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ má $\mathbf{\Lambda}$ nezáporné diagonální prvky. Položme $\mathbf{A} = \mathbf{\Lambda}^{1/2}\mathbf{V}^T$.
- 6.19.d) Ve spektrálním rozkladu $\mathbf{B} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ má $\mathbf{\Lambda}$ kladné diagonální prvky. Položme $\mathbf{A} = \mathbf{\Lambda}^{1/2}\mathbf{V}^T$.
- 6.20. Relace je reflexivní, když $\mathbf{A} \preceq \mathbf{A}$ pro každou \mathbf{A} , neboli $\mathbf{A} - \mathbf{A} \succeq \mathbf{0}$, což platí.
- Pro další si všimněte, že $\mathbf{A} \preceq \mathbf{B}$ znamená $(\forall \mathbf{x})(\mathbf{x}^T\mathbf{A}\mathbf{x} \leq \mathbf{x}^T\mathbf{B}\mathbf{x})$. Relace je antisymetrická, když $\mathbf{A} \preceq \mathbf{B} \preceq \mathbf{A} \Rightarrow \mathbf{A} = \mathbf{B}$. Ale $\mathbf{x}^T\mathbf{A}\mathbf{x} \leq \mathbf{x}^T\mathbf{B}\mathbf{x} \leq \mathbf{x}^T\mathbf{A}\mathbf{x}$ implikuje $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{B}\mathbf{x}$. To platí pro všechna \mathbf{x} , protože \mathbf{A}, \mathbf{B} jsou symetrické.
- Relace je tranzitivní, když $\mathbf{A} \preceq \mathbf{B} \preceq \mathbf{C} \Rightarrow \mathbf{A} \preceq \mathbf{C}$. To platí, neboť $\mathbf{x}^T\mathbf{A}\mathbf{x} \leq \mathbf{x}^T\mathbf{B}\mathbf{x} \leq \mathbf{x}^T\mathbf{C}\mathbf{x} \Rightarrow \mathbf{x}^T\mathbf{A}\mathbf{x} \leq \mathbf{x}^T\mathbf{C}\mathbf{x}$.
- 6.22. Algebraicky: Projektor \mathbf{P} splňuje $\mathbf{P}\mathbf{P} = \mathbf{P}$ (idempotence). Tedy z $\mathbf{P}\mathbf{v} = \lambda\mathbf{v}$ plyne (násobením maticí \mathbf{P}) $\mathbf{P}\mathbf{v} = \lambda\mathbf{P}\mathbf{v}$. To může platit jen pro $\lambda = 1$ nebo pro $\mathbf{P}\mathbf{v} = \mathbf{0}$. Ale když $\mathbf{P}\mathbf{v} = \mathbf{0}$, tak z rovnice $\mathbf{P}\mathbf{v} = \lambda\mathbf{v}$ musí být $\lambda = 0$. Tedy vlastní čísla projektoru mohou být buď 1 nebo 0.
- Úvaha: Projekce na podprostor $X = \text{rng } \mathbf{P}$ promítne každý vektor $\mathbf{x} \in X$ na sebe, tedy $\mathbf{P}\mathbf{x} = \mathbf{x}$. Každý vektor $\mathbf{x} \in X^\perp$ se promítne do $\mathbf{0}$, tedy $\mathbf{P}\mathbf{x} = \mathbf{0}$. Každý vektor $\mathbf{x} \notin X \cup X^\perp$ se promítne do vektoru, který není rovnoběžný s \mathbf{x} . Shrnutí, \mathbf{P} má jen dvě vlastní čísla 1 a 0 (předpokládáme nyní, že $X^\perp \neq \{\mathbf{0}\}$, neboli \mathbf{P} není regulární). Vlastní vektory příslušné číslu 1 tvoří podprostor X , vlastní vektory příslušné číslu 0 tvoří podprostor X^\perp .
- 6.23. Transformace $\mathbf{f}(\mathbf{x}) = \mathbf{H}\mathbf{x}$ je zrcadlení okolo nadroviny s normálovým vektorem \mathbf{v} . Tedy $\mathbf{H}\mathbf{x} = \mathbf{x}$ pro všechna \mathbf{x} ležící v této nadrovině, neboli $\mathbf{v}^T\mathbf{x} = 0$. Dále je $\mathbf{H}\mathbf{v} = -\mathbf{v}$, neboť normála nadroviny zrcadlením změnila orientaci.
- 6.25. Projektor \mathbf{P} splňuje $\mathbf{P}^2 = \mathbf{P}$ a $\mathbf{P}^T = \mathbf{P}$ (viz §4.5 a Cvičení 5.17), z čehož plyne $\mathbf{P} = \mathbf{P}^T\mathbf{P}$. Nyní $\mathbf{z}^T\mathbf{P}\mathbf{z} = \mathbf{z}^T\mathbf{P}^T\mathbf{P}\mathbf{z} = \|\mathbf{P}\mathbf{z}\|^2 \geq 0$ (viz Cvičení 6.19). Geometrický význam je takový, že libovolný vektor \mathbf{z} svírá se svou ortogonální projekcí $\mathbf{x} = \mathbf{P}\mathbf{z}$ ostrý úhel nebo jsou na sebe kolmé (když $\mathbf{z} = \mathbf{0}$ nebo $\mathbf{x} = \mathbf{0}$), tedy $\mathbf{z}^T\mathbf{x} = \mathbf{z}^T\mathbf{P}\mathbf{z} \geq 0$.

Kapitola 7

Proložení bodů podprostorem

7.1 Problém nejmenší stopy

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

Věta 7.1. *Platí*

$$\min\{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\} = \lambda_1 \quad (7.1)$$

a minimální hodnota se nabývá pro $\mathbf{x} = \mathbf{v}_1$, tedy pro normalizovaný vlastní vektor příslušný nejmenšímu vlastnímu číslu.

Důkaz. Pohledme na rovnost (6.11). 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 (7.1) má stejnou minimální hodnotu jako úloha

$$\min\{\mathbf{y}^T \mathbf{\Lambda} \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^n, \mathbf{y}^T \mathbf{y} = 1\} = \min\{\lambda_1 y_1^2 + \dots + \lambda_n y_n^2 \mid y_1, \dots, y_n \in \mathbb{R}, y_1^2 + \dots + y_n^2 = 1\}. \quad (7.2)$$

Tuto úlohu snadno vyřešíme úvahou (viz Cvičení 7.1): minimum se nabývá pro $y_1 = 1$ a $y_2 = \dots = y_n = 0$, tj. pro $\mathbf{y} = \mathbf{e}_1$ (první vektor standardní báze). To odpovídá $\mathbf{x} = \mathbf{V} \mathbf{e}_1 = \mathbf{v}_1$ a $\mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = \lambda_1$. \square

Je-li matice \mathbf{A} symetrická, ze spektrálního rozkladu a cykličnosti stopy (viz §2.6) máme

$$\text{tr } \mathbf{A} = \text{tr}(\mathbf{V} \mathbf{\Lambda} \mathbf{V}^T) = \text{tr}(\mathbf{V}^T \mathbf{V} \mathbf{\Lambda}) = \text{tr } \mathbf{\Lambda} = \lambda_1 + \dots + \lambda_n, \quad (7.3)$$

kde $\lambda_1, \dots, \lambda_n$ jsou vlastní čísla matice \mathbf{A} . Podotkněme, že rovnost $\text{tr } \mathbf{A} = \lambda_1 + \dots + \lambda_n$ platí pro každou (ne nutně symetrickou) matici. Důkaz tohoto tvrzení neuvádíme.

Uvažujme podprostor dimenze k s ortonormální bází tvořenou sloupci matice $\mathbf{X} \in \mathbb{R}^{n \times k}$, tedy $\mathbf{X}^T \mathbf{X} = \mathbf{I}$. Každý vektor z tohoto podprostoru je lineární kombinace $\mathbf{x} = \mathbf{X} \mathbf{y}$ vektorů báze, kde $\mathbf{y} \in \mathbb{R}^k$ jsou souřadnice vektoru \mathbf{x} v této bázi. Tedy funkce

$$g(\mathbf{y}) = f(\mathbf{X} \mathbf{y}) = \mathbf{y}^T \mathbf{X}^T \mathbf{A} \mathbf{X} \mathbf{y} \quad (7.4)$$

je zúžení (neboli *restrikce*) kvadratické formy $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ na daný podprostor, vyjádřená v souřadnicích \mathbf{y} . Funkce g je kvadratická forma s maticí $\mathbf{X}^T \mathbf{A} \mathbf{X}$. Její stopu $\text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X})$ lze tedy

vidět jako *stopu matice* \mathbf{A} na podprostoru $\text{rng } \mathbf{X}$. Všimněme si, že tato stopa nezávisí na tom, jakou ortonormální bázi podprostoru zvolíme (viz Cvičení 4.16), neboť pro každou ortogonální $\mathbf{U} \in \mathbb{R}^{k \times k}$ je $\text{tr}[(\mathbf{XU})^T \mathbf{A}(\mathbf{XU})] = \text{tr}(\mathbf{U}^T \mathbf{X}^T \mathbf{A} \mathbf{XU}) = \text{tr}(\mathbf{U} \mathbf{U}^T \mathbf{X}^T \mathbf{A} \mathbf{X}) = \text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X})$.

V následující úloze (7.5), která zobecňuje rovnosti (7.3) a (7.1), hledáme takový podprostor, na kterém je stopa matice \mathbf{A} nejmenší.

Věta 7.2. *Necht' $k \leq n$. Platí*

$$\min\{\text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) \mid \mathbf{X} \in \mathbb{R}^{n \times k}, \mathbf{X}^T \mathbf{X} = \mathbf{I}\} = \lambda_1 + \cdots + \lambda_k, \quad (7.5)$$

tedy¹ minimální hodnota je součet k nejmenších vlastních čísel matice \mathbf{A} . Minimum se nabývá pro $\mathbf{X} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k]$, tedy pro sloupce \mathbf{X} rovné (normalizovaným) vlastním vektorům příslušným těmto vlastním číslům.

Důkaz. (★) Máme $\text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) = \text{tr}(\mathbf{X}^T \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \mathbf{X}) = \text{tr}(\mathbf{Y}^T \mathbf{\Lambda} \mathbf{Y})$, kde $\mathbf{X} = \mathbf{V} \mathbf{Y}$. Protože \mathbf{V} je ortogonální, $\mathbf{X}^T \mathbf{X} = \mathbf{I}$ platí právě když $\mathbf{Y}^T \mathbf{V}^T \mathbf{V} \mathbf{Y} = \mathbf{Y}^T \mathbf{Y} = \mathbf{I}$. Tedy úloha (7.5) má stejnou optimální hodnotu jako úloha

$$\min\{\text{tr}(\mathbf{Y}^T \mathbf{\Lambda} \mathbf{Y}) \mid \mathbf{Y} \in \mathbb{R}^{n \times k}, \mathbf{Y}^T \mathbf{Y} = \mathbf{I}\}. \quad (7.6)$$

Řešme úlohu (7.6). Z vlastností stopy (viz §2.6) máme

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

kde p_{11}, \dots, p_{nn} jsou diagonální prvky matice $\mathbf{P} = \mathbf{Y} \mathbf{Y}^T$.

Pro každou $\mathbf{Y} \in \mathbb{R}^{n \times k}$ čísla p_{11}, \dots, p_{nn} splňují rovnost

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

Dále splňují nerovnosti $0 \leq p_{ii} \leq 1$ (viz Cvičení 4.20). Uvažujme úlohu

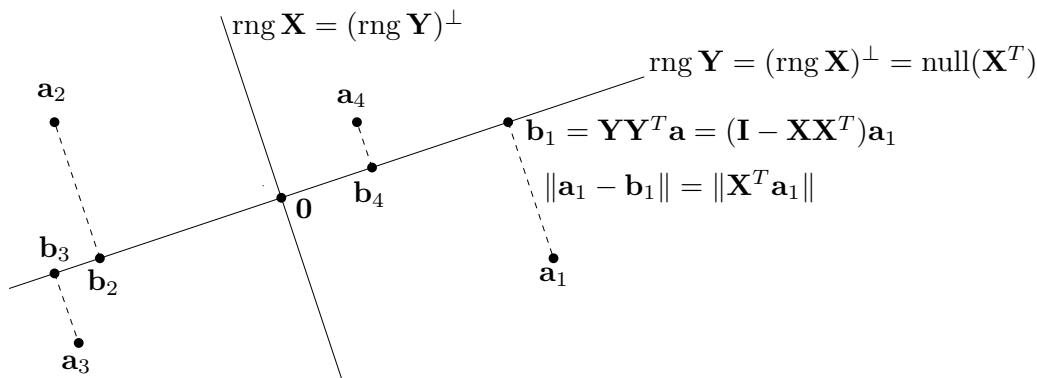
$$\min\{\lambda_1 p_{11} + \cdots + \lambda_n p_{nn} \mid 0 \leq p_{11}, \dots, p_{nn} \leq 1, p_{11} + \cdots + p_{nn} = k\}. \quad (7.7)$$

Tuto úlohu vyřešíme úvahou (viz Cvičení 11.3): její optimální hodnota je $\lambda_1 + \cdots + \lambda_k$ a optimální argument je $p_{11} = \cdots = p_{kk} = 1$ a $p_{k+1,k+1} = \cdots = p_{nn} = 0$. Ale matici $\mathbf{P} = \mathbf{Y} \mathbf{Y}^T$ s těmito diagonálními prvky lze realizovat volbou $\mathbf{Y} = \begin{bmatrix} \mathbf{I}_k \\ \mathbf{0} \end{bmatrix}$. Z toho $\mathbf{X} = \mathbf{V} \mathbf{Y} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k]$. \square

7.2 Proložení bodů podprostorem

Hledejme lineární podprostor dimenze k prostoru \mathbb{R}^n , který minimalizuje součet čtverců vzdáleností k daným bodům $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$. Podotkněme, že tuto úlohu nelze nijak převést na úlohu nejmenších čtverců (5.2). Pro pohodlí zavedeme matici $\mathbf{A} = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_m]^T \in \mathbb{R}^{m \times n}$ s řádky $\mathbf{a}_1^T, \dots, \mathbf{a}_m^T$. Ve statistice se této úloze říká *rozvoj podle hlavních komponent* (angl. *principal component analysis, PCA*) nebo *Karhunen-Loewův rozvoj*.

¹Namítnete, že úloha (7.5) není ve tvaru (1.8), protože optimalizujeme přes množinu *matic* a ne vektorů. To je ale jen záležitost značení, neboť matici rozměru $m \times n$ lze vnímat jako vektor o mn složkách.



Věta 7.3. Necht' $\mathbf{A}^T \mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ je spektrální rozklad matice

$$\mathbf{A}^T \mathbf{A} = \mathbf{a}_1 \mathbf{a}_1^T + \cdots + \mathbf{a}_n \mathbf{a}_n^T. \quad (7.8)$$

- Posledních k sloupců matice \mathbf{V} (tj. vlastní vektory matice (7.8) odpovídající k největším vlastním číslům) tvoří ortonormální bázi lineárního podprostoru dimenze k , který minimalizuje součet čtverců vzdáleností k bodům $\mathbf{a}_1, \dots, \mathbf{a}_m$.
- Prvních $n - k$ sloupců matice \mathbf{V} (tj. vlastní vektory matice (7.8) odpovídající $n - k$ nejmenším vlastním číslům) je ortonormální báze ortogonálního doplňku tohoto podprostoru.

Důkaz. Místo hledaného podprostoru je výhodnější hledat jeho ortogonální doplněk, který reprezentujeme ortonormální bází tvořenou sloupci matice $\mathbf{X} \in \mathbb{R}^{n \times (n-k)}$ (viz obrázek). Matice \mathbf{X} má $n - k$ sloupců, neboť dle Věty 4.1 je dimenze tohoto ortogonálního doplňku rovna $n - k$. Tedy hledaný podprostor je (dle (4.7a))

$$\text{rng } \mathbf{Y} = (\text{rng } \mathbf{X})^\perp = \text{null}(\mathbf{X}^T) = \{ \mathbf{a} \in \mathbb{R}^n \mid \mathbf{X}^T \mathbf{a} = \mathbf{0} \}. \quad (7.9)$$

Vzdálenost bodu $\mathbf{a} \in \mathbb{R}^n$ k podprostoru (7.9) je délka projekce na jeho ortogonální doplněk, tj. $\|\mathbf{X} \mathbf{X}^T \mathbf{a}\| = \|\mathbf{X}^T \mathbf{a}\|$ (viz §4.5). Součet čtverců vzdáleností všech bodů k podprostoru je

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

kde $\|\mathbf{A} \mathbf{X}\|$ zde značí maticovou (Frobeniovu) normu (4.27) matice $\mathbf{A} \mathbf{X}$. Řešíme tedy úlohu

$$\min \{ \|\mathbf{A} \mathbf{X}\| \mid \mathbf{X} \in \mathbb{R}^{n \times (n-k)}, \mathbf{X}^T \mathbf{X} = \mathbf{I} \}. \quad (7.10)$$

Jelikož $\|\mathbf{A} \mathbf{X}\|^2 = \text{tr}(\mathbf{X}^T \mathbf{A}^T \mathbf{A} \mathbf{X})$, řešení máme z Věty 7.2: rozdělíme-li (ortogonální) matici $\mathbf{V} = [\mathbf{X} \ \mathbf{Y}] \in \mathbb{R}^{n \times n}$ do dvou bloků $\mathbf{X} \in \mathbb{R}^{n \times (n-k)}$ a $\mathbf{Y} \in \mathbb{R}^{n \times k}$, pak \mathbf{X} je řešení úlohy (7.10) a (viz §4.5.1) sloupce \mathbf{Y} tvoří ortonormální bázi hledaného podprostoru (7.9). \square

Příklad 7.1. 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 $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3$. Přímka procházející počátkem (tedy $k = 1$), která minimalizuje součet čtverců vzdáleností k bodům, je množina

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

Rovina procházející počátkem (tedy $k = 2$), která minimalizuje součet čtverců vzdáleností k bodům, je množina

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

7.2.1 Nejblíží matice nižší hodnosti

Ukážeme, že úloha (7.10) je ekvivalentní následující úloze: k dané matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ hledáme nejblíží (ve smyslu Frobeniovy normy) matici \mathbf{B} hodnosti nejvýše k . Tedy

$$\min\{ \|\mathbf{A} - \mathbf{B}\| \mid \mathbf{B} \in \mathbb{R}^{m \times n}, \text{rank } \mathbf{B} \leq k \}. \quad (7.11)$$

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

Věta 7.4. *Optimální hodnota úlohy (7.11) je rovna optimální hodnotě úlohy (7.10) a nabývá se pro matici $\mathbf{B} = \mathbf{A}\mathbf{Y}\mathbf{Y}^T$, kde sloupce matice $\mathbf{Y} \in \mathbb{R}^{n \times k}$ tvoří (normalizované) vlastní vektory příslušné k největším vlastním číslům matice $\mathbf{A}^T\mathbf{A}$.*

Důkaz. Dle (3.13) má matice \mathbf{B} hodnost nejvýše k právě když její nulový prostor má hodnost nejméně $n - k$. To nastane právě když existuje matice $\mathbf{X} \in \mathbb{R}^{n \times (n-k)}$ s ortonormálními sloupci tak, že $\mathbf{B}\mathbf{X} = \mathbf{0}$. Tedy úloha (7.11) je ekvivalentní úloze

$$\min\{ \|\mathbf{A} - \mathbf{B}\| \mid \mathbf{B} \in \mathbb{R}^{m \times n}, \mathbf{X} \in \mathbb{R}^{n \times (n-k)}, \mathbf{X}^T\mathbf{X} = \mathbf{I}, \mathbf{B}\mathbf{X} = \mathbf{0} \}. \quad (7.12)$$

Označme řádky matice \mathbf{A} jako $\mathbf{a}_1^T, \dots, \mathbf{a}_m^T$ a řádky matice \mathbf{B} jako $\mathbf{b}_1^T, \dots, \mathbf{b}_m^T$. Podmínka $\mathbf{B}\mathbf{X} = \mathbf{0}$ jde psát jako $\mathbf{X}^T\mathbf{b}_1 = \dots = \mathbf{X}^T\mathbf{b}_m = \mathbf{0}$, což znamená, že body $\mathbf{b}_1, \dots, \mathbf{b}_m$ leží v podprostoru (7.9). Platí

$$\|\mathbf{A} - \mathbf{B}\|^2 = \|\mathbf{a}_i - \mathbf{b}_i\|^2 + \dots + \|\mathbf{a}_m - \mathbf{b}_m\|^2. \quad (7.13)$$

Tedy v úlohách (7.11) a (7.12) hledáme body $\mathbf{b}_1, \dots, \mathbf{b}_m$, které leží v podprostoru dimenze nejvýše k a jejichž součet čtverců vzdáleností k bodům \mathbf{a}_i je minimální. Viz obrázek v §7.2.

Ale pokud \mathbf{X} je pevné, máme

$$\min\{ \|\mathbf{a} - \mathbf{b}\| \mid \mathbf{b} \in \mathbb{R}^n, \mathbf{X}^T\mathbf{b} = \mathbf{0} \} = \|\mathbf{X}^T\mathbf{a}\| \quad (7.14)$$

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 (7.9). Minimalizace (7.13) za podmínky $\mathbf{B}\mathbf{X} = \mathbf{0}$ se pak rozpadá na m nezávislých úloh (7.14):

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

kde minimum se nabývá pro matici $\mathbf{B} = \mathbf{A}(\mathbf{I} - \mathbf{X}\mathbf{X}^T)$. Tedy úloha (7.12) (a také (7.11)) je ekvivalentní minimalizaci $\|\mathbf{A}\mathbf{X}\|$ za podmínky $\mathbf{X}^T\mathbf{X} = \mathbf{I}$, což je úloha (7.10).

Z úlohy (7.10) víme, že sloupce optimální matice $\mathbf{X} \in \mathbb{R}^{n \times (n-k)}$ jsou vlastní vektory příslušné $n - k$ nejmenším vlastním číslům matice $\mathbf{A}^T\mathbf{A}$. Dle (4.21) je $\mathbf{Y}\mathbf{Y}^T + \mathbf{X}\mathbf{X}^T = \mathbf{I}$, kde sloupce $\mathbf{Y} \in \mathbb{R}^{n \times k}$ jsou vlastní vektory příslušné k největším vlastním číslům matice $\mathbf{A}^T\mathbf{A}$. Tedy $\mathbf{B} = \mathbf{A}(\mathbf{I} - \mathbf{X}\mathbf{X}^T) = \mathbf{A}\mathbf{Y}\mathbf{Y}^T$. \square

7.2.2 Co když prokládáme afinním podprostorem?

Změňme nyní úlohu z Věty 7.3 tak, že místo lineárního podprostoru hledejme *afinní* podprostor dimenze k , který minimalizuje součet čtverců vzdáleností k bodům $\mathbf{a}_1, \dots, \mathbf{a}_m$.

Věta 7.5. *Afinní podprostor, který minimalizuje součet čtverců vzdáleností k bodům $\mathbf{a}_1, \dots, \mathbf{a}_m$, prochází těžištěm $\bar{\mathbf{a}} = \frac{1}{m}(\mathbf{a}_1 + \dots + \mathbf{a}_m)$ těchto bodů.*

Důkaz. Hledaný afinní podprostor parametrizujeme jako (srov. (7.9) a Věta 3.10)

$$\{\mathbf{a} \in \mathbb{R}^n \mid \mathbf{X}^T \mathbf{a} = \mathbf{y}\}, \quad (7.15)$$

kde $\mathbf{X} \in \mathbb{R}^{n \times k}$ je matice s ortonormálními sloupci a $\mathbf{y} \in \mathbb{R}^m$. Vzdálenost bodu $\mathbf{a} \in \mathbb{R}^n$ od tohoto podprostoru je $\|\mathbf{X}^T \mathbf{a} - \mathbf{y}\|$ (viz Cvičení 4.17). Součet čtverců vzdáleností všech bodů k podprostoru je tedy

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

Tento výraz tedy musíme minimalizovat přes proměnné \mathbf{X} a \mathbf{y} za podmínky $\mathbf{X}^T \mathbf{X} = \mathbf{I}$. Pokud je \mathbf{X} libovolné pevné a minimalizujeme pouze přes \mathbf{y} , minimum se nabývá v bodě $\mathbf{y} = \mathbf{X}^T \bar{\mathbf{a}}$ (viz Cvičení 5.3.a). Z toho plyne (proč?), že podprostor (7.15) obsahuje bod $\bar{\mathbf{a}}$. \square

Nyní je řešení jasné: nejprve posuneme body $\mathbf{a}_1, \dots, \mathbf{a}_m$ tak, aby jejich těžiště leželo v počátku, a potom najdeme (lineární) podprostor, který minimalizuje součet čtverců vzdáleností k posunutým bodům, dle Věty 7.3. Jeho posunutím zpět pak získáme kýžený afinní podprostor.

7.3 Přeurčené homogenní lineární soustavy

Řešme homogenní lineární soustavu

$$\mathbf{A}\mathbf{x} = \mathbf{0}, \quad (7.16)$$

kde $\mathbf{A} \in \mathbb{R}^{m \times n}$. Její množinou řešení je null \mathbf{A} , což je lineární podprostor \mathbb{R}^n dimenze $n - \text{rank } \mathbf{A}$ viz (3.13). 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

$$\min\{\|\mathbf{A}\mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\}. \quad (7.17)$$

Protože $\|\mathbf{A}\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x}$, dle (7.1) je řešením této úlohy vlastní vektor matice $\mathbf{A}^T \mathbf{A}$ příslušný jejímu nejmenšímu vlastnímu číslu.

Příklad 7.2 (Přibližné proložení bodů kuželosečkou). 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 nemusí existovat kuželosečka, která jimi prochází. 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 obtížné.

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

$$f(x_1, y_1)^2 + \cdots + f(x_m, y_m)^2. \quad (7.18)$$

Tato formulace ale nevyjadřuje to, co chceme, protože minimum výrazu (7.18) se nabývá pro $a = b = c = d = e = f = 0$. V tomto případě množina K není křivka, ale celá rovina \mathbb{R}^2 .

Ve skutečnosti navíc potřebujeme, aby aspoň jedno z čísel a, b, c, d, e, f bylo nenulové. Toho dosáhneme uvalením podmínky

$$a^2 + b^2 + c^2 + d^2 + e^2 + f^2 = 1. \quad (7.19)$$

Všimněte si, že pro každou kuželosečku můžeme dosáhnout splnění této podmínky, protože vynásobením vektoru (a, b, c, d, e, f) libovolným nenulovým číslem nezmění množinu K .

Minimalizace (7.18) za podmínky (7.19) se dá psát jako minimalizace $\|\mathbf{A}\boldsymbol{\theta}\|^2$ za podmínky $\boldsymbol{\theta}^T \boldsymbol{\theta} = 1$, kde

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

□

Obecněji, soustavu (7.16) nazveme přeúčtovanou tehdy, když dimenze jejího prostoru řešení je nižší než nějaká předem známá dimenze $n-k$. Přibližné řešení soustavy pak vede na úlohu (7.10), která je ovšem ekvivalentní úloze (7.11). Úloha (7.11) odpovídá tomu, že co nejméně změním matici \mathbf{A} , aby prostor řešení soustavy (7.16) měl kýženou dimenzi $n-k$. Neboli nejprve najdeme matici \mathbf{B} s hodnotami k nejbližší matici \mathbf{A} a potom řešíme soustavu $\mathbf{B}\mathbf{x} = \mathbf{0}$, jejíž prostor řešení již má dimenzi $n-k$.

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. Formulujme přibližné řešení nehomogenní soustavy takto: pokud soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ nemá řešení, změňme 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.

7.4 Singulární rozklad

Věta 7.6. Každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ lze rozložit jako

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (7.20)$$

kde matice $\mathbf{S} \in \mathbb{R}^{m \times n}$ je diagonální a $\mathbf{U} \in \mathbb{R}^{m \times m}$ a $\mathbf{V} \in \mathbb{R}^{n \times n}$ jsou ortogonální.

Připomeňme, že diagonální matice nemusí být čtvercová (viz začátek kapitoly 2). Diagonální prvky matice \mathbf{S} se nazývají **singulární čísla** matice \mathbf{A} a budeme je značit s_1, \dots, s_p , kde $p = \min\{m, n\}$. Je zvykem je zvolit nezáporné, což lze vždy zajistit vynásobením záporných singulárních čísel a odpovídajících sloupců \mathbf{U} a \mathbf{V} minus jedničkou. Sloupce matice \mathbf{U} příp. \mathbf{V} se nazývají levé příp. pravé **singulární vektory** matice \mathbf{A} . Rozklad (7.20) se nazývá **rozklad podle singulárních čísel** (*Singular Value Decomposition, SVD*). Všimneme si, že (např. použitím Věty 3.6) $\text{rank } \mathbf{A} = \text{rank } \mathbf{S}$, tedy hodnost matice je rovna počtu jejích nenulových singulárních čísel.

Singulární čísla a vektory matice \mathbf{A} mají úzký vztah k vlastním číslům a vektorům matic $\mathbf{A}^T \mathbf{A}$ a $\mathbf{A} \mathbf{A}^T$. Předpokládáme, že rozklad (7.20) existuje. Pak

$$\mathbf{A}^T \mathbf{A} = \mathbf{V} \mathbf{S} \mathbf{U}^T \mathbf{U} \mathbf{S} \mathbf{V}^T = \mathbf{V} \mathbf{S}^2 \mathbf{V}^T, \quad (7.21a)$$

$$\mathbf{A} \mathbf{A}^T = \mathbf{U} \mathbf{S} \mathbf{V}^T \mathbf{V} \mathbf{S} \mathbf{U}^T = \mathbf{U} \mathbf{S}^2 \mathbf{U}^T. \quad (7.21b)$$

Ale to jsou spektrální rozklady (6.7) symetrických pozitivně definitních matic $\mathbf{A}^T \mathbf{A}$ a $\mathbf{A} \mathbf{A}^T$. Tedy nenulová singulární čísla matice \mathbf{A} jsou druhé odmocniny nenulových vlastních čísel matic $\mathbf{A}^T \mathbf{A}$ a $\mathbf{A} \mathbf{A}^T$ (která jsou tudíž stejná, srov. Cvičení 6.5). Pravé a levé singulární vektory jsou vlastní vektory těchto matic příslušné nenulovým vlastním číslům.

Počítat SVD matice \mathbf{A} ze spektrálního rozkladu matice $\mathbf{A}^T \mathbf{A}$ nebo $\mathbf{A} \mathbf{A}^T$ ovšem není numericky vhodné, protože výpočet součinu $\mathbf{A}^T \mathbf{A}$ může vést ke zbytečným zaokrouhlovacím chybám (viz Příklad 5.2). Na SVD proto byly vymyšleny algoritmy, které se explicitnímu výpočtu $\mathbf{A}^T \mathbf{A}$ vyhýbají. Kdykoliv tedy hledáme vlastní čísla a vektory matice $\mathbf{A}^T \mathbf{A}$ (např. v úlohách (7.10) a (7.11)), měli bychom to dělat *pomocí SVD algoritmu*. Na druhou stranu, pokud nám možné snížení přesnosti nevadí, 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, $n \times n$).

7.4.1 Nejblížeší matice nižší hodnosti

Řešení úlohy (7.11) jsme odvodili ve Větě 7.4 pomocí spektrálního rozkladu matice $\mathbf{A}^T \mathbf{A}$. Následující klasická věta formuluje toto řešení elegantně pomocí SVD matice \mathbf{A} .

Věta 7.7 (Eckart-Young). *Necht' $\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T$ je SVD matice \mathbf{A} . Řešení úlohy (7.11) je*

$$\mathbf{B} = \mathbf{U} \mathbf{S}' \mathbf{V}^T = \sum_{i=1}^k s_i \mathbf{u}_i \mathbf{v}_i^T, \quad (7.22)$$

kde matice $\mathbf{S}' = \text{diag}(s_1, \dots, s_k, 0, \dots, 0)$, kde předpokládáme $s_1 \geq \dots \geq s_p \geq 0$. Tedy \mathbf{S}' má prvních k diagonálních prvků rovných diagonálním prvkům matice \mathbf{S} a ostatní diagonální prvky jsou vynulovány.

Důkaz. Ze vztahu SVD matice \mathbf{A} a spektrálního rozkladu matice $\mathbf{A}^T \mathbf{A}$ víme, že prvních k sloupců matice \mathbf{V} tvoří matici \mathbf{Y} . Tedy

$$\mathbf{B} = \mathbf{A} \mathbf{Y} \mathbf{Y}^T = \mathbf{U} \mathbf{S} \mathbf{V}^T \mathbf{Y} \mathbf{Y}^T = \mathbf{U} \mathbf{S} \begin{bmatrix} \mathbf{I}_k \\ \mathbf{0} \end{bmatrix} \mathbf{Y}^T = \mathbf{U} \mathbf{S} \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}^T = \mathbf{U} \mathbf{S}' \mathbf{V}^T. \quad \square$$

Optimální hodnota úlohy (7.11) je (viz (4.28))

$$\|\mathbf{A} - \mathbf{B}\| = \|\mathbf{USV}^T - \mathbf{US}'\mathbf{V}^T\| = \|\mathbf{U}(\mathbf{S} - \mathbf{S}')\mathbf{V}^T\| = \|\mathbf{S} - \mathbf{S}'\| = (s_{k+1}^2 + \dots + s_p^2)^{1/2}. \quad (7.23)$$

V tomto smyslu singulární čísla udávají vzdálenost matice k nejbližší matice dané nižší hodnosti. Všimněte si dále, že podle Cvičení 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 změnily, kdyby se matice nahradila nejbližší maticí dané nižší hodnosti.

7.5 Cvičení

- 7.1. Řekli jsme, že řešení úlohy (7.2) se nabývá pro $y_1 = 1$ a $y_2 = \dots = y_n = 0$ (kde předpokládáme, λ_1 je nejmenší z čísel $\lambda_1, \dots, \lambda_n$). Dokažte to přesně.
- 7.2. Vyřešte úlohu $\max\{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\}$. Musí být matice \mathbf{A} symetrická?
- 7.3. Pro čtvercovou matici \mathbf{A} je výraz $f(\mathbf{x}) = \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$ znám jako **Rayleighův kvocient**. Dokažte, že $\min_{\mathbf{x} \in \mathbb{R}^n \setminus \{0\}} f(\mathbf{x}) = \min\{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\}$.
- 7.4. 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.

Potom spočítejte ortogonální projekce bodů na podprostor X a souřadnice těchto projekcí v ortonormální bázi podprostoru X (uved'te i tuto ortonormální bázi). Použijte (a) spektrální rozklad, (b) SVD. Můžete použít počítač.

- 7.5. (★) Dokažte, že levé singulární vektory v rozkladu (7.20) tvoří ortonormální bázi prostoru $\text{rng } \mathbf{A}$ a pravé singulární vektory tvoří ortonormální bázi prostoru $\text{rng}(\mathbf{A}^T)$.
- 7.6. (★) V Matlabu se báze nulového prostoru matice najde funkcí `null`; vypište si implementaci této funkce matlabským příkazem `edit null` a najděte souvislost se Cvičením 7.5.
- 7.7. Je dán spektrální rozklad dané symetrické matice. Jak byste z něho jednoduše našli SVD této matice?
- 7.8. Pro matici $\mathbf{A} = \frac{1}{15} \begin{bmatrix} -13 & 2 & -22 \\ -16 & 14 & -4 \end{bmatrix}$ najděte matici \mathbf{B} hodnosti jedna takovou, že $\|\mathbf{A} - \mathbf{B}\|$ je minimální. Pak spočítejte $\|\mathbf{A} - \mathbf{B}\|$. Spočítejte pomocí (a) spektrálního rozkladu, (b) SVD. Použijte počítač.
- 7.9. Jsou dány matice

$$\mathbf{A} = \frac{1}{15} \begin{bmatrix} -13 & 2 & -22 \\ -16 & 14 & -4 \end{bmatrix}, \quad \mathbf{U} = \frac{1}{5} \begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix}, \quad \mathbf{V} = \frac{1}{3} \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}.$$

Bez počítače najděte matici \mathbf{B} hodnosti jedna takovou, že $\|\mathbf{A} - \mathbf{B}\|$ je minimální. Pak spočítejte co nejjednodušeji $\|\mathbf{A} - \mathbf{B}\|$.

- 7.10. Máme $n = 100$ bodů $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^{10^6}$. Chceme najít podprostor dimenze k (kde $k < 100$) minimalizující součet čtverců vzdáleností k bodům. Jak to uděláme co nejefektivněji? Náповěda: použijte výsledek Cvičení 6.5.
- 7.11. Dokažte, že následující dvě úlohy jsou stejné:
- Najdi podprostor takový, že součet čtverců vzdáleností daných bodů k tomuto podprostoru je minimální.
 - Najdi podprostor takový, že součet čtverců délek ortogonálních projekcí daných bodů na tento podprostor je maximální.
- 7.12. Zjednodušte Příklad 7.2 tak, že místo kuželosečky chceme dané body proložit přímkou $\{(x, y) \in \mathbb{R}^2 \mid ax + by + c = 0\}$. Děláme to opět tak, že minimalizujeme $\sum_i (ax_i + by_i + c)^2$ za podmínky $a^2 + b^2 + c^2 = 1$.
- Minimalizuje formulovaná úloha součet čtverců vzdáleností bodů k přímce, nebo jen jeho aproximaci (jako pro kuželosečku)?
 - Co když místo omezení $a^2 + b^2 + c^2 = 1$ použijeme omezení $a^2 + b^2 = 1$? Co úloha vyjadřuje a jak ji vyřešíme?
 - (*) Když v Příkladě 7.2 změníme omezení (7.19) na $a^2 + b^2 + c^2 + d^2 + e^2 = 1$ příp. na $a^2 + b^2 + c^2 = 1$, bude úloha minimalizovat přesný součet čtverců vzdáleností bodů ke kuželosečce?
- 7.13. (*) Dokažte, že $\max\{\mathbf{x}^T \mathbf{A} \mathbf{y} \mid \mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = \mathbf{y}^T \mathbf{y} = 1\} = s_1$, kde s_1 je největší singulární číslo matice $\mathbf{A} \in \mathbb{R}^{m \times n}$, a optimální argument je $\mathbf{u}_1, \mathbf{v}_1$, tedy levý a pravý singulární vektor příslušný singulárnímu číslu s_1 .
- 7.14. Do této chvíle jsme potkali již tři rozklady matic: QR, spektrální rozklad a SVD. Je ještě několik jiných užitečných rozkladů. Návrh algoritmů na operace s maticemi, řešení soustav lineárních rovnic a rozklady matic je předmětem *numerické lineární algebry*. Cílem je nalézt rychlé algoritmy, které jsou odolné vůči zaokrouhlovacím chybám. Existují volně dostupné softwarové balíky na numerickou lineární algebru. Matlab je postaven na balíku LAPACK. Jiný oblíbený balík je BLAS, který obsahuje funkce nižší úrovně než LAPACK. Najděte na internetu dokumentaci k balíkům LAPACK a BLAS a pochopte z ní co nejvíce!

Náповěda a řešení

- 7.1. Dokažte, že kdyby bylo $y_1 < 1$ (a tedy $y_2^2 + \dots + y_n^2 > 0$), tak bychom mohli hodnotu výrazu $\lambda_1 y_1^2 + \dots + \lambda_n y_n^2$ zmenšit bez porušení podmínky $y_1^2 + \dots + y_n^2 = 1$.
- 7.3. Klíčové je si všimnout, že $f(\alpha \mathbf{x}) = f(\mathbf{x})$ pro každé $\alpha \neq 0$ a $\mathbf{x} \neq \mathbf{0}$. To nám dovolí dokázat rovnost množin $\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}\} = \{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\}$. Každý prvek z pravé množiny zřejmě patří také do levé množiny. Ale také každý prvek z levé množiny patří do pravé množiny, protože když vektor \mathbf{x} vydělíme číslem $\|\mathbf{x}\|$, bude $\mathbf{x}^T \mathbf{x} = 1$ a hodnota $f(\mathbf{x})$ se nezmění.
- 7.4.a) $X = \text{span}\{(0.6912, -0.2181, 0.6889)\}$ (Uvědomte si ale, že báze vašeho podprostoru může být jiná než zde uvedená.)
- 7.4.b) $X = \text{span}\{(0.6912, -0.2181, 0.6889), (-0.3771, -0.9221, 0.0864)\}$
- 7.4.c) $X = (1.25, -1.25, 0.25) + \text{span}\{(0.6599, 0.0280, 0.7508)\}$
- 7.5. Chceme tedy dokázat, že $\text{rng } \mathbf{U} = \text{rng } \mathbf{A}$ a $\text{rng } \mathbf{V} = \text{rng}(\mathbf{A}^T)$. První rovnost plyne z Věty 3.6, neboť \mathbf{V}^T a \mathbf{S} mají l.n. řádky a proto $\text{rng } \mathbf{A} = \text{rng}(\mathbf{U} \mathbf{S} \mathbf{V}^T) = \text{rng}(\mathbf{U} \mathbf{S}) = \text{rng } \mathbf{U}$. Druhá rovnost je první rovnost použitá na matici \mathbf{A}^T .

- 7.9. Zadání napovídá, že by \mathbf{U}, \mathbf{V} mohly být součástí SVD matice \mathbf{A} . Pokud ano, bylo by $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ neboli $\mathbf{U}^T\mathbf{A}\mathbf{V} = \mathbf{S}$. Spočítáme $\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \end{bmatrix}$, tedy to je téměř pravda, jen musíme otočit znaménko druhého sloupce \mathbf{U} (šlo by i \mathbf{V}) a záporného prvku $s_{22} = -2$. Pak vynulujeme nejmenší diagonální prvek \mathbf{S} (tedy s_{11}) a máme $\mathbf{B} = \mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{A} = \frac{2}{15} \begin{bmatrix} -8 & 4 & -8 \\ -6 & 3 & -6 \end{bmatrix}$. Číslo $\|\mathbf{A} - \mathbf{B}\|$ je rovno vynulovanému singulárnímu číslu, tedy 1 (ověřte).
- 7.13. Dosad'te $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ (\mathbf{U}, \mathbf{V} ortogonální, $\mathbf{S} \in \mathbb{R}^{m \times n}$). Řešíme $\max\{\bar{\mathbf{x}}^T\mathbf{S}\bar{\mathbf{y}} \mid \bar{\mathbf{x}}^T\bar{\mathbf{x}} = \bar{\mathbf{y}}^T\bar{\mathbf{y}} = 1\}$. Máme $\bar{\mathbf{x}}^T\mathbf{S}\bar{\mathbf{y}} = \sum_{i=1}^p s_i \bar{x}_i \bar{y}_i = \sum_{i=1}^p s_i z_i$, kde $z_i = \bar{x}_i \bar{y}_i$. Platí $|z_i| \leq 1$ a $|\sum_i z_i| \leq 1$ (rozmyslete). Úloha $\max\{\sum_i s_i z_i \mid |z_i| \leq 1, |\sum_i z_i| \leq 1\}$ má optimální hodnotu s_1 a optimální argument je $z_1 = 1$ a $z_i = 0$ pro $i \neq 1$. Taková čísla $z_i = \bar{x}_i \bar{y}_i$ lze ale realizovat volbou $\bar{\mathbf{x}} = \mathbf{e}_1$ a $\bar{\mathbf{y}} = \mathbf{e}_1$. Optimální argument původní úlohy je $\mathbf{x} = \mathbf{U}\bar{\mathbf{x}} = \mathbf{u}_1$ a $\mathbf{y} = \mathbf{V}\bar{\mathbf{y}} = \mathbf{v}_1$.

Část II

Nelineární optimalizace

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 vždy bude očividné, jak by se látka zobecnila pro jiný definiční obor.

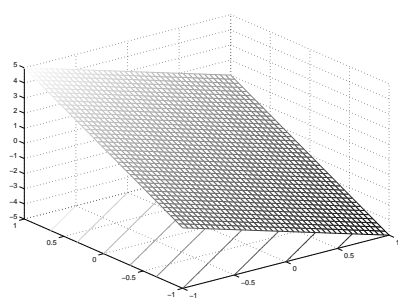
Pro funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ užíváme tyto pojmy:

- **Graf** funkce f je množina $\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid f(\mathbf{x}) = y\}$.
- **Vrstevnice** funkce f výšky y je množina $\{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) = y\}$.

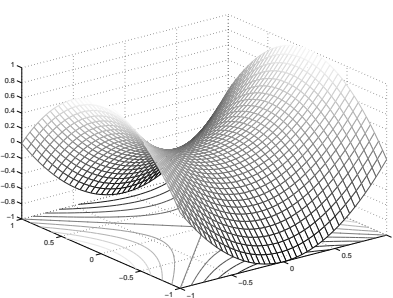
Příklad 8.1. Graf funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ dané vzorcem $f(x) = x^2$ je množina $\{(x, y) \mid x^2 = y\}$. Vrstevnice funkce f výšky 1 je množina $\{x \mid x^2 = 1\} = \{-1, 1\}$.

Graf funkce $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ dané vzorcem $f(x, y) = x^2 + y^2$ je množina $\{(x, y, z) \mid x^2 + y^2 = z\}$. Vrstevnice funkce f výšky 1 je množina $\{(x, y) \mid x^2 + y^2 = 1\}$. \square

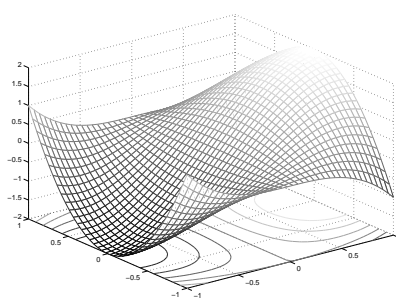
Příklad 8.2. Obrázek ukazuje část (na obdélníku $[-1, 1]^2$) grafu a vrtevnic funkcí $f: \mathbb{R}^2 \rightarrow \mathbb{R}$:



$$f(x, y) = -2x + 3y$$



$$f(x, y) = x^2 - y^2$$



$$f(x, y) = 3x - x^3 - 3xy^2$$

\square

Příklad 8.3. Příklady funkcí a zobrazení reálných proměnných:

1. $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2$
2. $f: \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = x^2 - y^2$
3. $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = a$ kde $a \in \mathbb{R}$ (konstantní funkce n proměnných)

4. $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(\mathbf{x}) = x_1$ (i když x_2, \dots, x_n chybí, f se rozumí jako funkce n proměnných)
5. $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} = a_1 x_1 + \dots + a_n x_n$ kde $\mathbf{a} \in \mathbb{R}^n$ (lineární funkce)
6. $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b = a_1 x_1 + \dots + a_n x_n + b$ kde $\mathbf{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$ (afinní funkce)
7. $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(\mathbf{x}) = e^{-\mathbf{x}^T \mathbf{x}}$ (nenormalizované Gaussovo rozdělení)
8. $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(\mathbf{x}) = \max_{i=1}^n x_i$
9. $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^2$, $\mathbf{f}(t) = (\cos t, \sin t)$ (parametrizace kružnice, množina $\mathbf{f}([0, 2\pi))$ je kružnice)
10. $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^3$, $\mathbf{f}(t) = (\cos t, \sin t, at)$ (parametrizace šroubovice neboli helixu)
11. $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^n$, $\mathbf{f}(t) = \mathbf{x} + t\mathbf{v}$ (parametrizace přímky jdoucí bodem \mathbf{x} ve směru \mathbf{v})
12. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\mathbf{f}(\mathbf{x}) = \mathbf{x}$ (identické zobrazení neboli identita)
13. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{f}(\mathbf{x}) = \mathbf{a}$ kde $\mathbf{a} \in \mathbb{R}^m$ (konstantní zobrazení)
14. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ kde $\mathbf{A} \in \mathbb{R}^{m \times n}$
(lineární zobrazení; je to také parametrizace lineárního podprostoru $\mathbf{f}(\mathbb{R}^n) = \text{rng } \mathbf{A} \subseteq \mathbb{R}^m$)
15. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ kde $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$
(afinní zobrazení; je to také parametrizace afinního podprostoru $\mathbf{f}(\mathbb{R}^n) \subseteq \mathbb{R}^m$)
16. $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $\mathbf{f}(u, v) = ((R + r \cos v) \cos u, (R + r \cos v) \sin u, r \sin v)$
(parametrizace toru neboli anuloidu, množina $\mathbf{f}([0, 2\pi) \times [0, 2\pi))$ je torus)
17. *Skalární pole* (pojem z fyziky) je funkce $\mathbb{R}^3 \rightarrow \mathbb{R}$. Např. teplotní pole přiřadí každému bodu prostoru teplotu.
18. *Vektorové pole* je zobrazení $\mathbb{R}^3 \rightarrow \mathbb{R}^3$. Např. elektrické pole přiřadí každému bodu v \mathbb{R}^3 vektor z \mathbb{R}^3 .
19. Při technice *image morphing* se obrázek např. obličeje zdeformuje na obrázek jiného obličeje. Morphing je realizován zobrazením $\mathbb{R}^2 \rightarrow \mathbb{R}^2$. □

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 fg jsou spojité v bodě \mathbf{x} . Pokud $g(\mathbf{x}) \neq 0$, je funkce f/g spojitá v bodě \mathbf{x} .
3. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je spojitá v bodě \mathbf{x} a $g: \mathbb{R} \rightarrow \mathbb{R}$ je spojitá v bodě $y = f(\mathbf{x})$. Pak složená funkce $g \circ f: \mathbb{R}^n \rightarrow \mathbb{R}$ 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 8.4. Z věty snadno ukážeme, že např. funkce dvou proměnných

$$f(x, y) = \sin(x + y^2) \quad (8.1)$$

je spojitá. Podle 1 je x spojitá funkce dvou proměnných (x, y) . Podobně, y^2 je spojitá funkce 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. \square

8.2 Derivace funkce jedné proměnné

Zopakujme pojem derivace funkce jedné proměnné $f: \mathbb{R} \rightarrow \mathbb{R}$. Existuje-li pro $x \in \mathbb{R}$ limita

$$\lim_{y \rightarrow x} \frac{f(y) - f(x)}{y - x} = \lim_{\alpha \rightarrow 0} \frac{f(x + \alpha) - f(x)}{\alpha}, \quad (8.2)$$

funkce f se nazývá *diferencovatelná* v bodě x a hodnota limity se nazývá její *derivace* funkce f v bodě x . Derivaci v bodě x značíme jedním ze symbolů

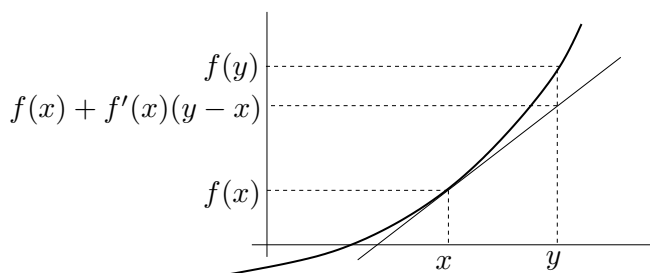
$$\frac{df(x)}{dx} = f'(x).$$

Pokud je funkce f je diferencovatelná pro všechna $x \in \mathbb{R}$, derivace v proměnném bodě x je funkce, kterou značíme $\frac{df}{dx}: \mathbb{R} \rightarrow \mathbb{R}$ nebo $f': \mathbb{R} \rightarrow \mathbb{R}$.

Pokud je funkce f v bodě x diferencovatelná, lze ji v blízkosti bodu x 'dobře' aproximovat afinní funkcí $\hat{f}: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(y) \approx \hat{f}(y) = f(x) + f'(x)(y - x) \quad (8.3)$$

(zde tedy x je pevné a y je proměnné)¹. Funkce \hat{f} je opravdu afinní, protože ji lze psát jako $\hat{f}(y) = ay + b$, kde $a = f'(x)$ a $b = f(x) - f'(x)x$. Viz obrázek:



8.3 Parciální derivace

Parciální derivaci funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ podle x_i značíme jedním ze symbolů

$$\frac{\partial f(\mathbf{x})}{\partial x_i} = f_{x_i}(\mathbf{x}).$$

¹Někdo raději píše $f(x + \alpha) \approx f(x) + f'(x)\alpha$, což je samozřejmě stejné jako (8.3) dosazením $y = x + \alpha$.

Spočítáme ji tak, že všechny proměnné x_j , $j \neq i$, považujeme za konstanty a zderivujeme funkci podle jediné proměnné x_i . Parciální derivace v proměnném bodě $\mathbf{x} \in \mathbb{R}^n$ je funkce, kterou značíme $\frac{\partial f}{\partial x_i}: \mathbb{R}^n \rightarrow \mathbb{R}$ nebo $f_{x_i}: \mathbb{R}^n \rightarrow \mathbb{R}$.

Příklad 8.5. Parciální derivace funkce (8.1) jsou

$$\frac{\partial f(x, y)}{\partial x} = f_x(x, y) = \cos(x + y^2), \quad (8.4a)$$

$$\frac{\partial f(x, y)}{\partial y} = f_y(x, y) = 2y \cos(x + y^2). \quad (8.4b)$$

□

8.4 Derivace zobrazení

Jak se dá pojem derivace zobecnit z funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ na zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$? Mohli bychom si myslet, že toho lze dosáhnout nějakým zobecněním limity (8.2), tudy ale cesta nevede. Je lépe vycházet ze vzorce (8.3): zkusme zobrazení v blízkosti bodu \mathbf{x} aproximovat jako

$$\mathbf{f}(\mathbf{y}) \approx \hat{\mathbf{f}}(\mathbf{y}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad (8.5)$$

kde symbol $\mathbf{f}'(\mathbf{x})$ označuje zatím neznámou matici rozměru $m \times n$. Zobrazení $\hat{\mathbf{f}}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je afinní (viz (3.5)), neboť lze psát $\hat{\mathbf{f}}(\mathbf{y}) = \mathbf{A}\mathbf{y} + \mathbf{b}$ kde $\mathbf{A} = \mathbf{f}'(\mathbf{x})$ a $\mathbf{b} = \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{x}$. Zobrazení \mathbf{f} 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 $\mathbf{f}(\mathbf{y}) - \hat{\mathbf{f}}(\mathbf{y})$ 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: zobrazení \mathbf{f} 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 funkce $\partial f_i/\partial x_j$ jsou v tomto bodě spojitě.

Věta 8.2. Je-li zobrazení v bodě **spojitě diferencovatelné**, je v tomto bodě **diferencovatelné**.

Příklad 8.6. Obě parciální derivace (8.4) funkce (8.1) jsou spojitě 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 . □

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 proměnných²:

$$\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}} = \mathbf{f}'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n}. \quad (8.6)$$

Matice (8.6) se nazývá **totální derivace**³ (nebo krátce jen **derivace**) zobrazení \mathbf{f} v bodě \mathbf{x} . Z historických důvodů se jí říká také **Jacobiho matice**. Totální derivace zobrazení \mathbf{f} v proměnném vektoru $\mathbf{x} \in \mathbb{R}^n$ je zobrazení, které značíme $\frac{d\mathbf{f}}{d\mathbf{x}}: \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$ nebo $\mathbf{f}': \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$.

Pamatujte speciální případy:

²Někdo matici (8.6) značí také $D\mathbf{f}(\mathbf{x})$, ale my to tak značit nebudeme.

³Někdy 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.

- Pro $f: \mathbb{R} \rightarrow \mathbb{R}$ je $f'(x)$ skalár a splývá s obyčejnou derivací (8.2).

- Pro $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^m$ je $\mathbf{f}'(x) = \begin{bmatrix} f'_1(x) \\ \vdots \\ f'_m(x) \end{bmatrix}$ sloupkový vektor, jehož složky jsou derivace složek \mathbf{f} .

- Pro $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je $f'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \end{bmatrix}$ řádkový vektor.

8.4.1 Derivace složeného zobrazení

Známé ‘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)

$$\mathbb{R}^n \xrightarrow{\mathbf{f}} \mathbb{R}^m \xrightarrow{\mathbf{g}} \mathbb{R}^l.$$

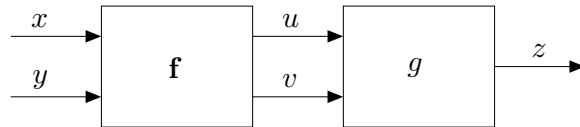
Pokud píšeme $\mathbf{u} = \mathbf{f}(\mathbf{x})$ a $\mathbf{y} = \mathbf{g}(\mathbf{u})$, pravidlo lze psát také v Leibnizově značení jako

$$\frac{d\mathbf{y}}{d\mathbf{x}} = \frac{d\mathbf{y}}{d\mathbf{u}} \frac{d\mathbf{u}}{d\mathbf{x}}, \quad (8.8)$$

což se dobře pamatuje, protože $d\mathbf{u}$ se ‘jakoby vykrátí’ (což ale není důkaz!). Zdůrazněme, že tato rovnost je *násobení 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 8.7. Necht' $g(u, v)$ je diferencovatelná funkce dvou proměnných. Hledejme (totální) derivaci funkce $h(x, y) = g(x + y, xy)$ podle vektoru (x, y) , tedy matici $\begin{bmatrix} \frac{\partial h(x, y)}{\partial x} & \frac{\partial h(x, y)}{\partial y} \end{bmatrix} \in \mathbb{R}^{1 \times 2}$.

Máme $\mathbb{R}^2 \xrightarrow{\mathbf{f}} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}$, kde $\mathbf{f}(x, y) = (u, v) = (x + y, xy)$. Označme $z = g(u, v)$. Viz obrázek:



Derivace zobrazení g podle vektoru (u, v) je matice 1×2 (řádkový vektor)

$$g'(u, v) = \begin{bmatrix} \frac{\partial g(u, v)}{\partial u} & \frac{\partial g(u, v)}{\partial v} \end{bmatrix} = \begin{bmatrix} g_u(u, v) & g_v(u, v) \end{bmatrix}.$$

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 8.8. 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 8.7 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 8.9. Máme diferencovatelnou funkci $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ a chceme spočítat derivaci funkce $g(t + t^2, \sin t)$ podle t .

Máme $\mathbb{R} \xrightarrow{\mathbf{f}} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}$, kde $\mathbf{f}(t) = (u, v) = (t + t^2, \sin t)$. Je

$$\frac{dg(t + t^2, \sin t)}{dt} = g'(u, v) \mathbf{f}'(t) = [g_u(u, v) \quad g_v(u, v)] \begin{bmatrix} 1 + 2t \\ \cos t \end{bmatrix} = g_u(u, v)(1 + 2t) + g_v(u, v) \cos t. \quad \square$$

8.4.2 Derivace maticových výrazů

Je-li zobrazení zadáno výrazem obsahujícím vektory a matice, jeho (totální) derivaci lze často napsat jako výraz obsahující tyto vektory a matice.

Příklad 8.10. Odvodme 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

$$\begin{aligned} f_1(\mathbf{x}) &= a_{11}x_1 + \cdots + a_{1n}x_n + b_1 \\ &\vdots \\ f_m(\mathbf{x}) &= a_{m1}x_1 + \cdots + a_{mn}x_n + b_m. \end{aligned}$$

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 zobrazení. Opravdu: pravá strana výrazu (8.5) je pro libovolné \mathbf{x} rovna $\mathbf{A}\mathbf{x} + \mathbf{b}$. □

Příklad 8.11. Počítejme derivaci zobrazení $\mathbf{g}(\mathbf{Ax} + \mathbf{b})$ 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 $\mathbf{f}(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$. Je $\mathbf{f}'(\mathbf{x}) = \mathbf{A}$, tedy

$$\frac{d\mathbf{g}(\mathbf{Ax} + \mathbf{b})}{d\mathbf{x}} = \mathbf{g}'(\mathbf{Ax} + \mathbf{b})\mathbf{A}. \quad \square$$

Příklad 8.12. Odvodme derivaci kvadratické formy $f: \mathbb{R}^n \rightarrow \mathbb{R}$,

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{Ax} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} x_i x_j, \quad (8.9)$$

kde $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n \times n}$ je (ne nutně symetrická) čtvercová matice. Derivace monomu je

$$\frac{\partial(x_i x_j)}{\partial x_k} = \begin{cases} x_j & \text{když } k = i, \\ x_i & \text{když } k = j, \\ 0 & \text{jinak.} \end{cases}$$

Tedy

$$\frac{\partial f(\mathbf{x})}{\partial x_k} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} \frac{\partial(x_i x_j)}{\partial x_k} = \sum_{i=1}^n (a_{ik} + a_{ki}) x_i = (\mathbf{a}_k + \mathbf{b}_k)^T \mathbf{x},$$

kde $\mathbf{a}_k = (a_{1k}, \dots, a_{nk})$ je k -tý sloupec a $\mathbf{b}_k^T = (a_{k1}, \dots, a_{kn})^T$ je k -tý řádek matice \mathbf{A} . Tedy

$$\mathbf{f}'(\mathbf{x}) = \left[\frac{\partial f(\mathbf{x})}{\partial x_1} \quad \dots \quad \frac{\partial f(\mathbf{x})}{\partial x_n} \right] = \mathbf{x}^T (\mathbf{A} + \mathbf{A}^T). \quad (8.10) \quad \square$$

Následující tabulka uvádí derivace jednoduchých zobrazení (odvodte je jako cvičení!), které je dobré si pamatovat. S pomocí zapamatovaných derivací jednoduchých zobrazení a řetězového pravidla nakonec totální derivace odvozujeme s trochou cviku rovnou v maticovém tvaru.

zobrazení	(totální) derivace	poznámka
$\mathbf{f}(\mathbf{x}) = \mathbf{x}$	$\mathbf{f}'(\mathbf{x}) = \mathbf{I}$	$\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$
$\mathbf{f}(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$	$\mathbf{f}'(\mathbf{x}) = \mathbf{A}$	$\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$
$f(\mathbf{x}) = \mathbf{a}^T \mathbf{x}$	$f'(\mathbf{x}) = \mathbf{a}^T$	$f: \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$
$f(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$	$f'(\mathbf{x}) = 2\mathbf{x}^T$	$f: \mathbb{R}^n \rightarrow \mathbb{R}$
$f(\mathbf{x}) = \mathbf{x}^T \mathbf{Ax}$	$f'(\mathbf{x}) = \mathbf{x}^T (\mathbf{A} + \mathbf{A}^T)$	$f: \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{A} \in \mathbb{R}^{n \times n}$
$f(\mathbf{x}) = \ \mathbf{x}\ $	$f'(\mathbf{x}) = \mathbf{x}^T / \ \mathbf{x}\ $	$f: \mathbb{R}^n \rightarrow \mathbb{R}$
$f(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T \mathbf{g}(\mathbf{x})$	$f'(\mathbf{x}) = 2\mathbf{g}(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$	$f: \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$
$f(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T \mathbf{h}(\mathbf{x})$	$f'(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T \mathbf{h}'(\mathbf{x}) + \mathbf{h}(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$	$f: \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{h}: \mathbb{R}^n \rightarrow \mathbb{R}^m$

8.5 Směrová derivace

Řez zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ v bodě $\mathbf{x} \in \mathbb{R}^n$ ve směru⁴ $\mathbf{v} \in \mathbb{R}^n$ je zobrazení $\varphi: \mathbb{R} \rightarrow \mathbb{R}^m$ dané jako

$$\varphi(\alpha) = \mathbf{f}(\mathbf{x} + \alpha \mathbf{v}). \quad (8.11)$$

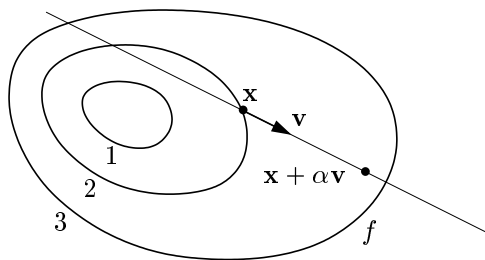
⁴Někdy se řez a směrová derivace uvažují jen pro *normalizované* směry \mathbf{v} , tj. $\|\mathbf{v}\| = 1$. My ale dovolujeme libovolný vektor \mathbf{v} .

Směrová derivace⁵ zobrazení \mathbf{f} v bodě \mathbf{x} ve směru \mathbf{v} je vektor

$$(\varphi'_1(0), \dots, \varphi'_m(0)) = \lim_{\alpha \rightarrow 0} \frac{\varphi(\alpha) - \varphi(0)}{\alpha} = \lim_{\alpha \rightarrow 0} \frac{\mathbf{f}(\mathbf{x} + \alpha \mathbf{v}) - \mathbf{f}(\mathbf{x})}{\alpha}, \quad (8.12)$$

kde $\varphi'_1(0)$ označuje derivaci i -té složky zobrazení φ v bodě $\alpha = 0$. Směrová derivace samozřejmě existuje jen tehdy, když derivace $\varphi'_1(0), \dots, \varphi'_m(0)$ existují.

Pojem směrové derivace se geometricky snadněji představí pro případ $m = 1$, tedy pro funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Obrázek ilustruje situaci pro funkci $f: \mathbb{R}^2 \rightarrow \mathbb{R}$:



Vidíme, že 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 zobrazení $\mathbf{y} = \mathbf{f}(\mathbf{u})$ a $\mathbf{u} = g(\alpha) = \mathbf{x} + \alpha \mathbf{v}$. Máme diagram $\mathbb{R} \xrightarrow{g(\alpha)=\mathbf{u}} \mathbb{R}^n \xrightarrow{\mathbf{f}(\mathbf{u})=\mathbf{y}} \mathbb{R}^m$. Je $g'(\alpha) = \mathbf{v}$. Podle řetězového pravidla je

$$\varphi'(\alpha) = \mathbf{f}'(g(\alpha))g'(\alpha) = \mathbf{f}'(g(\alpha))\mathbf{v}.$$

Ale $g(0) = \mathbf{x}$ a tedy $\varphi'(0) = \mathbf{f}'(\mathbf{x})\mathbf{v}$. □

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 8.13. Spočítáme směrovou derivaci funkce $f(x, y) = \sin(x + y^2)$ v bodě (x, y) ve směru (u, v) dvěma způsoby: nejdříve z definice (8.12) a pak z Věty 8.4. Z definice je

$$\begin{aligned} \varphi(\alpha) &= \sin(x + \alpha u + (y + \alpha v)^2), \\ \varphi'(\alpha) &= (u + 2v(y + \alpha v)) \cos(x + \alpha u + (y + \alpha v)^2), \\ \varphi'(0) &= (u + 2vy) \cos(x + y^2). \end{aligned}$$

Podle Věty 8.4 je směrová derivace rovna

$$uf_x(x, y) + vf_y(x, y) = u \cos(x + y^2) + 2vy \cos(x + y^2) = (u + 2vy) \cos(x + y^2),$$

což je stejné jako předešlý výsledek. □

⁵Přesněji jde o *oboustrannou* směrovou derivaci. Jednostrannou směrovou derivaci bychom dostali, kdybychom místo oboustranné limity (8.12) vzali jednostrannou limitu zprava.

8.6 Gradient

Sloupcový vektor parciálních derivací funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ se nazývá její **gradient** a značí se

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix} = f'(\mathbf{x})^T$$

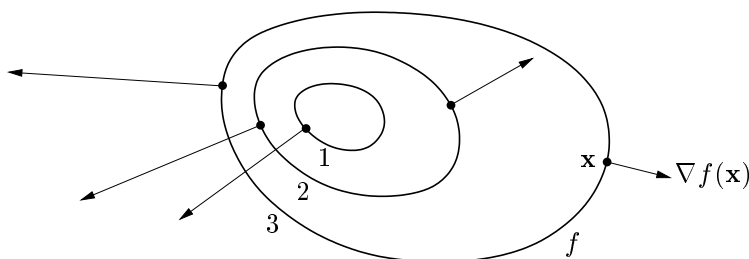
(∇ čteme ‘nabla’). Je to tedy transpozice Jacobiho matice $f'(\mathbf{x})$, což je řádkový vektor⁶.

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 §10.1), 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:



8.7 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. Jestliže druhé parciální derivace

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}, \quad \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i}$$

v bodě \mathbf{x} existují a jsou v tomto bodě spojité, pak jsou si rovny.

⁶Zavedení 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 $f: \mathbb{R}^n \rightarrow \mathbb{R}$. 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.

Příklad 8.14. Určeme všechny druhé parciální derivace funkce $f(x, y) = \sin(x + y^2)$ z Příkladu 8.4. 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

$$f''(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_n} \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Dle Věty 8.5 je tato matice symetrická. Často se jí říká **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 $m \times n \times n$.

8.8 Taylorův polynom

Necht' 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 aproximací funkce f v okolí bodu x .

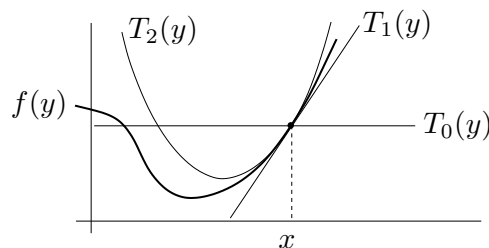
Taylorův polynom je tímto požadavkem definován jednoznačně a má tvar (odvoďte!)

$$T_k(y) = \sum_{i=0}^k \frac{1}{i!} f^{(i)}(x) (y - x)^i, \quad (8.13)$$

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 je roven funkci \hat{f} v (8.3). 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} společné všechny parciální derivace až do řádu k . Nebudeme uvádět vzorec pro polynom libovolného stupně, napíšeme jen polynomy do stupně dva:

$$T_0(\mathbf{y}) = f(\mathbf{x}), \quad (8.14a)$$

$$T_1(\mathbf{y}) = f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad (8.14b)$$

$$T_2(\mathbf{y}) = f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T f''(\mathbf{x})(\mathbf{y} - \mathbf{x}). \quad (8.14c)$$

Zde $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $f'(\mathbf{x}) \in \mathbb{R}^{1 \times n}$ je Jacobiho matice (řádkový vektor) a $f''(\mathbf{x}) \in \mathbb{R}^{n \times n}$ je Hessova matice. Funkce (8.14b) je afinní a funkce (8.14c) 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

$$\mathbf{T}_1(\mathbf{y}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad (8.15)$$

což je ale zobrazení $\hat{\mathbf{f}}$ ve (8.5). Polynom druhého stupně vede na zobrazení \mathbf{T}_2 , jehož složky jsou funkce (8.14c). To nejde napsat v maticové formě, protože všech $m \times n \times n$ druhých parciálních derivací se ‘nevejde’ do matice.

Příklad 8.15. Najděte Taylorův polynom druhého stupně funkce $f(x, y) = x^{-1} + y^{-1} + xy$ v bodě $(x_0, y_0) = (2, 1)$. Máme

$$\begin{aligned} f(x_0, y_0) &= x^{-1} + y^{-1} + xy|_{(x,y)=(2,1)} = \frac{7}{2}, \\ f'(x_0, y_0) &= [y - x^{-2} \quad x - y^{-2}]|_{(x,y)=(2,1)} = \left[\frac{3}{4} \quad 1\right], \\ f''(x_0, y_0) &= \begin{bmatrix} 2x^{-3} & 1 \\ 1 & 2y^{-3} \end{bmatrix}|_{(x,y)=(2,1)} = \begin{bmatrix} \frac{1}{4} & 1 \\ 1 & 2 \end{bmatrix}. \end{aligned}$$

Dle (8.14c) je (pozor, naše proměnné jsou označené jinak než v (8.14c))

$$\begin{aligned} T_2(x, y) &= f(x_0, y_0) + f'(x_0, y_0) \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}^T f''(x_0, y_0) \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\ &= \frac{7}{2} + \left[\frac{3}{4} \quad 1\right] \begin{bmatrix} x - 2 \\ y - 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x - 2 \\ y - 1 \end{bmatrix}^T \begin{bmatrix} \frac{1}{4} & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x - 2 \\ y - 1 \end{bmatrix} \\ &= \frac{1}{8}x^2 + xy + y^2 - \frac{3}{4}x - 3y + \frac{9}{2}. \end{aligned}$$

□

8.9 Vnitřek a hranice množiny

Označme jako⁷

$$B_\varepsilon(\mathbf{x}) = \{ \mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{y}\| \leq \varepsilon \} \quad (8.16)$$

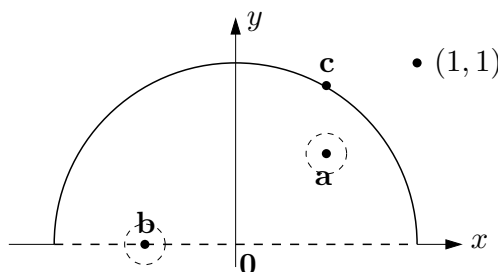
n -rozměrnou kouli se středem $\mathbf{x} \in \mathbb{R}^n$ a poloměrem $\varepsilon \geq 0$. Všimněte si, že pro speciální případ $n = 1$ (tedy $x \in \mathbb{R}$) je množina (8.16) interval $B_\varepsilon(x) = [x - \varepsilon, x + \varepsilon]$.

Mějme nyní množinu $X \subseteq \mathbb{R}^n$. Bod $\mathbf{x} \in \mathbb{R}^n$ se nazývá její

- **vnitřní bod**, jestliže existuje $\varepsilon > 0$ takové, že $B_\varepsilon(\mathbf{x}) \subseteq X$,
- **hraniční bod**, jestliže pro každé $\varepsilon > 0$ je $B_\varepsilon(\mathbf{x}) \cap X \neq \emptyset$ a $B_\varepsilon(\mathbf{x}) \cap (\mathbb{R}^n \setminus X) \neq \emptyset$.

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 8.16. 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 koule (s nenulovým poloměrem!) se středem **a**, která celé leží v množině. Bod **b** je hraniční, protože každá koule se středem **b** 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 $(1, 1)$ (který patří do množiny, viz její definice výše) je hraniční. \square

Příklad 8.17. Bod $1/2$ je vnitřní bod intervalu $(0, 1] \subseteq \mathbb{R}$ a body 0 a 1 jsou hraniční. \square

Příklad 8.18. 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

Příklad 8.19. Kružnice v rovině $\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \}$ nemá žádné vnitřní body, všechny její body jsou hraniční. Podobně pro n -rozměrnou sféru $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1 \}$. \square

Příklad 8.20. Mějme kruh bez hranice $\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1 \}$. Všechny jeho body jsou vnitřní. Jeho hranice je kružnice $\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \}$. Podobně pro n -rozměrnou kouli bez hranice $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| < 1 \}$. \square

⁷Norma v (8.16) může být eukleidovská, ale i libovolná jiná vektorová p -norma (viz §11.5.1). Vnitřek a hranice množiny na výběru normy nezávisí.

8.10 Lokální extrémy funkce na množině

Zopakujme (§1.1.4), že funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ nabývá na množině $X \subseteq \mathbb{R}^n$ v bodě $\mathbf{x} \in X$ svého

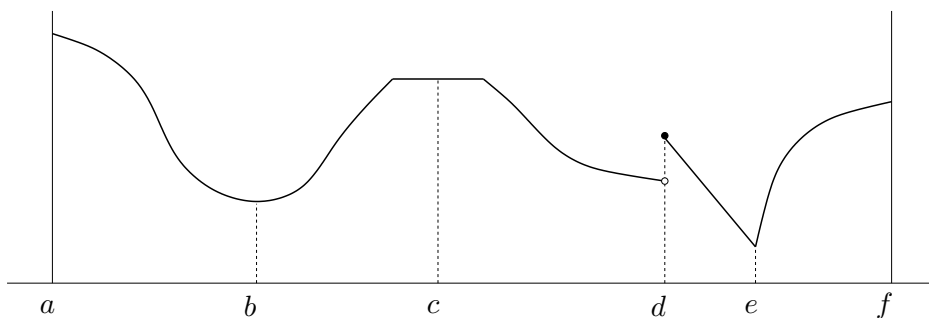
- *minima*, jestliže $f(\mathbf{x}) \leq f(\mathbf{x}')$ pro všechna $\mathbf{x}' \in X$,
- *ostrého minima*, jestliže $f(\mathbf{x}) < f(\mathbf{x}')$ pro všechna $\mathbf{x}' \in X \setminus \{\mathbf{x}\}$.

Funkce f nabývá na množině X v bodě \mathbf{x} svého **[ostrého] lokálního minima**, jestliže existuje $\varepsilon > 0$ tak, že funkce f nabývá na množině $B_\varepsilon(\mathbf{x})$ svého [ostrého] minima. Maximum, lokální maximum a jejich ostré verze se definují obdobně.

Každé minimum funkce f je zároveň lokální minimum funkce f (proč?), naopak to ale obecně neplatí. Mluvíme-li o lokálních extrémech, pro zdůraznění někdy ‘obyčejné’ extrémy nazýváme **globální extrémy**. Pokud odkaz na množinu X chybí, myslí se jí celý definiční obor funkce f .

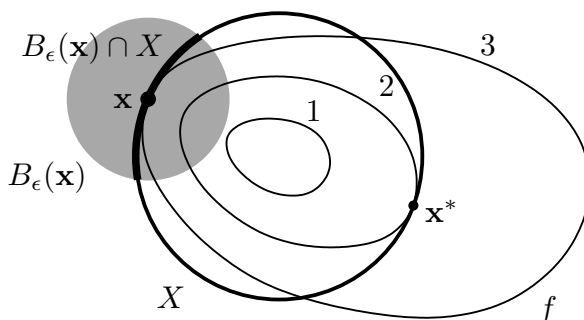
Každý bod množiny $X \subseteq \mathbb{R}^n$ je buď vnitřní nebo hraniční. Extrém \mathbf{x} (globální či lokální) funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině X je **volný** když \mathbf{x} je vnitřní bod množiny X , a **vázaný** když \mathbf{x} je hraniční bod množiny X .

Příklad 8.21. Funkce jedné proměnné na obrázku nabývá na uzavřeném intervalu $[a, f] \subseteq \mathbb{R}$ v bodě a globálního (a tedy i lokálního) maxima, v bodě b lokálního minima, v bodě c lokálního maxima a zároveň lokálního minima, v bodě d lokálního maxima, v bodě e globálního (a tedy i lokálního) minima, v bodě f lokálního maxima. Extrémy v bodech a, b, d, e, f jsou ostré. Extrémy v bodech a, f jsou vázané, v bodech b, c, d, e jsou volné.



□

Příklad 8.22. 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 $\varepsilon > 0$ tak, že funkce f nabývá na části kružnice $B_\varepsilon(\mathbf{x}) \cap X$ svého (globálního) minima. Oba extrémy jsou vázané, protože body \mathbf{x}^* a \mathbf{x} jsou hraniční body množiny X (množina X vnitřní body nemá, viz Příklad 8.19). □

Příklad 8.23. Funkce $f(\mathbf{x}) = x_1$ má na množině $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq 1\}$ vázané globální (a tedy i lokální) minimum v bodě $(-1, 0, \dots, 0)$. \square

Příklad 8.24. Libovolná funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ má na množině \mathbb{Z} (množina celých čísel) v libovolném bodě $x \in \mathbb{Z}$ lokální minimum i lokální maximum. \square

8.11 Cvičení

8.1. Najděte parametrizaci válce. Přesněji, najděte zobrazení $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ tak, aby množina $\mathbf{f}([0, 2\pi) \times \mathbb{R})$ byla nekonečný válec o poloměru r bez podstav.

8.2. 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:

- a) $\{\mathbf{x} \in \mathbb{R}^2 \mid 1 \geq \min_{\mathbf{y} \in X} \|\mathbf{x} - \mathbf{y}\|\}$
- b) $\{\mathbf{x} \in \mathbb{R}^2 \mid 2 \geq \max_{\mathbf{y} \in X} \|\mathbf{x} - \mathbf{y}\|\}$
- c) vrstevnice výšky 1 funkce $f(\mathbf{x}) = \min_{\mathbf{y} \in Y} \|\mathbf{x} - \mathbf{y}\|$
- d) vrstevnice výšky $\sqrt{2}$ funkce $f(\mathbf{x}) = \max_{\mathbf{y} \in Y} \|\mathbf{x} - \mathbf{y}\|$

/hiddenTohle cvičení změnit na nakreslení vrstevnic několika funkcí. Studenti neumí nakreslit vrstevnice lineární funkce!

8.3. Máme funkci $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ danou vzorcem $f(x, y) = \ln(1 + xy)$. Máme bod $(x_0, y_0) = (1, 2)$.

- a) Je funkce f v bodě (x_0, y_0) spojitá?
- b) Je funkce f v bodě (x_0, y_0) spojitě diferencovatelná?
- c) Je funkce f v bodě (x_0, y_0) diferencovatelná?
- d) Najdi totální derivaci (Jacobiho matici) $f'(x, y)$ funkce f v bodě (x_0, y_0) .
- e) Najdi gradient $\nabla f(x, y)$ funkce f v bodě (x_0, y_0) .
- f) Najdi řez a směrovou derivaci funkce f v bodě (x_0, y_0) ve směru $(1, -1)$.
- g) Najdi Hessovu matici funkce f v bodě (x_0, y_0) .

8.4. Funkce $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ je dána vzorcem $f(x, y) = \max\{x, y\}$. Ve kterých bodech je tato funkce spojitě diferencovatelná? Odpověď odůvodněte.

8.5. Najdi totální derivaci (Jacobiho matici) zobrazení 1, 2, 4, 6, 9, 12, 14, 15 z Příkladu 8.3.

8.6. Necht' $f(x, y)$ je diferencovatelná funkce dvou proměnných.

- a) Spočítej derivaci f podle polárních souřadnic (φ, r) , kde $x = r \cos \varphi$, $y = r \sin \varphi$.
- b) 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.

8.7. Necht' $f(x, y)$ je diferencovatelná funkce dvou proměnných. Spočítej derivaci funkce $g(\mathbf{u}) = f(\mathbf{a}^T \mathbf{u}, \|\mathbf{u}\|)$ podle vektoru \mathbf{u} .

8.8. Odvod' co nejjednodušší vzorec pro totální derivaci (Jacobiho matici) těchto zobrazení. Kde je to možné, odvod' nejdříve přímým výpočtem a pak řetězovým pravidlem.

- a) $\mathbf{f}(\mathbf{x}) = \mathbf{x}^T \mathbf{x} = \|\mathbf{x}\|^2$
- b) $f(\mathbf{x}) = \|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}}$
- c) $f(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T \mathbf{h}(\mathbf{x})$ (kde $\mathbf{g}, \mathbf{h}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ jsou dána)

d) $f(\mathbf{x}) = \|\mathbf{g}(\mathbf{x})\|$

- 8.9. Odvod' totální derivaci (8.10) kvadratické formy, nepoužij ale postup z Příkladu 8.12 ale použij vzorec $(\mathbf{g}(\mathbf{x})^T \mathbf{h}(\mathbf{x}))' = \mathbf{g}(\mathbf{x})^T \mathbf{h}'(\mathbf{x}) + \mathbf{h}(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$.
- 8.10. 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.

- 8.11. Spočítejte druhou derivaci $f''(\mathbf{x})$ (Hessovu matici) těchto funkcí:

- $f(x, y) = e^{-x^2-y^2}$
- $f(x, y) = \log(e^x + e^y)$
- $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ (matice \mathbf{A} je dána, nemusí být symetrická)
- $f(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T \mathbf{g}(\mathbf{x})$ (zobrazení $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je dáno)
- $f(\mathbf{x}) = \|\mathbf{x}\|$

- 8.12. 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.

- 8.13. *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 $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ a $\mathbf{g}: \mathbb{R}^m \rightarrow \mathbb{R}^l$ pro nějaké navzájem různé dimenze $n, m, l > 1$. Zvolte bod $\mathbf{x} \in \mathbb{R}^n$. Spočítejte přibližně totální derivace (Jacobiho matice) $\mathbf{f}'(\mathbf{x})$ a $\mathbf{g}'(\mathbf{f}(\mathbf{x}))$ v Matlabu metodou konečných diferencí. Potom spočítejte derivaci složeného zobrazení $(\mathbf{g} \circ \mathbf{f})'(\mathbf{x})$ jednak metodou konečných diferencí a jednak vynásobením matic $\mathbf{f}'(\mathbf{x})$ a $\mathbf{g}'(\mathbf{f}(\mathbf{x}))$. Porovnejte.

- 8.14. (★) Necht' zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je definováno výrazem $\mathbf{f}(\mathbf{x})$ obsahujícím vektor \mathbf{x} , konstantní matice a vektory, a operace součet matic, maticový součin a transpozice (např. $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{x}\mathbf{b}^T \mathbf{x}$). Lze vždy derivaci $\mathbf{f}'(\mathbf{x})$ vyjádřit výrazem obsahujícím vektor \mathbf{x} , ty samé konstantní matice a vektory, a operace součet matic, maticový součin a transpozice?

- 8.15. Načrtněte následující podmnožiny \mathbb{R}^2 (proměnné x, y patří do \mathbb{R}):

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

- 8.16. Co je vnitřek a hranice těchto množin?

- a) Množina reálných čísel \mathbb{R}
- b) Uzavřený interval $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$
- c) Množina racionálních čísel \mathbb{Q}
- d) Množina (8.16)
- e) $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1, y \geq 0\}$
- f) $\{(x, y) \in \mathbb{R}^2 \mid y = x^2, -1 < x \leq 1\}$
- g) $\{(x, y) \in \mathbb{R}^2 \mid xy < 1, x > 0, y > 0\}$
- h) $\{\mathbf{x} \in \mathbb{R}^n \mid \max_{i=1}^n x_i \leq 1\}$
- i) $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}$, kde $\mathbf{a} \in \mathbb{R}^n, b \in \mathbb{R}$ (nadrovina)
- j) $\{\mathbf{x} \in \mathbb{R}^n \mid b \leq \mathbf{a}^T \mathbf{x} \leq c\}$, kde $\mathbf{a} \in \mathbb{R}^n, b, c \in \mathbb{R}$ (panel)
- k) $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}\}$ (afinní podprostor \mathbb{R}^n)

8.17. Každá z následujících množin je sjednocením konečného počtu (otevřených, uzavřených či polouzavřených) intervalů. Najděte tyto intervaly. Příklad: $\{x^2 \mid x \in \mathbb{R}\} = [0, +\infty)$.

- a) $\{1/x \mid x \geq 1\}$
- b) $\{1/x \mid |x| \geq 1\}$
- c) $\{e^{-x^2} \mid x \in \mathbb{R}\}$
- d) $\{x + y \mid x^2 + y^2 < 1\}$
- e) $\{x + y \mid x^2 + y^2 = 1\}$
- f) $\{x - y \mid x^2 + y^2 = 1\}$
- g) $\{|x| + |y| \mid x^2 + y^2 = 1\}$
- h) $\{x_1 + \dots + x_n \mid \mathbf{x} \in \mathbb{R}^n, x_1^2 + \dots + x_n^2 = 1\}$
- i) $\{|x - y| \mid x \in [0, 1], y \in (1, 2]\}$
- j) $\{x + y \mid |x| \geq 1, |y| \geq 1\}$

8.18. Najdi (úvahou, bez použití derivací) všechny extrémy funkce

- a) $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x}$ (vektor $\mathbf{a} \in \mathbb{R}^n$ je dán)
- b) $f(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$

na množině

- a) \mathbb{R}^n
- b) $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$
- c) $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq 1\}$
- d) $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| < 1\}$
- e) daný lineární podprostor prostoru \mathbb{R}^n
- f) daný afinní podprostor prostoru \mathbb{R}^n
- g) $\{\mathbf{x} \in \mathbb{R}^n \mid -1 \leq \mathbf{1}^T \mathbf{x} \leq 1\}$

U každého extrému urči, zda je lokální/globální, ostrý/neostrý, volný/vázaný.

8.19. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a $\mathbf{x} \in X' \subseteq X \subseteq \mathbb{R}^n$. Uvažujme dva výroky:

1. Funkce f má v bodě \mathbf{x} lokální minimum na množině X .
2. Funkce f má v bodě \mathbf{x} lokální minimum na množině X' .

Vyplývá druhý výrok z prvního? Vyplývá první výrok z druhého? Odpovědi dokažte.

- 8.20. Může se stát, že funkce má na množině lokální minimum ale nemá na ní globální minimum? Odpověď dokažte.

Nápověda a řešení

- 8.3.a) Ano, dle Věty 8.1.
- 8.3.b) Ano, protože její parciální derivace $\partial f(x, y)/\partial x = y/(1 + xy)$ a $\partial f(x, y)/\partial y = x/(1 + xy)$ jsou (dle Věty 8.1) spojitě funkce.
- 8.3.c) Ano, dle Věty 8.2.
- 8.3.d) $f'(x, y) = [y/(1 + xy) \quad x/(1 + xy)] \in \mathbb{R}^{1 \times 2}$, tedy $f'(x_0, y_0) = [2 \quad 1]/3$.
- 8.3.e) $\nabla f(x, y) = [f'(x, y)]^T = \begin{bmatrix} y/(1 + xy) \\ x/(1 + xy) \end{bmatrix} \in \mathbb{R}^{2 \times 1}$, tedy $\nabla f(x_0, y_0) = (2, 1)/3 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}/3$.
- 8.3.f) Řez je $\varphi(\alpha) = \ln[1 + (1 + \alpha)(2 - \alpha)]$. Směrová derivace je $1/3$.
- 8.4. Je spoj. diferencovatelná na množině $\mathbb{R}^2 \setminus \{(x, y) \mid x = y\}$.
- 8.6.a) $f_\varphi(x, y) = -f_x(x, y)r \sin \varphi + f_y(x, y)r \cos \varphi$, $f_r(x, y) = f_x(x, y) \cos \varphi + f_y(x, y) \sin \varphi$
- 8.6.b) $f_t(x, y) = 2f_x(x, y)(t + 1) + 2tf_y(x, y)/(t^2 + 1)$
- 8.9. Dosad' $\mathbf{g}(\mathbf{x}) = \mathbf{x}$ a $\mathbf{h}(\mathbf{x}) = \mathbf{Ax}$.
- 8.10.a) $(5, 1)/\sqrt{26}$
- 8.10.b) $(5, 1)^T(1, -1)/\sqrt{2} = 2\sqrt{2}$
- 8.11.a) $2e^{-x^2-y^2} \begin{bmatrix} 2x^2 - 1 & 2xy \\ 2xy & 2y^2 - 1 \end{bmatrix}$
- 8.11.b) $\frac{e^{x+y}}{(e^x + e^y)^2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$
- 8.11.c) $f''(\mathbf{x}) = \mathbf{A} + \mathbf{A}^T$
- 8.11.d) $f''(\mathbf{x}) = 2\mathbf{g}'(\mathbf{x})^T \mathbf{g}'(\mathbf{x}) + 2 \sum_{i=1}^m g_i(\mathbf{x}) g_i''(\mathbf{x})$ (výraz nejde napsat čistě v maticové formě, nutno použít sumu)
- 8.12. $T_0(x, y) = 46$, $T_1(x, y) = 18x - 60y - 92$, $T_2(x, y) = -6x^2 - 24xy - 18x + 24y^2 + 60y + 46$
- 8.16.c) Vnitřek je \emptyset a hranice je \mathbb{R} .
- 8.19. Druhý výrok plyne z prvního. Naopak to ale neplatí, protipříklad je $X = \mathbb{R}$, $X' = [0, 1]$, $f(x) = x$.

Kapitola 9

Volné lokální extrémy

9.1 Analytické podmínky

9.1.1 Pro jednu proměnnou

Zopakujme nejdříve z analýzy potřebné pojmy pro funkce jedné proměnné.

Funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ je **rostoucí** [**klesající**] na intervalu $[a, b] \subseteq \mathbb{R}$, když pro každé $x, y \in [a, b]$ platí, že $x < y$ implikuje $f(x) < f(y)$ [$f(x) > f(y)$].

Věta 9.1. *Necht' je funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ v bodě $x \in \mathbb{R}$ diferencovatelná a necht' $f'(x) > 0$ [$f'(x) < 0$]. Pak existuje $\delta > 0$ tak, že na intervalu $[x - \delta, x + \delta] = B_\delta(x)$ je f rostoucí [klesající].*

Důkaz. Necht' $f'(x) = a$. Z definice limity funkce (zopakujte z analýzy!) a definice derivace (8.2) plyne, že pro každé $\varepsilon > 0$ existuje $\delta > 0$ tak, že

$$y \in [x - \delta, x + \delta] \implies \frac{f(y) - f(x)}{y - x} \in [a - \varepsilon, a + \varepsilon].$$

Zvolme $\varepsilon = \frac{1}{2}a$, tedy $[a - \varepsilon, a + \varepsilon] = [\frac{1}{2}a, \frac{3}{2}a]$. Necht' $y - x > 0$. Když $a > 0$ [$a < 0$], interval $[\frac{1}{2}a, \frac{3}{2}a]$ obsahuje jen kladná [záporná] čísla a proto $f(y) - f(x) > 0$ [$f(y) - f(x) < 0$]. \square

Důsledek 9.2 (Fermatova věta). *Necht' funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ je v bodě $x \in \mathbb{R}$ diferencovatelná a má v tomto bodě lokální extrém. Pak $f'(x) = 0$.*

Jestliže $f'(x) = 0$, bodu x říkáme **stacionární bod** funkce f . Věta 9.2 říká, že stacionární body jsou body 'podezřelé' z volného lokálního extrému. Věta samozřejmě neříká, že každý stacionární bod je lokální extrém (uvažte např. funkci $f(x) = x^3$ v bodě $x = 0$).

Věta 9.2 udává podmínku *prvního řádu* na volné extrémy, protože obsahuje první derivace. Následující podmínka *druhého řádu* (uvedená bez důkazu) pomůže zjistit, zda je stacionární bod lokálním extrémem, případně jakým.

Věta 9.3. *Necht' funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ je v bodě $x \in \mathbb{R}$ dvakrát diferencovatelná a necht' $f'(x) = 0$.*

- *Je-li x lokální minimum [maximum] funkce f , pak $f''(x) \geq 0$ [$f''(x) \leq 0$].*
- *Je-li $f''(x) > 0$ [$f''(x) < 0$], pak x je ostré lokální minimum [maximum] funkce f .*

Všimněte si, že je-li $f''(x) = 0$, funkce v bodě x lokální extrém mít může nebo nemusí (příkladem jsou funkce $f(x) = x^3$ a $f(x) = x^4$ v bodě $x = 0$).

9.1.2 Pro více proměnných

Věta 9.2 má snadný důsledek pro funkce více proměnných:

Věta 9.4. *Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a $X \subseteq \mathbb{R}^n$. Necht' \mathbf{x} je vnitřní bod množiny X a zároveň lokální extrém funkce f na množině X . Necht' $\mathbf{v} \in \mathbb{R}^n$. Jestliže existuje směrová derivace funkce f v bodě \mathbf{x} ve směru \mathbf{v} , pak je nulová.*

Důkaz. Jelikož \mathbf{x} je lokální extrém f na X a zároveň vnitřní bod X , existuje $\varepsilon > 0$ tak, že $B_\varepsilon(\mathbf{x}) \subseteq X$ a \mathbf{x} je (globální) extrém funkce f na kouli $B_\varepsilon(\mathbf{x})$. Pro řez $\varphi(\alpha) = f(\mathbf{x} + \alpha\mathbf{v})$ (viz §8.5) z toho plyne, že bod 0 je (globální) extrém funkce φ na intervalu $[-\varepsilon/\|\mathbf{v}\|, \varepsilon/\|\mathbf{v}\|] = B_{\varepsilon/\|\mathbf{v}\|}(0)$. Tedy funkce φ má v bodě $\alpha = 0$ lokální extrém. Dle Věty 9.2 je tedy $\varphi'(0) = 0$. \square

Jestliže je funkce v bodě \mathbf{x} diferencovatelná, tak existují směrové derivace ve všech směrech a jsou nulové, tedy speciálně všechny parciální derivace jsou nulové neboli $f'(\mathbf{x}) = \mathbf{0}$. V tom případě bod \mathbf{x} opět nazýváme **stacionární bod** funkce f .

Věta 9.4 svádí k tomu, aby se použila v situacích, kdy nejsou splněny její předpoklady. Uved'me příklady tohoto chybého použití (jeden už jsme uvedli v §9.1.1 pro $n = 1$).

Příklad 9.1. V Příkladu 8.21 jsou předpoklady Věty 9.4 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 9.2. 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 X a vrstevnice funkce f pro $n = 1$ a pro $n = 2$!). Nemá tam ale stacionární bod, protože tam není diferencovatelná. Dále má funkce na množině X vázaná lokální maxima ve všech jejích rozích (např. v bodě $\mathbf{1}$), což jsou její hraniční body. Bod $\mathbf{1}$ ovšem není stacionární bod funkce f . \square

A nyní podmínky druhého řádu:

Věta 9.5. *Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a $X \subseteq \mathbb{R}^n$. Necht' \mathbf{x} je vnitřní bod množiny X . Necht' je funkce f v bodě \mathbf{x} dvakrát diferencovatelná a platí $f'(\mathbf{x}) = \mathbf{0}$. Pak:*

- *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 .*

Bod \mathbf{x} , ve kterém $f'(\mathbf{x}) = \mathbf{0}$ a matice $f''(\mathbf{x})$ je indefinitní, se nazývá **sedlový bod**. Všimněte si, že je-li matice $f''(\mathbf{x})$ (pozitivně či negativně) semidefinitní, věta neříká nic o tom, zda funkce v bodě \mathbf{x} má nebo nemá lokální extrém. Příklad jsme uvedli v §9.1.1 pro $n = 1$.

Větu 9.5 nebudeme dokazovat, uvedeme jen náznak důkazu. Místo funkce f vyšetřujme v blízkosti bodu \mathbf{x} její Taylorův polynom druhého stupně (8.14c),

$$T_2(\mathbf{y}) = f(\mathbf{x}) + \underbrace{f'(\mathbf{x})(\mathbf{y} - \mathbf{x})}_0 + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T f''(\mathbf{x})(\mathbf{y} - \mathbf{x}).$$

Protože $f'(\mathbf{x}) = \mathbf{0}$, lineární člen je nulový a polynom je tedy kvadratická forma posunutá do bodu \mathbf{x} . Zda funkce T_2 má či nemá v bodě \mathbf{x} extrém bychom tedy mohli dokázat určit z podle Věty 6.2 z definitnosti matice $f''(\mathbf{x})$. Zde ovšem vyšetřujeme funkci f a ne její aproximaci T_2 , proto pro lokální extrém nestačí (pozitivní či negativní) semidefinitnost $f''(\mathbf{x})$.

Příklad 9.3. Extrémy kvadratické funkce (6.12) 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 (6.14a). 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

9.2 Sestupné metody

Dále se budeme věnovat numerickým iteračním algoritmům na nalezení volného lokálního minima diferencovatelných funkcí na množině \mathbb{R}^n .

Iterační algoritmy na hledání lokálního minima spojitě funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ mají tvar

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{v}_k, \quad (9.1)$$

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, $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$.

Necht' je funkce f diferencovatelná. Směr \mathbf{v}_k se nazývá **sestupný** v bodě \mathbf{x}_k , jestliže

$$f'(\mathbf{x}_k) \mathbf{v}_k < 0, \quad (9.2)$$

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

$$\varphi(\alpha_k) = f(\mathbf{x}_k + \alpha_k \mathbf{v}_k) \quad (9.3)$$

přes všechny $\alpha_k \geq 0$. Tato úloha je v kontextu vícerozměrné optimalizace nazývána *jednorozměrné hledání* (angl. *line search*). Tuto úlohu stačí řešit přibližně. Takovou přibližnou metodu není obtížné vymyslet a proto se jí dále nebudeme zabývat.

Dále uvedeme nejznámější zástupce sestupných metod.

9.3 Gradientní metoda

Tato nejjednodušší metoda volí směr sestupu jako záporný gradient funkce f v bodě \mathbf{x}_k :

$$\mathbf{v}_k = -f'(\mathbf{x}_k)^T = -\nabla f(\mathbf{x}_k). \quad (9.4)$$

Tento směr je sestupný, což je okamžitě vidět dosazením do (9.2).

Výhodou gradientní metody je spolehlivost, daná tím, že směr (9.4) je vždy sestupný. Nevýhodou gradientní metody je často pomalá konvergence. 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 $f''(\mathbf{x})$ jsou velmi různá).

¹Existují 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*).

Příklad 9.4. Hledejme minimum kvadratické formy $f(x, y) = (ax^2 + y^2)/2$ (kde $a > 0$) gradientní metodou s počátečním bodem $(x_0, y_0) = (1, a)$. Minimum se nabývá v bodě $(x, y) = (0, 0)$. Při přesném řešení problému (9.3) je k -tá iterace rovna (odvod'te!)

$$x_k = \left(-\frac{a-1}{a+1}\right)^k, \quad y_k = a \left(\frac{a-1}{a+1}\right)^k. \quad (9.5)$$

Vidíme, že konvergence je velmi pomalá pro $a \ll 1$ nebo $a \gg 1$. □

9.3.1 (★) 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

$$\tilde{\mathbf{x}}_{k+1} = \tilde{\mathbf{x}}_k - \alpha_k \tilde{f}'(\tilde{\mathbf{x}}_k)^T. \quad (9.6)$$

Zkoumejme, jaké iteraci to odpovídá v původních proměnných. K tomu potřebujeme vyjádřit (9.6) 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 (9.6) a úpravou dostaneme

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k (\mathbf{A}^T \mathbf{A})^{-1} f'(\mathbf{x}_k)^T. \quad (9.7)$$

To lze napsat ve tvaru (9.1) se směrem hledání

$$\mathbf{v}_k = -(\mathbf{A}^T \mathbf{A})^{-1} f'(\mathbf{x}_k)^T. \quad (9.8)$$

Tento směr se liší od původního směru (9.4) vynásobením maticí $(\mathbf{A}^T \mathbf{A})^{-1}$. Vidíme, že gradientní metoda *není invariantní* vůči lineární transformaci souřadnic.

Ovšem lze ukázat, že nový směr (9.8) je také sestupný. Dosazením (9.4) do (9.2) 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 6.19 a 6.18).

Na vzorec (9.8) 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.

9.4 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.

9.4.1 Použití na soustavy nelineárních rovnic

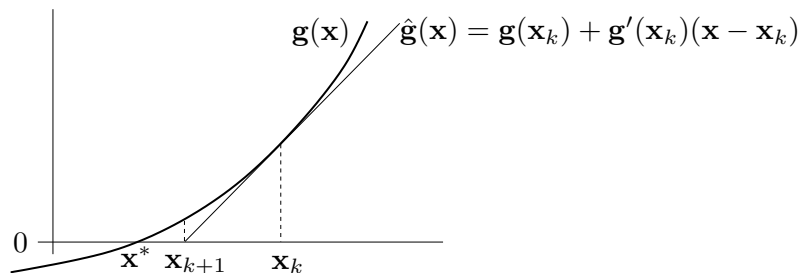
Řešme rovnici $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, kde $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ je diferencovatelné zobrazení, což je soustava n rovnic s n neznámými. Zobrazení \mathbf{g} aproximujeme v okolí bodu \mathbf{x}_k afinním zobrazením $\hat{\mathbf{g}}$ (viz (8.5))

$$\mathbf{g}(\mathbf{x}) \approx \hat{\mathbf{g}}(\mathbf{x}) = \mathbf{g}(\mathbf{x}_k) + \mathbf{g}'(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k). \quad (9.9)$$

Další iteraci \mathbf{x}_{k+1} najdeme řešením nehomogenní lineární soustavy $\hat{\mathbf{g}}(\mathbf{x}_{k+1}) = \mathbf{0}$. Pokud je Jacobiho matice $\mathbf{g}'(\mathbf{x}_k)$ regulární, řešením je

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{g}'(\mathbf{x}_k)^{-1} \mathbf{g}(\mathbf{x}_k). \quad (9.10)$$

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 9.5. *Babylónská metoda* na výpočet druhé odmocniny čísla $a \geq 0$ opakuje iteraci

$$x_{k+1} = \frac{1}{2} \left(x_k + \frac{a}{x_k} \right).$$

To není nic jiného než Newtonova metoda pro rovnici $0 = g(x) = x^2 - a$. 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 9.6. Hledeme průsečík $(x, y) \in \mathbb{R}^2$ dvou rovinných křivek daných rovnicemi

$$\begin{aligned} (x-1)^2 + y^2 &= 1, \\ x^4 + y^4 &= 1. \end{aligned}$$

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 (9.10) je

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} - \begin{bmatrix} 2(x_k-1) & 2y_k \\ 4x_k^3 & 4y_k^3 \end{bmatrix}^{-1} \begin{bmatrix} (x_k-1)^2 + y_k^2 - 1 \\ x_k^4 + y_k^4 - 1 \end{bmatrix}.$$

Načrtne-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

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 2 \\ 4 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 1 \end{bmatrix}.$$

Pro šestou iteraci $(x_6, y_6) = (0.671859751039018, 0.944629015546222)$ jsou rovnice splněny se strojovou přesností. \square

Příklad 9.7. 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 9.8. 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

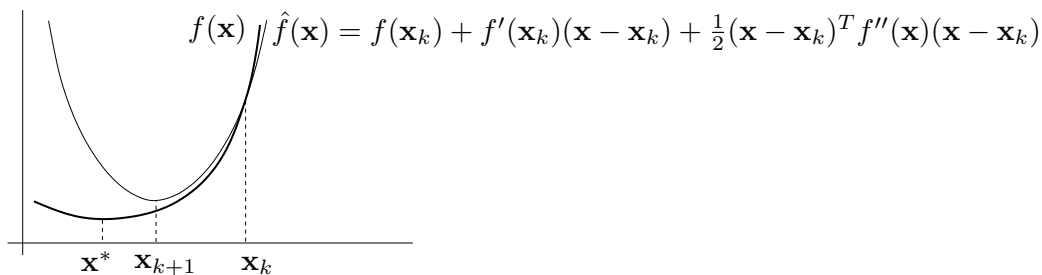
9.4.2 Použití na minimalizaci funkce

Newtonovu metodu lze použít pro hledání lokálního extrému dvakrát diferencovatelné funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ tak, že v algoritmu (9.10) položíme $\mathbf{g}(\mathbf{x}) = f'(\mathbf{x})^T$. Tím dostaneme iteraci

$$\mathbf{x}_{k+1} = \mathbf{x}_k - f''(\mathbf{x}_k)^{-1} f'(\mathbf{x}_k)^T, \quad (9.11)$$

kde $f''(\mathbf{x}_k)$ je Hessova matice funkce f v bodě \mathbf{x}_k .

Význam iterace (9.10) byl takový, že se zobrazení \mathbf{g} v bodě \mathbf{x}_k aproximovalo afinním zobrazením a pak se našel nulový bod tohoto zobrazení. Význam iterace (9.11) je takový, že se funkce f aproximuje Taylorovým polynomem druhého stupně (tedy kvadratickou funkcí) a pak se najde minimum této kvadratické funkce (viz Cvičení 9.8).



Iteraci (9.11) lze napsat v obecnějším tvaru (9.1), kde

$$\mathbf{v}_k = -f''(\mathbf{x}_k)^{-1} f'(\mathbf{x}_k)^T. \quad (9.12)$$

Výhodou tohoto zobecnění je možnost zvolit optimální (ne nutně jednotkovou) délku kroku pomocí jednorozměrné minimalizace (9.3). Algoritmu (9.11) s jednotkovou délkou kroku se pak říká **čistá** Newtonova metoda.

Vektoru (9.12) říkáme **Newtonův směr**. Vidíme, že se od gradientního směru (9.4) 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 §9.4.1 je zde matice $\mathbf{g}'(\mathbf{x}_k) = f''(\mathbf{x}_k)$ symetrická, což může řešení soustavy ulehčit.

9.5 Nelineární metoda nejmenších čtverců

Mějme soustavu rovnic $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, kde $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ (tedy je to soustava m rovnic s n neznámými). Soustavu nazveme přeuredenou, jestliže nemá žádné řešení. Chceme takovou přeuredenou soustavu řešit přibližně ve smyslu nejmenších čtverců. Tedy chceme minimalizovat funkci

$$f(\mathbf{x}) = \|\mathbf{g}(\mathbf{x})\|^2 = \mathbf{g}(\mathbf{x})^T \mathbf{g}(\mathbf{x}) = \sum_{i=1}^m g_i(\mathbf{x})^2, \quad (9.13)$$

kde g_i jsou složky zobrazení \mathbf{g} . Speciálním případem je přibližné řešení lineární nehomogenní soustavy $\mathbf{Ax} = \mathbf{b}$, kde $\mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{Ax}$ (viz §5.1).

Zatímco v §9.3 a §9.4.2 bylo cílem minimalizovat *obecnou* funkci, zde chceme minimalizovat funkci ve speciálním tvaru (9.13). Nyní máme dvě možnosti. Bud' můžeme nasadit na funkci (9.13) jednu z metod pro minimalizaci obecné funkce, k čemuž se vrátíme v §9.5.2. Nebo můžeme být chytřejší a využít speciálního tvaru funkce (9.13), což popíšeme v §9.5.1.

9.5.1 Gauss-Newtonova metoda

Aproximujme opět zobrazení \mathbf{g} afinním zobrazením (9.9). Úloha (9.13) 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) má tvar

$$\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k) = -\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k). \quad (9.14)$$

Její řešení napíšeme pomocí pseudoinverze:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{g}'(\mathbf{x}_k)^+ \mathbf{g}(\mathbf{x}_k). \quad (9.15)$$

Algoritmus (9.15) je znám jako (čistá) **Gauss-Newtonova metoda**. Můžeme jej opět napsat obecněji ve tvaru (9.1) se směrem hledání (Gauss-Newtonův směr)

$$\mathbf{v}_k = -\mathbf{g}'(\mathbf{x}_k)^+ \mathbf{g}(\mathbf{x}_k). \quad (9.16)$$

Všimněte si, že pokud $m = n$ a Jakobiho matice $\mathbf{g}'(\mathbf{x}_k)$ je regulární, je $\mathbf{g}'(\mathbf{x}_k)^+ = \mathbf{g}'(\mathbf{x}_k)^{-1}$, tedy Gauss-Newtonova metoda (9.15) se redukuje na Newtonovu metodu (9.10).

Pokud má Jacobiho matice $\mathbf{g}'(\mathbf{x}_k)$ lineárně nezávislé sloupce (viz §5.1), výraz (9.16) můžeme napsat jako

$$\mathbf{v}_k = -(\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} \mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k) = -\frac{1}{2}(\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} f'(\mathbf{x}_k)^T, \quad (9.17)$$

kde jsme dosadili derivaci $f'(\mathbf{x}) = 2\mathbf{g}(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$ účelové funkce (9.13) (viz §8.4.2). Vidíme, že Gauss-Newtonův směr (9.17) se liší od gradientního směru (9.4) 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

$$f'(\mathbf{x}_k) \mathbf{v}_k = -\frac{1}{2} f'(\mathbf{x}_k) (\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} f'(\mathbf{x}_k)^T < 0.$$

To 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 když $\mathbf{g}'(\mathbf{x}_k)$ má lineárně nezávislé sloupce (viz Cvičení 6.19), 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 (9.13). Protože ale Gauss-Newtonův směr je vždy sestupný, vhodnou volbou délky kroku α_k lze vždy zajistit konvergenci.

Příklad 9.9. Hledáme přibližné řešení soustavy tří rovnic o dvou neznámých

$$\begin{aligned}(x-1)^2 + y^2 &= 1, \\ x^4 + y^4 &= 1, \\ x^2 + (y-1)^2 &= 1/2.\end{aligned}$$

Oba průsečíky křivek daných prvními dvěma rovnicemi již známe z Příkladu 9.6. Ani jeden z těchto průsečíků neleží na třetí křivce (i když je jí blízko), tedy soustava je přeurčená. Nezbývá 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 (9.15) 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 9.10. 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 přeurčenou nelineární soustavu ve smyslu nejmenších čtverců, tedy minimalizujme 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 §8.4.2, 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 (9.15). \square

9.5.2 Rozdíl oproti Newtonově metodě

Předpokládejme, že bychom optimalizovali naši účelovou funkci (9.13) přímo Newtonovou metodou z §9.4.2. Spočítejme (viz Cvičení 8.11.d) Hessián funkce (9.13):

$$f''(\mathbf{x}) = 2\mathbf{g}'(\mathbf{x})^T \mathbf{g}'(\mathbf{x}) + 2 \sum_{i=1}^m g_i(\mathbf{x}) g''_i(\mathbf{x}). \quad (9.18)$$

Hessián je součtem členu obsahujícího derivace prvního řádu a členu obsahujícího derivace druhého řádu. Vidíme, že Gauss-Newtonův směr (9.17) se liší od Newtonova směru (9.12) zanedbáním členu druhého řádu v Hessiánu (9.18). Jinými slovy, Gauss-Newtonovu metodu je možno vnímat jako aproximaci Newtonovy metody na minimalizaci funkce (9.13) spočívající v tom, že zanedbáme členy druhého řádu, tedy skutečný Hessián (9.18) aproximujeme 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 (9.13). Ovšem vyhnuli jsme se počítání druhých derivací funkce \mathbf{g} , což je hlavní výhoda Gauss-Newtonovy metody.

9.5.3 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

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (\mathbf{g}'(\mathbf{x}_k)^T\mathbf{g}'(\mathbf{x}_k) + \mu_k\mathbf{I})^{-1}\mathbf{g}'(\mathbf{x}_k)^T\mathbf{g}(\mathbf{x}_k) \quad (9.19)$$

kde $\mu_k > 0$. Přidání členu $\mu_k\mathbf{I}$ je vlastně regularizace, viz §5.1.5. Potom:

- Pro malé μ_k se (9.19) 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 (9.19) 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 (9.15) 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 (9.19) je ale vždy pozitivně definitní (viz Cvičení 6.16), a tedy regulární.

9.6 Cvičení

9.1. 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

- $\{2, 3, -1\}$
- $\{2, 3, 0\}$
- $\{2, 1, 1\}$

9.2. Pro následující funkce najděte stacionární body. Pro každý stacionární bod určete, zda je to lokální minimum, lokální maximum, či ani jedno. Pokud to určit neumíte, odůvodněte.

- a) $f(x, y) = x(1 - \frac{2}{3}x^2 - y^2)$
- b) $f(x, y) = 1/x + 1/y + xy$
- c) $f(x, y) = e^y(y^2 - x^2)$
- d) $f(x, y) = 3x - x^3 - 3xy^2$
- e) $f(x, y) = 6xy^2 - 2x^3 - 3y^4$
- f) $f(x, y) = x^4/3 + y^4/2 - 4xy^2 + 2x^2 + 2y^2 + 3$
- g) $f(x, y, z) = x^3 + y^3 + 2xyz + z^2$

- 9.3. Najděte všechna řešení rovnice $\sin x = \frac{1}{2}x$ (sinus je v radiánech) na kalkulačce s největší přesností, jakou dokážete.
- 9.4. Najděte lokální extrém funkce $f(x, y) = x^2 - y + \sin(y^2 - 2x)$ čistou Newtonovou metodou. Počáteční odhad zvolte $(x_0, y_0) = (1, 1)$. Můžete použít počítač.
- 9.5. 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. Napište iteraci (a) Gauss-Newtonovy, (b) Levenberg-Marquardtovy metody.
- 9.6. Soustavu dvou rovnic o jedné neznámé

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

chceme řešit přibližně ve smyslu nejmenších čtverců. Napište iteraci

- a) čisté Gauss-Newtonovy metody,
- b) čisté Newtonovy metody.

Výsledné vzorce co nejvíce zjednodušte.

- 9.7. 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 (9.13). Napište iteraci (a) gradientní, (b) Newtonovy, (c) Gauss-Newtonovy, (d) Levenberg-Marquardtovy metody.

- 9.8. Čistá Newtonova metoda (9.11) na minimalizaci funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je Newtonova metoda (9.10) na řešení soustavy $f'(\mathbf{x})^T = \mathbf{0}$. Takto jsme ji odvodili. Ukažte, že iteraci (9.11) lze odvodit také tak, že funkci f aproximujeme okolo bodu \mathbf{x}_k Taylorovým polynomem druhého řádu a najdeme \mathbf{x}_{k+1} jako minimum tohoto polynomu.
- 9.9. (★) V §9.3.1 jsme ukázali, že iterace gradientní metody není invariantní vůči lineární transformaci souřadnic $\tilde{\mathbf{x}} = \mathbf{A}\mathbf{x}$ (pro regulární \mathbf{A}). Ukažte, že iterace Newtonovy metody (9.11) je invariantní vůči této transformaci.
- 9.10. Může Newtonova metoda (9.10) konvergovat, když soustava $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ nemá řešení?
- 9.11. Rovnice $x^2 + 1 = 0$ nemá řešení, řešme ji tedy přibližně Gauss-Newtonovou metodou. Protože $m = n$, čistá Gauss-Newtonova iterace (9.15) je vlastně Newtonova iterace (9.10). Bude Newtonova metoda (9.10) konvergovat? Víme, že Gauss-Newtonovův směr je sestupný, tedy Gauss-Newtonova metoda by konvergovat měla. Vysvětlete zdánlivý rozpor.

- 9.12. (★) Gauss-Newtonův směr (9.16) se získá řešením normální rovnice (9.14). Ukázali jsme, že když Jacobiho matice $\mathbf{g}'(\mathbf{x}_k)$ má lineárně nezávislé sloupce, je Gauss-Newtonův směr sestupný. Má-li $\mathbf{g}'(\mathbf{x}_k)$ lineárně závislé sloupce, rovnice (9.14) má nekonečně mnoho řešení, tj. směrů hledání je nekonečně mnoho. Dokažte, že každý takový směr je sestupný.

Nápověda a řešení

- 9.1.a) funkce nemá v tomto bodě lokální extrém
 9.1.b) nemůžeme rozhodnout, zda má funkce má v tomto bodě lokální extrém
 9.1.c) funkce má v tomto bodě lokální minimum
 9.2.d) Stacionární body jsou 4.
 9.2.e) Stacionární body jsou 3.
 9.2.f) Stacionárních bodů je 5.
 9.2.g) Stacionární body jsou $(0, 0, 0)$ a $(3/2, 3/2, -9/4)$.
 9.3. Jeden kořen je $x = 0$ a pak dva další liší se znaménkem. Jeden z nich získáme Newtonovou metodou: $x_{k+1} = x_k - (2 \sin x_k - x_k) / (2 \cos x_k - 1)$. Načtneme si grafy funkcí $\sin x$ a $1/\overline{2x}$ a z toho zvolíme počáteční odhad, např. $x_0 = 2$. Po několika iteracích máme $x_k = 1.895494267033981$.
 9.6.a) Minimalizujeme $\mathbf{g}(x)^T \mathbf{g}(x)$ kde $\mathbf{g}(x) = \begin{bmatrix} x^2 + x - 1 \\ x^2 - x - 1 \end{bmatrix}$. Máme $\mathbf{g}'(x) = \begin{bmatrix} 2x + 1 \\ 2x - 1 \end{bmatrix}$. Iterace je

$$x \leftarrow x - (\mathbf{g}'(x)^T \mathbf{g}'(x))^{-1} \mathbf{g}'(x)^T \mathbf{g}(x) = x - (8x^2 + 2)^{-1} \begin{bmatrix} 2x + 1 & 2x - 1 \end{bmatrix} \begin{bmatrix} x^2 + x - 1 \\ x^2 - x - 1 \end{bmatrix} = \frac{2x(x^2 + 1)}{4x^2 + 1}.$$

 9.6.b) Minimalizujeme $f(x) = \mathbf{g}(x)^T \mathbf{g}(x) = (x^2 + x - 1)^2 + (x^2 - x - 1)^2 = 2(x^4 - x^2 + 1)$. Máme $f'(x) = 8x^3 - 4x$, $f''(x) = 24x^2 - 4$. Iterace je $x \leftarrow x - f'(x)/f''(x) = (4x^3)/(6x^2 - 1)$. Vidíme, že pro tuto jednoduchou soustavu s jednou proměnnou není Newtonova iterace složitější než Gauss-Newtonova iterace – obvykle je to ale naopak.
 9.7. 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.
 9.8. Po přejmenování proměnných $\mathbf{x}_k, \mathbf{x}_{k+1}$ na \mathbf{x}, \mathbf{y} máme ukázat, že stacionární bod \mathbf{y} Taylorova polynomu (8.14c) je právě bod splňující $\mathbf{g}(\mathbf{x}) + \mathbf{g}'(\mathbf{x})(\mathbf{y} - \mathbf{x}) = \mathbf{0}$ kde $\mathbf{g}(\mathbf{x}) = f'(\mathbf{x})^T$. To se snadno dokáže výpočtem derivace $T'_2(\mathbf{y}) = f'(\mathbf{x}) + (\mathbf{y} - \mathbf{x})^T f''(\mathbf{x})$ (s užitím Cvičení 8.11).
 9.10. Nemůže, protože $\mathbf{g}'(\mathbf{x}_k)^{-1}$ (pokud existuje) je regulární a $\mathbf{g}(\mathbf{x}_k) \neq \mathbf{0}$, tedy v (9.10) je vždy $\mathbf{x}_{k+1} \neq \mathbf{x}_k$.
 9.11. Čistá (tj. s jednotkovými délkami kroku $\alpha_k = 1$) Gauss-Newtonova metoda konvergovat nebude. Ale pokud délky kroku α_k v Gauss-Newtonově metodě volíme tak, aby funkce $f(x) = g(x)^2 = (x^2 + 1)^2$ klesala, metoda bude konvergovat k minimu funkce f .
 9.12. Pišme normální rovnici (9.14) jako $\mathbf{A}^T \mathbf{A} \mathbf{v} = \mathbf{A}^T \mathbf{b}$ kde $\mathbf{A} = \mathbf{g}'(\mathbf{x}_k)$, $\mathbf{b} = -\mathbf{g}(\mathbf{x}_k)$ a $\mathbf{v} = \mathbf{x}_{k+1} - \mathbf{x}_k$ je směr hledání. Je-li \mathbf{v} řešení normální rovnice, pak $\mathbf{A} \mathbf{v} = \mathbf{P} \mathbf{b}$ je projekce vektoru \mathbf{b} na podprostor $\text{rng } \mathbf{A}$. Je $f'(\mathbf{x}_k) = -\mathbf{A}^T \mathbf{b}$, tedy podmínka na sestupnost směru \mathbf{v} zní $\mathbf{b}^T \mathbf{A} \mathbf{v} = \mathbf{b}^T \mathbf{P} \mathbf{b} > 0$. Ale \mathbf{P} je pozitivně semidefinitní (viz Cvičení 6.25), tedy $\mathbf{b}^T \mathbf{P} \mathbf{b} \geq 0$. Zde ovšem rovnost nastane jen když $\mathbf{b} = \mathbf{0}$ nebo $\mathbf{A}^T \mathbf{b} = \mathbf{0}$. Tedy \mathbf{v} je sestupný, za přirozených předpokladů $\mathbf{g}(\mathbf{x}_k) \neq \mathbf{0}$ a $f'(\mathbf{x}) \neq \mathbf{0}$.

Kapitola 10

Lokální extrémý vázané rovnostmi

Hledejme minimum funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) = \mathbf{0} \}, \quad (10.1)$$

kde $\mathbf{g} = (g_1, \dots, g_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$. To odpovídá úloze (1.9) s omezeními typu rovnosti:

$$\begin{aligned} \min \quad & f(x_1, \dots, x_n) \\ \text{za podmínek} \quad & g_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, m. \end{aligned} \quad (10.2)$$

Mluvíme o minimu funkce f *vázaném rovnostmi* $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

Množina X je množina řešení soustavy m rovnic (obecně nelineárních) o n neznámých. Tato množina obvykle nemá žádné vnitřní body, proto nelze použít Větu 9.4. V některých případech ale lze vyjádřit všechna řešení soustavy $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ parametricky a úlohu tak převést na úlohu bez omezení. To jsme použili v Příkladu 1.2, uveďme další příklady.

Příklad 10.1. Minimalizujme funkci $f(x, y) = x + y$ na množině $X = \{ (x, y) \in \mathbb{R}^2 \mid g(x, y) = x^2 + y^2 - 1 = 0 \}$, máme tedy $m = 1$ a $n = 2$. Tj. řešíme úlohu

$$\begin{aligned} \min \quad & x + y \\ \text{za podmínky} \quad & x^2 + y^2 = 1 \end{aligned} \quad (10.3)$$

Množinu X lze 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 10.2. 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 a \mathbf{C} jsou 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. Nevýhodou je nutnost spočítat matici \mathbf{C} (pokud má \mathbf{C} lineárně nezávislé sloupce, znamená to najít bázi nulového prostoru matice \mathbf{A}). \square

Někdy ovšem množinu (10.1) parametrizovat nejde nebo je to složité. Nyní proto odvodíme obecnější metodu, *metodu Lagrangeových multiplikátorů*.

10.1 Tečný a ortogonální prostor k povrchu

Zapomeňme nejprve na účelovou funkci f a zkoumejme jen množinu (10.1). Nejprve definujme pojem *tečný* k množině X v bodě $\mathbf{x} \in X$. Tento pojem je intuitivně zřejmý, ale jeho přesná definice není očividná.

Uvažujme diferencovatelné zobrazení

$$\gamma: (\alpha_1, \alpha_2) \rightarrow X \quad (10.4)$$

na intervalu $(\alpha_1, \alpha_2) \subseteq \mathbb{R}$. Všimněte si, že obor hodnot zobrazení je množina X , tedy (10.4) je vlastně zobrazení $\gamma: (\alpha_1, \alpha_2) \rightarrow \mathbb{R}^n$ splňující

$$\mathbf{g}(\gamma(\alpha)) = \mathbf{0} \quad (10.5)$$

pro každé $\alpha \in (\alpha_1, \alpha_2)$. Množina $\{\gamma(\alpha) \mid \alpha \in (\alpha_1, \alpha_2)\} \subseteq X$ je *hladká křivka* na množině X . Pro každé $\alpha \in (\alpha_1, \alpha_2)$ je vektor $\gamma'(\alpha) \in \mathbb{R}^n$ tečný ke křivce v bodě $\gamma(\alpha) \in X$. Tento vektor je zároveň **tečný k množině X** v bodě $\gamma(\alpha)$.

Tedy: vektor $\mathbf{v} \in \mathbb{R}^n$ je tečný k množině X v bodě $\mathbf{x} \in X$, existují-li čísla $\alpha_1 < \alpha < \alpha_2$ a diferencovatelné zobrazení (10.4) tak, že $\mathbf{x} = \gamma(\alpha)$ a $\mathbf{v} = \gamma'(\alpha)$. Množina X v daném bodě $\mathbf{x} \in X$ může mít více tečných vektorů nebo tam také nemusí mít žádný – to záleží na vlastnostech zobrazení \mathbf{g} v okolí bodu \mathbf{x} .

Tvrzení 10.1. *Necht' je zobrazení \mathbf{g} spojitě diferencovatelné v bodě $\mathbf{x} \in X$. Necht' vektor $\mathbf{v} \in \mathbb{R}^n$ je tečný k množině X v bodě \mathbf{x} . Pak $\mathbf{g}'(\mathbf{x})\mathbf{v} = \mathbf{0}$.*

Důkaz. Derivací rovnice (10.5) v bodě α s použitím řetězového pravidla máme

$$\frac{d\mathbf{g}(\gamma(\alpha))}{d\alpha} = \mathbf{g}'(\gamma(\alpha))\gamma'(\alpha) = \mathbf{g}'(\mathbf{x})\mathbf{v} = \mathbf{0}. \quad \square$$

Podmínka $\mathbf{g}'(\mathbf{x})\mathbf{v} = \mathbf{0}$ znamená, že vektor \mathbf{v} je kolmý k řádkům Jacobiho matice $\mathbf{g}'(\mathbf{x})$, což jsou gradienty $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$ složek zobrazení \mathbf{g} v bodě \mathbf{x} . Pro $m = 1$ tedy tvrzení zobecňuje skutečnost, kterou jsme bez důkazu uvedli v §8.6, totiž že gradient funkce je v každém bodě kolmý k její vrstevnici.

Následující věta říká, že za jisté podmínky to platí i naopak, tedy že *každý* vektor kolmý ke gradientům $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$ je tečný k množině X . Důkaz věty neuvádíme (plyne z tzv. *věty o implicitní funkci*).

Věta 10.2. *Necht' je zobrazení \mathbf{g} spojitě diferencovatelné v bodě $\mathbf{x} \in X$. Necht'*

$$\text{rank } \mathbf{g}'(\mathbf{x}) = m. \quad (10.6)$$

Necht' vektor $\mathbf{v} \in \mathbb{R}^n$ splňuje $\mathbf{g}'(\mathbf{x})\mathbf{v} = \mathbf{0}$. Pak vektor \mathbf{v} je tečný k množině X v bodě \mathbf{x} .

Podmínka (10.6) znamená, že gradienty $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$ musí být lineárně nezávislé. Bodu $\mathbf{x} \in X$ splňující podmínku (10.6) se někdy říká **regulární bod**¹ množiny X .

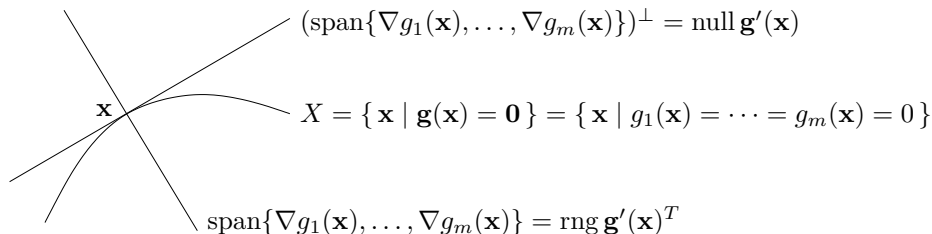
Věta 10.2 říká, že v každém regulárním bodě $\mathbf{x} \in X$ je množina všech tečných vektorů rovna nulovému prostoru Jacobiho matice $\mathbf{g}'(\mathbf{x})$. To zároveň dokazuje, že množina všech tečných

¹Jsou-li pro nějaké $\varepsilon > 0$ všechny body množiny $X \cap B_\varepsilon(\mathbf{x})$ regulární, tato množina je 'hladký povrch' v \mathbb{R}^n dimenze $n - m$. Studium 'hladkých povrchů' se zabývá zajímavá disciplína *diferenciální geometrie*.

vektorů tvoří podprostor \mathbb{R}^n . Tomuto podprostoru se říká **tečný prostor** k množině X v bodě \mathbf{x} . Jeho ortogonální doplněk (viz Věta 4.2)

$$(\text{null } \mathbf{g}'(\mathbf{x}))^\perp = \text{rng } \mathbf{g}'(\mathbf{x})^T = \text{span}\{\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})\} \quad (10.7)$$

je **ortogonální prostor** k množině X v bodě \mathbf{x} . Viz obrázek:



Příklad 10.3. 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 Věty 10.2 jsou splněny a ortogonální prostor k X v bodě (x, y) je množina $\text{span}\{\nabla g(x, y)\}$, což je přímka kolmá ke kružnici. Tečný prostor v bodě (x, y) je ortogonální doplněk této přímky, tedy přímka tečná ke kružnici. \square

Příklad 10.4. 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 $\text{span}\{\nabla g(x, y, z)\}$, 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 10.5. 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 $\text{span}\{\nabla g_1(x, y, z), \nabla g_2(x, y, z)\}$, 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 10.6. Necht' $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ je funkce $g(x, y) = (x^2 + y^2 - 1)^2$. Množina X je úplně stejná kružnice jako v Příkladě 10.3. 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 bod (x, y) není regulární. Neplatí, že každý vektor kolmý ke gradientu $\nabla g(x, y)$ (což je úplně každý vektor) je tečný ke kružnici v bodě (x, y) . Platí pouze Tvzení 10.1, totiž že každý vektor tečný ke kružnici v bodě (x, y) je kolmý ke gradientu (protože k nulovému vektoru je každý vektor kolmý). \square

10.2 Podmínky prvního řádu

Nyní přidáme do našich úvah i naši účelovou funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$.

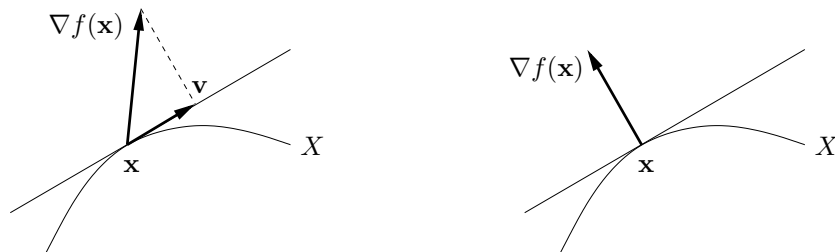
Tvrzení 10.3. Necht' $\mathbf{x} \in X$ je lokální extrém funkce f na množině X . Necht' f a \mathbf{g} jsou v bodě \mathbf{x} spojitě diferencovatelné. Necht' \mathbf{v} je tečný vektor k množině X v bodě \mathbf{x} . Pak $f'(\mathbf{x})\mathbf{v} = 0$.

Důkaz. Protože \mathbf{v} je tečný vektor k množině X v bodě \mathbf{x} , existují čísla $\alpha_1 < \alpha < \alpha_2$ a diferencovatelné zobrazení (10.4) tak, že $\mathbf{x} = \gamma(\alpha)$ a $\mathbf{v} = \gamma'(\alpha)$. Protože \mathbf{x} je lokální extrém funkce f na množině X , je to lokální extrém funkce f také na každé hladké křivce na X procházející bodem \mathbf{x} (to plyne z definice lokálního extrému, viz §8.10). Dle Věty 9.2 tedy

$$\frac{df(\gamma(\alpha))}{d\alpha} = f'(\gamma(\alpha))\gamma'(\alpha) = f'(\mathbf{x})\mathbf{v} = 0.$$

□

Tvrzení 10.3 říká, že je-li $\mathbf{x} \in X$ lokální extrém funkce f na množině X , pak gradient $\nabla f(\mathbf{x}) = f'(\mathbf{x})^T$ je kolmý na každý vektor tečný k množině X v bodě \mathbf{x} . Neboli (dle Věty 8.4) směrová derivace $f'(\mathbf{x})\mathbf{v} = \nabla f(\mathbf{x})^T \mathbf{v}$ je nulová v každém směru tečném k množině X v bodě \mathbf{x} . Toto tvrzení je intuitivně jasné. Pokud je směrová derivace nenulová (obrázek dole vlevo), bod \mathbf{x} není lokální extrém, protože bychom bodem \mathbf{x} mohli pohnout ve směru vektoru \mathbf{v} (příp. $-\mathbf{v}$) a funkci tak zvětšit (příp. zmenšit). Aby byla směrová derivace nulová, musí být gradient $\nabla f(\mathbf{x})$ kolmý na tečný vektor \mathbf{v} (obrázek vpravo).



Kombinací Tvrzení 10.3 a Věty 10.2 už snadno obdržíme hlavní výsledek této kapitoly:

Věta 10.4. *Necht' $\mathbf{x} \in X$ je lokální extrém funkce f na množině X . Necht' f a \mathbf{g} jsou v bodě \mathbf{x} spojitě diferencovatelné. Necht' platí (10.6). Pak gradient $\nabla f(\mathbf{x}) = f'(\mathbf{x})^T$ patří do ortogonálního prostoru množiny X v bodě \mathbf{x} , tj.*

$$\nabla f(\mathbf{x}) \in \text{span}\{\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})\}. \quad (10.8)$$

Důkaz. Protože \mathbf{x} je lokální extrém, dle Tvrzení 10.3 je gradient $\nabla f(\mathbf{x})$ kolmý na každý vektor tečný k množině X v bodě \mathbf{x} . Protože platí (10.6), dle Věty 10.2 je tedy $\nabla f(\mathbf{x})$ kolmý na tečný prostor null $\mathbf{g}'(\mathbf{x})$, neboli patří do ortogonálního prostoru (10.7). □

Věta 10.4 se k řešení úlohy (10.2) obvykle používá následujícím způsobem. Podmínka (10.8) říká, že existují čísla $(\lambda_1, \dots, \lambda_m) = \boldsymbol{\lambda} \in \mathbb{R}^m$ (tzv. **Lagrangeovy multiplikátory**) taková, že

$$f'(\mathbf{x}) + \lambda_1 g'_1(\mathbf{x}) + \dots + \lambda_m g'_m(\mathbf{x}) = f'(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}'(\mathbf{x}) = 0. \quad (10.9)$$

Zaved'me **Lagrangeovu funkci** $L: \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ jako

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x}) + \lambda_1 g_1(\mathbf{x}) + \dots + \lambda_m g_m(\mathbf{x}). \quad (10.10)$$

Očividně je $L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x})$ pro všechna $\mathbf{x} \in X$ a $\boldsymbol{\lambda} \in \mathbb{R}^m$. Dále si všimneme, že:

- Podmínku (10.9) lze psát jako² $\partial L(\mathbf{x}, \boldsymbol{\lambda}) / \partial \mathbf{x} = \mathbf{0}$.
- Rovnost $\partial L(\mathbf{x}, \boldsymbol{\lambda}) / \partial \boldsymbol{\lambda} = \mathbf{g}(\mathbf{x})^T = \mathbf{0}$ je ekvivalentní omezením.

²Výraz $\partial L(\mathbf{x}, \boldsymbol{\lambda}) / \partial \mathbf{x}$ označuje řádkový vektor parciálních derivací funkce L podle x_1, \dots, x_n .

Platí-li předpoklady Věty 10.4, existuje tedy vektor $\lambda \in \mathbb{R}^m$ takový, že $L'(\mathbf{x}, \lambda) = \mathbf{0}$, tj. bod $(\mathbf{x}, \lambda) \in \mathbb{R}^{m+n}$ je stacionární bod funkce L .

Příklad 10.7. Řešme znovu Příklad 10.1. 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 dají $x = y = 1/(2\lambda)$. 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, $(x, y) = \pm(1, 1)/\sqrt{2}$. \square

Příklad 10.8. Řešme Příklad 10.1, kde ale omezení změníme na $g(x, y) = (1 - x^2 - y^2)^2 = 0$. Podle Příkladu 10.6 máme $g'(x, y) = (0, 0)$ pro každé $(x, y) \in X$, čekáme tedy potíže.

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 10.9. Vraťme se k úloze (5.17), tedy k hledání řešení nehomogenní lineární soustavy s nejmenší normou. Lagrangeova funkce je

$$L(\mathbf{x}, \lambda) = \mathbf{x}^T \mathbf{x} + 2\lambda^T (\mathbf{b} - \mathbf{A}\mathbf{x}),$$

kde přidaná dvojka nemění situaci. Je $\partial L(\mathbf{x}, \lambda) / \partial \mathbf{x} = 2\mathbf{x}^T - 2\lambda^T \mathbf{A}$ (odvoďte!). Stacionární body funkce L tedy získáme řešením soustavy (5.19), kterou jsme v 5.2 odvodili úvahou. \square

Příklad 10.10. Vraťme se k úloze (7.1), kde minimalizujeme kvadratickou formu $\mathbf{x}^T \mathbf{A} \mathbf{x}$ za podmínky $\mathbf{x}^T \mathbf{x} = 1$. Předpokládáme, že \mathbf{A} je symetrická. Máme

$$L(\mathbf{x}, \lambda) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \lambda(1 - \mathbf{x}^T \mathbf{x}).$$

Rovnice $\partial L(\mathbf{x}, \lambda) / \partial \mathbf{x} = \mathbf{0}$ dá (po transpozici) $\mathbf{A} \mathbf{x} = \lambda \mathbf{x}$. Tedy (\mathbf{x}, λ) je stacionární bod funkce L , právě když λ je vlastní číslo matice \mathbf{A} a \mathbf{x} je normalizovaný (kvůli podmínce $\mathbf{x}^T \mathbf{x} = 1$) vlastní vektor příslušný λ .

Pouze z podmínek prvního řádu nelze rozhodnout, které stacionární body funkce L odpovídají minimu úlohy (7.1), které maximum a které ani jednomu. Můžeme to ale udělat úvahou: když (λ, \mathbf{x}) je stacionární bod L , máme $\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T (\lambda \mathbf{x}) = \lambda$. Tedy nejmenší [největší] vlastní číslo odpovídá minimu [maximum] a ostatní vlastní čísla neodpovídají extrémům úlohy. \square

Příklad 10.11. Řešme úlohu

$$\begin{aligned} \min \quad & \| \mathbf{Ax} - \mathbf{b} \|^2 \\ \text{za podmínek} \quad & \mathbf{Cx} = \mathbf{d} \end{aligned} \quad (10.11)$$

Tedy řešíme úlohu (5.2) s lineárními omezeními (angl. *linearly constrained least squares*).

Máme

$$L(\mathbf{x}, \boldsymbol{\lambda}) = (\mathbf{Ax} - \mathbf{b})^T (\mathbf{Ax} - \mathbf{b}) + 2\boldsymbol{\lambda}^T (\mathbf{Cx} - \mathbf{d}).$$

Rovnice $\partial L(\mathbf{x}, \boldsymbol{\lambda}) / \partial \mathbf{x} = \mathbf{0}$ dá po transpozici $\mathbf{A}^T (\mathbf{Ax} - \mathbf{b}) + \mathbf{C}^T \boldsymbol{\lambda} = \mathbf{0}$. Tedy řešíme soustavu

$$\begin{aligned} \mathbf{A}^T \mathbf{Ax} + \mathbf{C}^T \boldsymbol{\lambda} &= \mathbf{A}^T \mathbf{b} \\ \mathbf{Cx} &= \mathbf{d}. \end{aligned}$$

Napíšeme-li tuto soustavu maticově, můžeme vyjádřit

$$\begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T \mathbf{A} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \mathbf{b} \\ \mathbf{d} \end{bmatrix}.$$

Zde předpokládáme, že invertovaná matice je regulární (k tomu viz Cvičení 10.23). \square

Předchozí příklady vyžadují od studenta nejen znalost metody Lagrangeových multiplikátorů, ale i jistou zručnost v manipulaci s maticovými výrazy. Cvičte tuto zručnost ve Cvičcích 10.15–10.18!

Věta 10.4 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 §10.3. Zde pouze zdůrazníme, že druh lokálního extrému *nelze* zjistit podle definitnosti Hessovy matice $L''(\mathbf{x}, \boldsymbol{\lambda})$, tedy je chybou použít Větu 9.5 na funkci L .

10.3 (★) 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 $\mathbf{x}^T \mathbf{Ax} \geq 0$ 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 $\mathbf{y} = \mathbf{Bx}$. Protože $\mathbf{y}^T \mathbf{Ay} = \mathbf{x}^T \mathbf{B}^T \mathbf{ABx}$, převedli jsme problém na ověřování pozitivní semidefinitnosti matice $\mathbf{B}^T \mathbf{AB}$. Podobně definujeme pozitivní a negativní (semi)definitnost a indefinitnost matice \mathbf{A} na podprostoru X .

Věta 10.5. *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 $L'(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$,*
- *f a \mathbf{g} jsou dvakrát diferencovatelné v bodě \mathbf{x} .*

Pak platí:

- *Je-li \mathbf{x} lokální minimum [maximum] funkce f vázané podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, pak matice $\partial^2 L(\mathbf{x}, \boldsymbol{\lambda}) / \partial \mathbf{x}^2$ je pozitivně [negativně] semidefinitní na nulovém prostoru matice $\mathbf{g}'(\mathbf{x})$.*
- *Je-li $\partial^2 L(\mathbf{x}, \boldsymbol{\lambda}) / \partial \mathbf{x}^2$ pozitivně [negativně] definitní na nulovém prostoru matice $\mathbf{g}'(\mathbf{x})$, pak \mathbf{x} je ostré lokální minimum [maximum] funkce f 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 $\mathbf{g}'(\mathbf{x})$, pak \mathbf{x} není lokální extrém funkce f vázaný podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

Zde

$$\frac{\partial^2 L(\mathbf{x}, \boldsymbol{\lambda})}{\partial \mathbf{x}^2} = f''(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i''(\mathbf{x})$$

značí druhou derivaci (Hessovu matici) funkce $L(\mathbf{x}, \boldsymbol{\lambda})$ podle \mathbf{x} v bodě $(\mathbf{x}, \boldsymbol{\lambda})$.

Příklad 10.12. Najděme strany kvádrů 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

$$\frac{\partial^2 L(x, y, z, \lambda)}{\partial (x, y, z)^2} = \begin{bmatrix} 0 & 1 - \lambda z & 1 - \lambda y \\ 1 - \lambda z & 0 & 1 - \lambda x \\ 1 - \lambda y & 1 - \lambda x & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}. \quad (10.12)$$

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 štestí, zda matice (10.12) 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. □

10.4 Cvičení

Následující úlohy vyřešte nejdříve převodem na hledání extrémů funkcí jedné proměnné dle Příkladu 10.1 (je-li to možné) a potom metodou Lagrangeových multiplikátorů. V druhém případě nemusíte ověřovat podmínky druhého řádu, stačí jen najít stacionární body Lagrangeovy funkce (lze-li ale usoudit na druh extrému nějakou snadnou úvahou, udělejte to).

10.1. Na kružnici $x^2 + y^2 = 1$ najděte lokální extrémy funkce

- a) $f(x, y) = 2x - y$
- b) $f(x, y) = x(y - 1)$
- c) $f(x, y) = x^2 + 2y^2$
- d) $f(x, y) = x^2y$
- e) $f(x, y) = x^4 + y^2$
- f) $f(x, y) = \sin(xy)$
- g) $f(x, y) = e^{xy}$

10.2. Na sféře $x^2 + y^2 + z^2 = 1$ najděte lokální extrémy funkce

- a) $f(x, y, z) = (x + y)(y + z)$
- b) $f(x, y, z) = a/x + b/y + c/z$, kde $a, b, c > 0$ jsou dány
- c) $f(x, y, z) = x^3 + y^2 + z$
- d) $f(x, y, z) = x^3 + y^3 + z^3 + 2xyz$
- e) $(\star) f(x, y, z) = x^3 + y^3 + z^3 - 3xyz$
- f) $(\star) f(x, y, z) = x^3 + 2xyz - z^3$

10.3. Najděte lokální extrémy funkce

- a) $f(x, y, z) = x + yz$ za podmínek $x^2 + y^2 + z^2 = 1$ a $z^2 = x^2 + y^2$
- b) $f(x, y, z) = xyz$ za podmínek $x^2 + y^2 + z^2 = 1$ a $xy + yz + zx = 1$

10.4. Najděte bod nejblíže počátku na křivce

- a) $x + y = 1$
- b) $x + 2y = 5$
- c) $x^2y = 1$
- d) $x^2 + 2y^2 = 1$

10.5. 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) (\star) 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 $S = \pi(R + r)\sqrt{(R - r)^2 + h^2}$. Použijte vhodný numerický algoritmus na řešení vzniklé soustavy rovnic.

10.6. Rozložte dané kladné reálné číslo na součin n kladných reálných čísel tak, aby jejich součet byl co nejmenší.

- 10.7. Dokažte, že funkce $f(x, y) = x$ nabývá za podmínky $x^3 = y^2$ minimum pouze v počátku. Ukažte, že metoda Lagrangeových multiplikátorů toto minimum nenajde.
- 10.8. 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}^* .
- 10.9. Do elipsy o daných délkách os vepište obdélník s maximálním obsahem. Předpokládejte přitom, že strany obdélníku jsou rovnoběžné s osami elipsy.
- 10.10. *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:
- Zákon odrazu od zrcadla: úhel dopadu se rovná úhlu odrazu.
 - 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:

- Pro rovinné zrcadlo a rovinné rozhraní (což vede na minimalizaci bez omezení).
 - Pro zrcadlo a rozhraní tvaru obecné plochy s rovnicí $g(\mathbf{x}) = 0$. Dokažte najít situaci, kdy skutečná dráha paprsku má čas *větší* než jí blízké dráhy?
- 10.11. Rozdělení pravděpodobnosti diskrétní náhodné proměnné je funkce $p: \{1, \dots, n\} \rightarrow \mathbb{R}_+$ (tj. soubor nezáporných čísel $p(1), \dots, p(n)$) splňující $\sum_{x=1}^n p(x) = 1$.
- 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.
 - Dokažte *Gibbsovu nerovnost* (též zvanou *informační nerovnost*): pro každé dvě rozdělení p, q platí

$$\sum_{x=1}^n p(x) \log q(x) \leq \sum_{x=1}^n p(x) \log p(x),$$

přičemž rovnost nastává jen tehdy, když $p = q$.

- 10.12. (★) 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?
- 10.13. (★) 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.
- 10.14. (★) 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 $x, y, z, u, v, w > 0$.

- 10.15. Minimalizujte $\mathbf{x}^T \mathbf{x}$ za podmínky $\mathbf{a}^T \mathbf{x} = 1$. Jaký je geometrický význam úlohy?

- 10.16. Maximalizujte $\mathbf{a}^T \mathbf{x}$ za podmínky $\mathbf{x}^T \mathbf{x} = 1$. Jaký je geometrický význam úlohy?
- 10.17. Minimalizujte $\mathbf{x}^T \mathbf{A} \mathbf{x}$ za podmínky $\mathbf{b}^T \mathbf{x} = 1$, kde \mathbf{A} je symetrická pozitivně definitní.
- 10.18. Minimalizujte $\mathbf{x}^T \mathbf{C} \mathbf{x}$ za podmínky $\mathbf{A} \mathbf{x} = \mathbf{b}$, kde \mathbf{A} má lineárně nezávislé řádky a \mathbf{C} je symetrická pozitivně definitní.
- 10.19. (★) Minimalizujte $\mathbf{x}^T \mathbf{A} \mathbf{x}$ za podmínek $\mathbf{B} \mathbf{x} = \mathbf{0}$ a $\mathbf{x}^T \mathbf{x} = 1$, kde \mathbf{A} je pozitivně definitní.
- 10.20. (★) Minimalizujte $\mathbf{x}^T \mathbf{A} \mathbf{x}$ za podmínky $\mathbf{x}^T \mathbf{B} \mathbf{x} = 1$, kde \mathbf{A} a \mathbf{B} jsou pozitivně definitní.
- 10.21. (★) Minimalizujte $\mathbf{a}^T (\mathbf{x} - \mathbf{b})$ za podmínky $\mathbf{x}^T \mathbf{C} \mathbf{x} = 1$, kde \mathbf{C} je pozitivně definitní.
- 10.22. (★) Jaké musí být vlastnosti matice \mathbf{A} a vektoru \mathbf{b} , aby $\max\{\|\mathbf{A} \mathbf{x}\| \mid \mathbf{b}^T \mathbf{x} = 0\} = 0$?
- 10.23. (★) Dokažte, že matice $\begin{bmatrix} \mathbf{A}^T \mathbf{A} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix}$ je regulární právě tehdy, když matice $\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix}$ má lineárně nezávislé sloupce a matice \mathbf{C} má lineárně nezávislé řádky (viz Příklad 10.11).

Nápověda a řešení

- 10.1. Někdy je možné účelovou funkci zjednodušit. Např. extrémy funkce e^{xy} se nabývají ve stejných bodech jako extrémy funkce xy , protože funkce e^t je rostoucí.
- 10.4.c) Jsou dva body nejbližší počátku: $x = \pm 2^{1/6}$, $y = 2^{-1/3}$.
- 10.10.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 pólí úhel mezi nimi. Pro důkaz druhého tvrzení násobte rovnici vektory $(\mathbf{x} - \mathbf{a})^0$ a $(\mathbf{x} - \mathbf{b})^0$ a porovnejte.
- 10.11.b) Maximalizujte levou stranu přes q za podmínky $\sum_x q(x) = 1$. (Podmínka $q(x) \geq 0$ je implicitně zajištěna tím, že když se některá složka q blíží shora nule, účelová funkce se zhoršuje nade všechny meze.)
- 10.17. $\mathbf{x} = \mathbf{A}^{-1} \mathbf{b} / (\mathbf{b}^T \mathbf{A}^{-1} \mathbf{b})$
- 10.18. $\mathbf{x} = \mathbf{C}^{-1} \mathbf{A}^T (\mathbf{A} \mathbf{C}^{-1} \mathbf{A}^T)^{-1} \mathbf{b}$.

Část III

Lineární programování

Kapitola 11

Lineární programování

Úloha **lineárního programování**¹ (LP, také zvané lineární optimalizace) je definována jako minimalizace nebo maximalizace lineární funkce za omezujících podmínek ve tvaru lineárních rovnic a nerovnic. Zde **lineární rovnici** rozumíme výrok

$$a_1x_1 + \cdots + a_nx_n = b,$$

neboli krátce $\mathbf{a}^T \mathbf{x} = b$. **Lineární nerovnici** rozumíme jeden z výroků

$$a_1x_1 + \cdots + a_nx_n \leq b, \quad a_1x_1 + \cdots + a_nx_n \geq b,$$

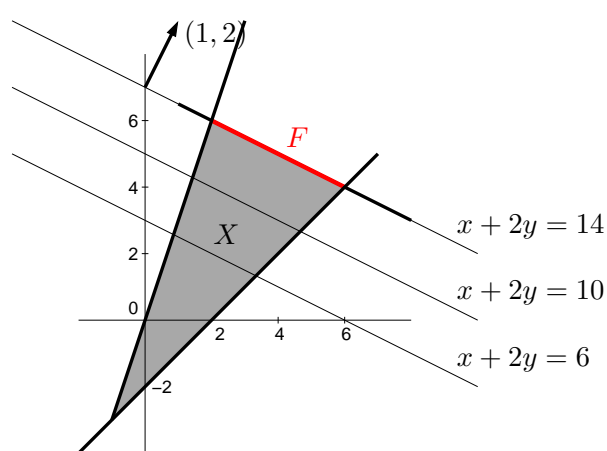
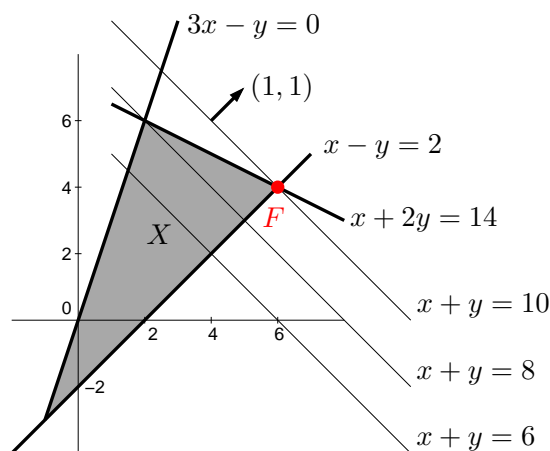
neboli $\mathbf{a}^T \mathbf{x} \leq b$ či $\mathbf{a}^T \mathbf{x} \geq b$. Lineární program je tedy úloha (1.9) ve které je funkce f lineární (tj. tvaru (3.6)) a funkce g_i, h_i jsou afinní (tj. tvaru (3.17)).

Lineární programy s jednou nebo dvěma proměnnými lze řešit graficky.

Příklad 11.1. Příkladem lineárního programu je optimalizační úloha

$$\begin{aligned} \max \quad & x + y \\ \text{za podmínek} \quad & x + 2y \leq 14 \\ & 3x - y \geq 0 \\ & x - y \leq 2 \end{aligned} \tag{11.1}$$

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:



¹Slovo *programování* zde má trochu jiný význam než znáte: místo tvoření sekvenčního počítačového kódu, který řeší daný problém, to znamená hledání vhodné účelové funkce a omezujících podmínek tak, aby optimální řešení bylo řešením daného problému. Optimalizaci obecně se také někdy říká *matematické programování*.

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 obrázek vlevo) účelová funkce na množině X nabývá maxima v bodě $(x, y) = (6, 4)$. Úloha má tedy jediné optimální řešení.

Pokud bychom účelovou funkci úlohy změnili na $x + 2y$, množina optimálních řešení úlohy by byla úsečka spojující body $(2, 6)$ a $(6, 4)$ (viz obrázek vpravo), úloha by tedy měla nekonečně mnoho optimálních řešení. Pokud by účelová funkce byla nulová (tj. $0x + 0y$), množina optimálních řešení úlohy by byla celý trojúhelník X . \square

Z našich úvah je patrné (přesně ukážeme 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 bez porušení omezení libovolně zlepšovat).

Jestliže úloha má optimální řešení, pak množina optimálních řešení je buď vrchol mnohoúhelníku nebo jeho hrana nebo celý mnohoúhelník.

11.1 Transformace úloh LP

Často je užitečné formulovat lineární program v nějakém speciálním tvaru, kdy jsou dovoleny jen určité typy omezení. To obvykle vyžadují např. algoritmy na řešení LP.

Jeden speciální tvar je tvar, kdy minimalizujeme (ne tedy maximalizujeme) a dovolíme pouze omezení typu \geq (větší nebo rovno):

$$\begin{aligned} \min \quad & c_1 x_1 + \cdots + c_n x_n \\ \text{za podmíněk} \quad & a_{i1} x_1 + \cdots + a_{in} x_n \geq b_i, \quad i = 1, \dots, m \end{aligned} \quad (11.2a)$$

To se pohodlněji napíše v maticovém tvaru

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A} \mathbf{x} \geq \mathbf{b} \}, \quad (11.2b)$$

kde $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$. Zápis $\mathbf{A} \mathbf{x} \geq \mathbf{b}$ značí, že i -tá složka vektoru $\mathbf{A} \mathbf{x}$ není menší než i -tá složka vektoru \mathbf{b} , pro všechna $i = 1, \dots, m$. Jakýkoliv lineární program snadno převedeme na tvar (11.2) těmito 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 $\mathbf{a}^T \mathbf{x} \geq b$, $-\mathbf{a}^T \mathbf{x} \geq -b$.

Jiný často užívaný speciální tvar je **rovníkový tvar**², ve kterém všechna omezení jsou rovnosti a všechny proměnné jsou nezáporné, tedy

$$\begin{aligned} \min \quad & c_1 x_1 + \cdots + c_n x_n \\ \text{za podmíněk} \quad & a_{i1} x_1 + \cdots + a_{in} x_n = b_i, \quad i = 1, \dots, m \\ & x_j \geq 0, \quad j = 1, \dots, n \end{aligned} \quad (11.3a)$$

neboli

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A} \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}. \quad (11.3b)$$

Tvar (11.2) lze převést na tvar (11.3) dvěma úpravami, ve kterých zavedeme nové proměnné:

²Tomuto 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.

- Nerovnost $\mathbf{a}^T \mathbf{x} \leq b$ nahradíme dvěma omezeními $\mathbf{a}^T \mathbf{x} + u = b$, $u \geq 0$. Pomocné proměnné u se říká **slacková proměnná**³. Podobně převedeme nerovnost $\mathbf{a}^T \mathbf{x} \geq b$ na rovnost.
- Neomezenou proměnnou $x \in \mathbb{R}$ rozdělíme na dvě nezáporné proměnné $x^+ \geq 0$, $x^- \geq 0$ přidáním podmínky $x = x^+ - x^-$.

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 a rychle’⁴ získat z argumentu optima nové úlohy.

Příklad 11.2. Úlohu (11.1) převedme na tvar, ve kterém minimalizujeme (místo maximalizujeme) a omezení jsou rovnosti, přičemž některé proměnné mohou být nezáporné. To uděláme zavedením slackových proměnných. Nová úloha je

$$\begin{aligned} \min \quad & -x - y \\ \text{za podmínek} \quad & x + 2y + u = 14 \\ & 3x - y - v = 0 \\ & x - y + w = 2 \\ & u, v, w \geq 0 \end{aligned}$$

Zde proměnné x, y mohou mít libovolné znaménko. Převedme úlohu na tvar (11.3), kde všechny proměnné jsou nezáporné. Dosadíme $x = x_+ - x_-$ a $y = y_+ - y_-$, kde $x_+, x_-, y_+, y_- \geq 0$. Výsledná úloha je

$$\begin{aligned} \min \quad & -x_+ + x_- - y_+ + y_- \\ \text{za podmínek} \quad & x_+ - x_- + 2y_+ - 2y_- + u = 14 \\ & 3x_+ - 3x_- - y_+ + y_- - v = 0 \\ & x_+ - x_- - y_+ + y_- + w = 2 \\ & x_+, x_-, y_+, y_-, u, v, w \geq 0 \end{aligned}$$

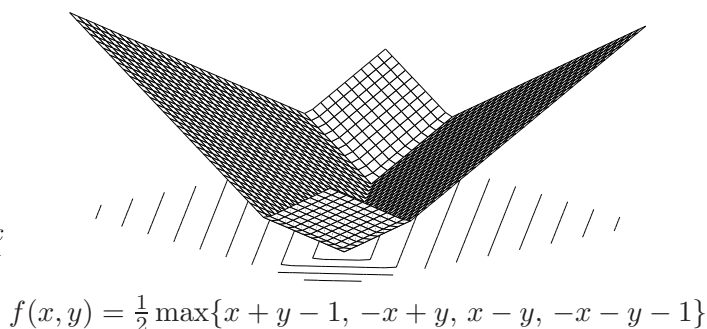
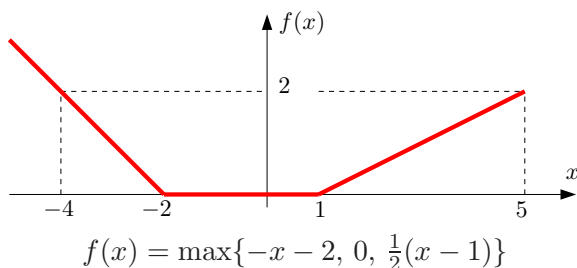
□

11.1.1 Po částech afinní funkce

Mějme funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ danou vzorcem

$$f(\mathbf{x}) = \max_{i=1}^k (\mathbf{c}_i^T \mathbf{x} + d_i), \quad (11.4)$$

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. Na obrázku jsou příklady (vlevo pro $n = 1$ a $k = 3$, vpravo pro $n = 2$ a $k = 4$):



³Slack znamená anglicky např. mezeru mezi zdí a skříní, která není zcela přiřažená ke zdi. Termín *slack variable* nemá ustálený český ekvivalent, někdy se překládá jako *skluzová proměnná*.

⁴Přesněji v *lineárním čase*, ve smyslu teorie algoritmů.

Minimalizujeme funkci (11.4) za podmínek $\mathbf{Ax} \geq \mathbf{b}$. To není úloha LP, neboť její účelová funkce není lineární. Ovšem lze ji převést na LP zavedením pomocné proměnné:

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} \geq \mathbf{b}\} = \min\{z \mid (\mathbf{x}, z) \in \mathbb{R}^{n+1}, f(\mathbf{x}) \leq z, \mathbf{Ax} \geq \mathbf{b}\} \quad (11.5a)$$

$$= \min\{z \mid (\mathbf{x}, z) \in \mathbb{R}^{n+1}, \max_i(\mathbf{c}_i^T \mathbf{x} + d_i) \leq z, \mathbf{Ax} \geq \mathbf{b}\} \quad (11.5b)$$

$$= \min\{z \mid (\mathbf{x}, z) \in \mathbb{R}^{n+1}, (\forall i)(\mathbf{c}_i^T \mathbf{x} + d_i \leq z), \mathbf{Ax} \geq \mathbf{b}\} \quad (11.5c)$$

kde (11.5c) je již úloha LP. Rovnost (11.5a) platí proto, že v optimu pravé úlohy je $f(\mathbf{x}) = z$. Kdyby totiž bylo $f(\mathbf{x}) < z$, mohli bychom z zmenšit bez porušení omezení a tedy (\mathbf{x}, z) by nebylo optimum. Rovnost (11.5c) platí proto, že pro libovolná čísla a_i, b platí (samozřejmě!)

$$\max_i a_i \leq b \iff (\forall i)(a_i \leq b). \quad (11.6)$$

Příklad 11.3. Úloha

$$\begin{aligned} \min \quad & \max\{3x + 4y + 1, 2x - 3y\} \\ \text{za podm.} \quad & x + 2y \leq 14 \\ & 3x - y \geq 0 \\ & x - y \leq 2 \end{aligned}$$

není LP, protože účelová funkce $f(x, y) = \max\{3x + 4y + 1, 2x - 3y\}$ není lineární ani afinní (načrtněte si na papír několik jejích vrstevnic!). Úlohu lze ale převést na LP

$$\begin{aligned} \min \quad & z \\ \text{za podm.} \quad & 3x + 4y - z \leq -1 \\ & 2x - 3y - z \leq 0 \\ & x + 2y \leq 14 \\ & 3x - y \geq 0 \\ & x - y \leq 2 \end{aligned}$$

□

Tento převod lze užít i pro funkce obsahující absolutní hodnoty, neboť $|x| = \max\{-x, x\}$, a na různé jiné případy (trénujte to ve Cvičení 11.4!). Buďte ale opatrní: 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.

Ve zbytku této kapitoly uvedeme některá z mnoha použití lineárního programování.

11.2 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

Chceme zjistit, kolik jakých výrobků máme vyrobit, aby zisk byl největší. Tuto úlohu lze formalizovat lineárním programem

$$\max \sum_{j=1}^n c_j x_j \quad (11.7a)$$

$$\text{za podm. } \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = 1, \dots, m \quad (11.7b)$$

$$x_j \geq 0, \quad j = 1, \dots, n \quad (11.7c)$$

což jde kratčeji napsat v maticové formě jako $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$.

Příklad 11.4. 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 $(120 - 2 \cdot 12 - 0.4 \cdot 40)l + (76 - 1.5 \cdot 12 - 0.2 \cdot 40)h = 80l + 50h$.

V obou případech je optimální řešení $l = 20$ kg lupínků a $h = 40$ kg hranolků. \square

11.3 Směšovací (výživová) úloha

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

Tuto úlohu lze formalizovat lineárním programem

$$\min \sum_{j=1}^n c_j x_j \quad (11.8a)$$

$$\text{za podm. } \sum_{j=1}^n a_{ij} x_j \geq b_i, \quad i = 1, \dots, m \quad (11.8b)$$

$$x_j \geq 0, \quad j = 1, \dots, n \quad (11.8c)$$

neboli $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$.

Příklad 11.5. 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 varíte 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?

Řešíme LP

$$\begin{aligned}
& \min \quad 25b + 50m + 80z \\
& \text{za podmínek} \quad \begin{aligned}
2b + m + z &\geq 8 \\
2b + 6m + z &\geq 16 \\
b + 3m + 6z &\geq 8 \\
b, m, z &\geq 0
\end{aligned}
\end{aligned}$$

Optimální řešení je $b = 3.2$, $m = 1.6$, $z = 0$ s hodnotou 160. □

11.4 Dopravní úloha

Máme m skladů a n spotřebitelů.

- a_i = množství ve skladě i
- b_j = množství zboží požadované spotřebitelem j
- c_{ij} = cena dopravy jednotky zboží ze skladu i ke spotřebiteli j
- x_{ij} = množství zboží vezené ze skladu i ke spotřebiteli j

Chceme co nejlevněji rozvézt zboží ze skladů ke spotřebitelům. Řešením je lineární program

$$\min \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \quad (11.9a)$$

$$\text{za podm.} \quad \sum_{j=1}^n x_{ij} = a_i, \quad i = 1, \dots, m \quad (11.9b)$$

$$\sum_{i=1}^m x_{ij} = b_j, \quad j = 1, \dots, n \quad (11.9c)$$

$$x_{ij} \geq 0, \quad i = 1, \dots, m, j = 1, \dots, n \quad (11.9d)$$

Ten jde také napsat jako (rozmyslete!)

$$\min \{ \mathbf{C} \cdot \mathbf{X} \mid \mathbf{X} \in \mathbb{R}^{m \times n}, \mathbf{X}\mathbf{1} = \mathbf{a}, \mathbf{X}^T \mathbf{1} = \mathbf{b}, \mathbf{X} \geq \mathbf{0} \}, \quad (11.10)$$

kde $\mathbf{a} = (a_1, \dots, a_m)$, $\mathbf{b} = (b_1, \dots, b_n)$, $\mathbf{C} = [c_{ij}]$, $\mathbf{C} \cdot \mathbf{X}$ značí skalární součin matic (4.25) a optimalizujeme přes matice $\mathbf{X} = [x_{ij}]$.

Zadání musí splňovat $\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$ (nabídka je rovna poptávce), jinak je úloha nepřipustná. To snadno vidíme: součet levých stran rovnic (11.9b) je roven součtu levých stran rovnic (11.9c). Úloha jde modifikovat tak, že dovolíme $\sum_{i=1}^m a_i \geq \sum_{j=1}^n b_j$, tedy zboží na skladech může být více než žádají spotřebitelé. Pak by se omezení (11.9b) muselo změnit na $\sum_{j=1}^n x_{ij} \leq a_i$ (promyslete!).

11.5 Přeurčené lineární soustavy

11.5.1 Vektorové normy

Norma formalizuje pojem ‘délky’ vektoru \mathbf{x} . Známe již eukleidovskou normu, ale existují i jiné normy. Obecně pojem norma definujeme takto: 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 výšky 1. Díky homogenitě je jednotková sféra středově symetrická a její tvar zcela určuje normu.

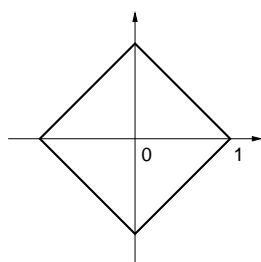
Uveďme příklady norem. Důležitou skupinou norem jsou **p -normy**

$$\|\mathbf{x}\|_p = (|x_1|^p + \dots + |x_n|^p)^{1/p}.$$

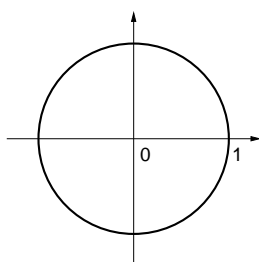
Musí být $p \geq 1$, jinak neplatí trojúhelníková nerovnost. Nejčastěji narazíte na:

- $\|\mathbf{x}\|_1 = |x_1| + \dots + |x_n|$. 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*.
- $\|\mathbf{x}\|_\infty = \lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \max\{|x_1|, \dots, |x_n|\}$ (spočtete limitu ve Cvičení 11.12!). Někdy se jí říká *Čebyševova norma* nebo *max-norma*.

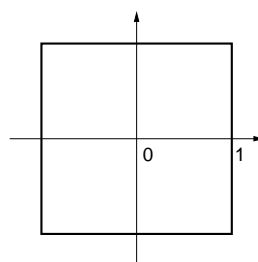
Jednotkové sféry těchto norem v \mathbb{R}^2 vypadají takto:



$$\|\mathbf{x}\|_1 = 1$$



$$\|\mathbf{x}\|_2 = 1$$



$$\|\mathbf{x}\|_\infty = 1$$

Všimněte si, že pro $n = 1$ se p -norma pro libovolné $p \geq 1$ redukuje na *absolutní hodnotu* $|x|$.

Existují ovšem 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} matice s lineárně nezávislými sloupci, je také $\|\mathbf{A}\mathbf{x}\|$ norma.

11.5.2 Přibližné řešení přeuerčených lineárních soustav

Mějme přeuerčenou 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 formulujme jako úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_p. \quad (11.11)$$

Uvažujme tři případy:

- Pro $p = \infty$ hledáme takové \mathbf{x} , které minimalizuje výraz

$$\|\mathbf{Ax} - \mathbf{b}\|_\infty = \max_{i=1}^m |\mathbf{a}_i^T \mathbf{x} - b_i|, \quad (11.12)$$

tedy minimalizuje maximální residuum. Toto řešení je známé pod názvem *minimaxní* nebo *Čebyševovo*. Podle §11.1.1 je tato úloha ekvivalentní lineárnímu programu

$$\begin{aligned} \min \quad & z \\ \text{za podm.} \quad & -z \leq \mathbf{a}_i^T \mathbf{x} - b_i \leq z, \quad i = 1, \dots, m \end{aligned} \quad (11.13a)$$

což lze napsat kratěji v maticové formě

$$\min\{z \mid \mathbf{x} \in \mathbb{R}^n, z \in \mathbb{R}, -z\mathbf{1} \leq \mathbf{Ax} - \mathbf{b} \leq z\mathbf{1}\}. \quad (11.13b)$$

- 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

$$\|\mathbf{Ax} - \mathbf{b}\|_1 = \sum_{i=1}^m |\mathbf{a}_i^T \mathbf{x} - b_i|, \quad (11.14)$$

kde $\mathbf{a}_1, \dots, \mathbf{a}_m$ jsou řádky matice \mathbf{A} . Úloha je ekvivalentní lineárnímu programu

$$\begin{aligned} \min \quad & z_1 + \dots + z_m \\ \text{za podm.} \quad & -z_i \leq \mathbf{a}_i^T \mathbf{x} - b_i \leq z_i, \quad i = 1, \dots, m \end{aligned} \quad (11.15a)$$

neboli

$$\min\{\mathbf{1}^T \mathbf{z} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{R}^m, -\mathbf{z} \leq \mathbf{Ax} - \mathbf{b} \leq \mathbf{z}\}. \quad (11.15b)$$

11.5.3 Lineární regrese

Vraťme se k lineární regresi z §5.1.3 (znovu přečtěte!). Funkční závislost přibližně popsanou naměřenými dvojicemi (x_i, y_i) , $i = 1, \dots, m$, jsme aproximovali regresní funkcí

$$f(x, \boldsymbol{\theta}) = \theta_1 \varphi_1(x) + \dots + \theta_n \varphi_n(x) = \boldsymbol{\varphi}(x)^T \boldsymbol{\theta},$$

kde parametry $\boldsymbol{\theta}$ jsou takové, aby $y_i \approx f(x_i, \boldsymbol{\theta})$ pro všechna i . Přibližné rovnosti \approx jsme chápali ve smyslu nejmenších čtverců, tedy hledali jsme takové $\boldsymbol{\theta}$ které minimalizovalo funkci

$$\sum_{i=1}^m (y_i - f(x_i, \boldsymbol{\theta}))^2 = \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|_2, \quad (11.16)$$

kde $\mathbf{y} = (y_1, \dots, y_m)$ a prvky matice \mathbf{A} jsou $a_{ij} = \varphi_j(x_i)$. Tedy řešíme úlohu (11.11) pro $p = 2$. Můžeme ale použít i jiné normy než eukleidovskou. Pro $p = 1$ minimalizujeme

$$\sum_{i=1}^m |y_i - f(x_i, \boldsymbol{\theta})| = \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|_1 \quad (11.17)$$

a pro $p = \infty$ minimalizujeme

$$\max_{i=1}^m |y_i - f(x_i, \boldsymbol{\theta})| = \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|_\infty. \quad (11.18)$$

Dále ukážeme, k čemu to může být dobré.

Regrese ve smyslu ∞ -normy je vhodná např. při aproximaci funkcí.

Příklad 11.6. Na počítači bez matematického koprocessoru potřebujeme mnohokrát vyhodnocovat funkci sinus na intervalu $[0, \frac{\pi}{2}]$. Protože výpočet hodnot této funkce by trval příliš dlouho, chceme ji aproximovat polynomem třetího stupně $\theta_1 + \theta_2 x + \theta_3 x^2 + \theta_4 x^3$, jehož hodnoty se spočítají rychleji. Spočítejme hodnoty $y_i = \sin x_i$ funkce v dostatečném počtu bodů $x_i = \frac{\pi i}{2m}$ pro $i = 1, \dots, m$. Koeficienty polynomu hledáme minimalizací Čebyševova kritéria (11.18), 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 chybně (např. se někdo při zapisování výsledků měření 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 velmi jednoduchém případě regrese: odhadu hodnoty jediného čísla ze souboru jeho nepřesných měření. To je vlastně polynomiální regrese polynomem nultého stupně $f(x, \theta) = \theta$, tj. konstantní funkcí (viz Příklad 5.4). Pro daná čísla $\mathbf{y} = (y_1, \dots, y_m) \in \mathbb{R}^m$ hledáme $\theta \in \mathbb{R}$ minimalizující výraz

$$\|\mathbf{y} - \mathbf{1}\theta\|_p = \|(y_1 - \theta, \dots, y_m - \theta)\|_p. \quad (11.19)$$

- Pro $p = \infty$ minimalizujeme $\max_{i=1}^m |y_i - \theta|$. Řešením je $\theta = \frac{1}{2}(\min_{i=1}^m y_i + \max_{i=1}^m y_i)$, tedy bod v polovině mezi krajními body.
- Pro $p = 2$ minimalizujeme $\sqrt{\sum_{i=1}^m (y_i - \theta)^2}$. Řešením je aritmetický průměr, $\theta = \frac{1}{m} \sum_{i=1}^m y_i$ (viz Příklad 5.4).
- Pro $p = 1$ minimalizujeme $\sum_{i=1}^m |y_i - \theta|$. Řešením je *medián* z čísel y_1, \dots, y_m (viz Cvičení 11.7). Medián se vypočte tak, že seřadíme čísla y_1, \dots, y_m 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 θ pro různá p budou chovat různě. Např. aritmetický průměr se bude zvětšovat, a to tak, že zvětšováním hodnoty y_1 dosáhneme *libovolně velké* hodnoty θ . Pro medián to ovšem neplatí – zvětšováním jediného bodu y_1 ovlivníme θ jen natolik, nakolik to změní pořadí bodů. Jeho libovolným zvětšováním nedosáhneme libovolně velké hodnoty θ .

Příklad 11.7. Šuplérrou změříme průměr ocelové kuličky v několika místech, dostaneme hodnoty $y_1 = 1.02$, $y_2 = 1.04$, $y_3 = 0.99$, $y_4 = 2.03$ (cm). Při posledním měření jsme se na stupnici přehlédli, proto je poslední hodnota úplně špatně. Z těchto měření chceme odhadnout skutečný průměr. Máme

$$\frac{1}{2} \left(\min_{i=1}^m y_i + \max_{i=1}^m y_i \right) = 1.51, \quad \frac{1}{m} \sum_{i=1}^m y_i = 1.27, \quad \max_{i=1}^m y_i = 1.03.$$

Je zjevné, že medián je neovlivněn vychýleným bodem, zatímco ostatní odhady ano. \square

Ve složitějším případě, např. prokládání dat polynomem jako v Příkladu 5.4, se nedá robustnost řešení ve smyslu 1-normy takto jednoduše formálně ukázat a analýza je obtížnější. Výsledek ale bude podobný: řešení ve smyslu 1-normy je méně citlivé na vychýlené body než řešení ve smyslu 2-normy.

11.6 Celočíselné lineární programování, LP relaxace

Celočíselné lineární programování (angl. *integer linear programming*, ILP) se liší od lineárního programování dodatečným omezením, že proměnné musí nabývat pouze celočíselných hodnot. Nejčastěji používané jsou *binární* proměnné, které mohou nabývat jen dvou hodnot 0 a 1. ILP s binárními proměnnými je tedy úloha

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \{0, 1\}^n, \mathbf{Ax} \geq \mathbf{b} \}. \quad (11.20)$$

Množina přípustných řešení této úlohy obsahuje konečný (avšak obvykle obrovský) počet izolovaných bodů, jedná se o typickou úlohu *kombinatorické optimalizace* (viz §1.2). Zatímco vyřešit obyčejný lineární program je relativně snadné (LP je řešitelné v polynomiálním čase), vyřešit ILP je obecně nesmírně obtížné (ILP je tzv. *NP-těžké*⁵).

Mohli bychom zkusit úlohu (11.20) jednoduše nahradit lineárním programem

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in [0, 1]^n, \mathbf{Ax} \geq \mathbf{b} \}. \quad (11.21)$$

To je opravdu lineární program, protože omezení $\mathbf{x} \in [0, 1]^n$ lze napsat jako $0 \leq x_i \leq 1$ ($i = 1, \dots, n$), což jsou lineární nerovnosti. Úloze (11.21) se říká **LP relaxace** (angl. *relaxation* znamená uvolnění, myšleno jako ‘uvolnění’ omezení) úlohy (11.20). LP relaxace je tedy nahrazení (těžké) úlohy ILP (snadnou) úlohou LP, ve které jsme vypustili omezení na celočíselnost proměnných.

Mohli bychom si myslet, že když optimální řešení $\mathbf{x} \in [0, 1]^n$ LP relaxace (11.21) nějak šikovně zaokrouhlíme na celočíselný vektor $\mathbf{x}' \in \{0, 1\}^n$, dostaneme řešení (nebo aspoň přibližné řešení) úlohy (11.20). To ale bohužel obecně neplatí. Zaokrouhlení totiž může zničit přípustnost, tedy zaokrouhlené proměnné \mathbf{x}' nemusí splňovat podmínky $\mathbf{Ax}' \geq \mathbf{b}$. I to nejšíkovnější zaokrouhlení nám nepomůže, neboť už jen odpovědět na otázku, zda úloha (11.20) je přípustná (tedy zda existuje alespoň jedno $\mathbf{x} \in \{0, 1\}^n$ splňující $\mathbf{Ax} \geq \mathbf{b}$), je obecně stejně těžké jako

⁵Tento pojem patří do *teorie složitosti algoritmů*, kterou jste ještě nebrali. Zde jen řekneme, že pro NP-těžkou úlohu je velmi malá naděje, že bude někdy nalezen algoritmus, který by ji řešil v polynomiálním čase. Algoritmus řeší úlohu (11.20) v *polynomiálním čase*, jestliže existuje polynom p takový, že algoritmus pro každou trojici $(\mathbf{A}, \mathbf{b}, \mathbf{c})$ vyřeší úlohu v čase menším než $p(L)$, kde L je počet bitů potřebných k zápisu všech prvků matic $\mathbf{A}, \mathbf{b}, \mathbf{c}$ v binárním kódu.

najít optimální řešení úlohy (11.20). I kdybychom však měli štěstí a zaokrouhlený vektor \mathbf{x}' byl přípustný, hodnota $\mathbf{c}^T \mathbf{x}'$ účelové funkce může být obecně libovolně daleko od optimální hodnoty úlohy (11.20).

Jediný vztah, který obecně platí mezi optimálním řešením původní a relaxované úlohy, je tento: optimální hodnota relaxované úlohy není větší než optimální hodnota původní úlohy,

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in [0, 1]^n, \mathbf{Ax} \geq \mathbf{b}\} \leq \min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \{0, 1\}^n, \mathbf{Ax} \geq \mathbf{b}\}. \quad (11.22)$$

To proto, že v relaxované úloze minimalizujeme přes *větší* množinu než v původní úloze:

$$\{\mathbf{x} \in [0, 1]^n \mid \mathbf{Ax} \geq \mathbf{b}\} \supseteq \{\mathbf{x} \in \{0, 1\}^n \mid \mathbf{Ax} \geq \mathbf{b}\}.$$

Pro některé konkrétní kombinatorické problémy však je situace příznivější, např. LP relaxace řeší původní úlohu vždy přesně (tedy nerovnost (11.22) platí s rovností), nebo dovoluje sestavit její přibližné řešení se zárukou přesnosti (dále uvedeme příklady). I když toto štěstí nemáme, dolní mez je často velmi užitečná při hledání přesného nebo přibližného řešení úlohy nějakým jiným způsobem (např. metodou *větví a mezí* nebo metodou *sečných nadrovin*).

11.6.1 Nejlepší přiřazení

V **přiřazovací úloze** (*assignment problem*, též znám jako *bipartitní párování*) jsou dána čísla c_{ij} pro $i, j \in \{1, \dots, n\}$ a cílem je najít bijektivní zobrazení $\pi: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ (tedy permutaci množiny $\{1, \dots, n\}$), která maximalizuje součet $\sum_{i=1}^n c_{i, \pi(i)}$.

Každé dvojici $i, j \in \{1, \dots, n\}$ přiřadíme binární proměnnou $x_{ij} \in \{0, 1\}$. Tyto proměnné reprezentují permutaci π tak, že $x_{ij} = 1$ právě když $\pi(i) = j$. Protože π přiřazuje každému i právě jedno j , musí být mezi proměnnými x_{i1}, \dots, x_{in} právě jedna rovna 1. To zapíšeme jako $\sum_{j=1}^n x_{ij} = 1$. Protože π je bijekce, každý obraz π má právě jeden vzor, tedy $\sum_{i=1}^n x_{ij} = 1$. Nyní úlohu můžeme psát jako ILP

$$\min \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \quad (11.23a)$$

$$\text{za podm. } \sum_{j=1}^n x_{ij} = 1, \quad i = 1, \dots, n \quad (11.23b)$$

$$\sum_{i=1}^n x_{ij} = 1, \quad j = 1, \dots, n \quad (11.23c)$$

$$x_{ij} \in \{0, 1\}, \quad i, j = 1, \dots, n \quad (11.23d)$$

LP relaxaci úlohy (11.23) získáme nahrazením omezení $x_{ij} \in \{0, 1\}$ omezeními $0 \leq x_{ij} \leq 1$. Vlastně stačí jen $x_{ij} \geq 0$, neboť (11.23b) a $x_{ij} \geq 0$ implikuje $x_{ij} \leq 1$. Takže LP relaxace je speciální případ dopravního problému (11.9).

Pro tuto úlohu máme štěstí, protože platí toto (důkaz neuvádíme, není krátký):

Věta 11.1. Úloha (11.23) má stejnou optimální hodnotu jako její LP relaxace a mezi optimálními řešeními LP relaxace je aspoň jedno celočíselné.

Vyslovme tuto větu trochu jinak. Zapišme všechny proměnné x_{ij} do matice \mathbf{X} velikosti $n \times n$. Účelová funkce (11.23a) je lineární funkce matice \mathbf{X} (jde psát jako $\mathbf{C} \cdot \mathbf{X}$, viz §4.7). Přípustná

řešení úlohy (11.23) jsou všechny matice, které mají v každém řádku a každém sloupci právě jednu jedničku. To jsou *permutační matice* (viz Příklad 4.6). Přípustná řešení LP relaxace jsou matice s nezápornými prvky, ve kterých součet prvků v každém řádku a v každém sloupci je roven 1. Takové matice se nazývají *dvojitě stochastické matice*. Takže Věta 11.1 říká, že minimum lineární funkce na množině permutačních matic je rovno minimu této lineární funkce na množině dvojitě stochastických matic. To je klasický výsledek *Birkhoffa a von Neumanna*.

11.6.2 Nejmenší vrcholové pokrytí

Vrcholové pokrytí neorientovaného grafu (V, E) je podmnožina $X \subseteq V$ vrcholů taková, že každá hrana má alespoň jeden vrchol v X . V úloze na **nejmenší vrcholové pokrytí** hledáme pro daný graf vrcholové pokrytí s nejmenším počtem vrcholů. Je to jedna z klasických NP-těžkých úloh kombinatorické optimalizace.

Každému vrcholu $i \in V$ přiřadíme proměnnou $x_i \in \{0, 1\}$ s tímto významem: $x_i = 1$ když $i \in X$ a $x_i = 0$ když $i \notin X$. Úlohu můžeme formulovat jako ILP

$$\min \sum_{i \in V} x_i \quad (11.24a)$$

$$\text{za podm. } x_i + x_j \geq 1, \quad \{i, j\} \in E \quad (11.24b)$$

$$x_i \in \{0, 1\}, \quad i \in V \quad (11.24c)$$

Označíme-li *incidenční matici* grafu (V, E) jako $\mathbf{A} \in \{0, 1\}^{n \times m}$ kde $n = |V|$ je počet vrcholů a $m = |E|$ je počet hran, úlohu (11.24) můžeme zapsat také jako

$$\min \{ \mathbf{1}^T \mathbf{x} \mid \mathbf{x} \in \{0, 1\}^n, \mathbf{A}^T \mathbf{x} \geq \mathbf{1} \}. \quad (11.25)$$

LP relaxace úlohy je lineární program

$$\min \sum_{i \in V} x_i \quad (11.26a)$$

$$\text{za podm. } x_i + x_j \geq 1, \quad \{i, j\} \in E \quad (11.26b)$$

$$0 \leq x_i \leq 1, \quad i \in V \quad (11.26c)$$

tedy v úloze (11.24) jsme nahradili omezení $x_i \in \{0, 1\}$ omezeními $x_i \in [0, 1]$ neboli $0 \leq x_i \leq 1$.

Necht' \mathbf{x} je optimální řešení relaxované úlohy (11.26). Zaokrouhleme to řešení nahoru, tedy položíme $\bar{x}_i = \lceil x_i \rceil$. Vektor $\bar{\mathbf{x}}$ je očividně přípustný pro úlohu (11.24). Platí toto:

Věta 11.2. Pro každý problém minimálního vrcholového pokrytí je hodnota $\sum_{i \in V} \bar{x}_i$ nejvýše dvakrát větší než optimální hodnota původní úlohy (11.24).

Důkaz. Pro každé číslo $x \geq 0$ platí $\lceil x \rceil \leq 2x$. Tedy $\bar{x}_i \leq 2x_i$. Dle (11.22) je $\text{OPT}_{\text{LP}} \leq \text{OPT}_{\text{ILP}}$, kde OPT_{LP} a OPT_{ILP} označuje optimální hodnotu úlohy (11.26) a (11.24). Dáno dohromady,

$$\sum_{i \in V} \bar{x}_i \leq \sum_{i \in V} 2x_i = 2 \cdot \text{OPT}_{\text{LP}} \leq 2 \cdot \text{OPT}_{\text{ILP}}. \quad \square$$

Příklad 11.8. Necht' $V = \{1, 2, 3\}$ a $E = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$, tedy (V, E) je úplný graf se třemi vrcholy. Úloha (11.24) má tři optimální řešení $\mathbf{x} = (1, 1, 0)$, $\mathbf{x} = (1, 0, 1)$ a $\mathbf{x} = (0, 1, 1)$, odpovídající minimálním pokrytím $X = \{1, 2\}$, $X = \{1, 3\}$ a $X = \{2, 3\}$. Optimální hodnota úlohy je 2. LP relaxace (11.26) má jediné optimální řešení $\mathbf{x} = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ s optimální hodnotou $\frac{3}{2}$, což je ve shodě s nerovností (11.22). Nahoru zaokrouhlené řešení je $\bar{\mathbf{x}} = (1, 1, 1)$ s optimální hodnotou 3, což je ve shodě s Větou 11.2. \square

11.6.3 Největší nezávislá množina

Podmnožina vrcholů $X \subseteq V$ neorientovaného grafu (V, E) se nazývá **nezávislá**, když žádné dva vrcholy z X nejsou spojeny hranou. V úloze na **největší nezávislou množinu** hledáme pro daný graf nezávislou množinu s největším počtem vrcholů.

Formulace této úlohy pomocí ILP je velmi podobné (11.24):

$$\max \sum_{i \in V} x_i \quad (11.27a)$$

$$\text{za podm. } x_i + x_j \leq 1, \quad \{i, j\} \in E \quad (11.27b)$$

$$x_i \in \{0, 1\}, \quad i \in V \quad (11.27c)$$

V LP relaxaci opět nahradíme podmínky $x_i \in \{0, 1\}$ nerovnostmi $0 \leq x_i \leq 1$.

Ukážeme, že tato LP relaxace není příliš užitečná. LP relaxace má vždy přípustné řešení, ve kterém $x_i = \frac{1}{2}$ pro všechna $i \in V$. Tomu odpovídá hodnota účelové funkce $\sum_i x_i = |V|/2$. Tedy optimální hodnota LP relaxace nemůže být menší než $|V|/2$. Bohužel, optimální hodnota původní úlohy (11.27) může libovolně krát menší než $|V|/2$. Např. pro každý úplný graf je velikost největší nezávislé množiny 1 (protože každé dva vrcholy jsou spojeny hranou).

Mohli bychom doufat, že šikovným zaokrouhlením optimálního řešení relaxované úlohy získáme analogii Věty 11.2. To je ale nemožné: je známo, že problém největší nezávislé množiny nelze aproximovat s jakoukoliv konstantní (tj. nezávislou na velikosti grafu) zárukou (je tzv. *APX-těžký*).

11.7 Cvičení

11.1. Najděte graficky množinu optimálních řešení úlohy

$$\begin{aligned} \min \quad & c_1x_1 + c_2x_2 + c_3x_3 \\ \text{za podm.} \quad & x_1 + x_2 \geq 1 \\ & x_1 + 2x_2 \leq 3 \\ & x_1 + x_2 \leq 10 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

pro následující případy: (a) $\mathbf{c} = (-1, 0, 1)$, (b) $\mathbf{c} = (0, 1, 0)$, (c) $\mathbf{c} = (0, 0, -1)$.

11.2. Následující úlohy nejprve převed'te na rovnicový tvar (tj. tvar s nezápornými proměnnými a omezeními typu lineární rovnice, viz §11.1). Potom je převed'te do maticové formy $\min\{\mathbf{r}^T \mathbf{u} \mid \mathbf{P}\mathbf{u} = \mathbf{q}, \mathbf{u} \geq \mathbf{0}\}$ (výsledkem tedy budou $\mathbf{u}, \mathbf{P}, \mathbf{q}, \mathbf{r}$).

$$\begin{aligned} \text{a)} \quad & \min \quad 2x_1 \quad \quad \quad - 3x_3 + x_4 \\ \text{za podmínek} \quad & x_1 - x_2 - x_3 \quad \quad \geq 0 \\ & -x_1 + 2x_2 - 3x_3 \quad \quad \leq 5 \\ & 2x_1 - x_2 - x_3 + 2x_4 = 6 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

b) lineární program (11.9)

c) lineární program (11.13)

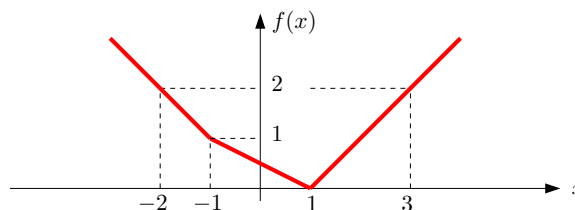
d) lineární program (11.15)

11.3. Vyřešte úvahou tyto jednoduché lineární programy a napište (jednoduchý) výraz pro optimální hodnotu. Odpovědi dokažte. Vektor $\mathbf{c} \in \mathbb{R}^n$ a číslo $k \in \{1, \dots, n\}$ jsou dány.

- a) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}$
- b) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}\}$
- c) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} = 1\}$
- d) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} \leq 1\}$
- e) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, -1 \leq \mathbf{1}^T \mathbf{x} \leq 1\}$
- f) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} = k\}$
- g) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{1}^T \mathbf{x} = k\}$
- h) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{1}^T \mathbf{x} \leq k\}$
- i) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, 0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq 1\}$
- j) $(\star) \max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, -\mathbf{y} \leq \mathbf{x} \leq \mathbf{y}, \mathbf{1}^T \mathbf{y} = k, \mathbf{y} \leq \mathbf{1}\}$

11.4. Pokuste se úlohy transformovat na LP. Pokud to nedokážete, vysvětlete proč.

- a) $\min\{|x_1| + |x_2| \mid x_1, x_2 \in \mathbb{R}, 2x_1 - x_2 \geq 1, -x_1 + 2x_2 \geq 1\}$
- b) $\max\{|x_1 - c_1| + \dots + |x_n - c_n| \mid x_1, \dots, x_n \in \mathbb{R}, a_1 x_1 + \dots + a_n x_n \geq b\}$
- c) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}, |\mathbf{d}^T \mathbf{x}| \leq 1, \mathbf{x} \geq \mathbf{0}\}$
- d) $\min_{\mathbf{x} \in \mathbb{R}^n} \sum_{l=1}^L \max_{k=1}^K (\mathbf{c}_{kl}^T \mathbf{x} + d_{kl})$
- e) $\min_{\mathbf{x} \in \mathbb{R}^n} \sum_{i=1}^m f(\mathbf{a}_i^T \mathbf{x} - b_i)$, kde funkce f je definována obrázkem



- f) $\min\{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_1 \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\|_\infty \leq 1\}$
- g) $\min\{\|\mathbf{x}\|_1 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}\}$
- h) $\min\{\|\mathbf{x}\|_1 \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_\infty \leq 1\}$
- i) $\min_{\mathbf{x} \in \mathbb{R}^n} (\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_1 + \|\mathbf{x}\|_\infty)$

11.5. Máme algoritmus (černou skříňku) na řešení LP, kterou můžeme zavolat i vícekrát. S pomocí tohoto algoritmu vyřešte úlohu $\max\{|\mathbf{c}^T \mathbf{x}| \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$.

11.6. Dokažte nebo vyvrát'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 $\mathbf{x} \in \mathbb{R}^n$.

- a) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| = 1\} = \max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| \leq 1\}$
- b) $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| = 1\} = \min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| \leq 1\}$
- c) $\max\{\|\mathbf{A}\mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| = 1\} = \max\{\|\mathbf{A}\mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| \leq 1\}$

11.7. Dokažte, že minimum funkce $f(x) = \sum_{i=1}^m |x - y_i|$ se nabývá v mediánu čísel y_1, \dots, y_m .

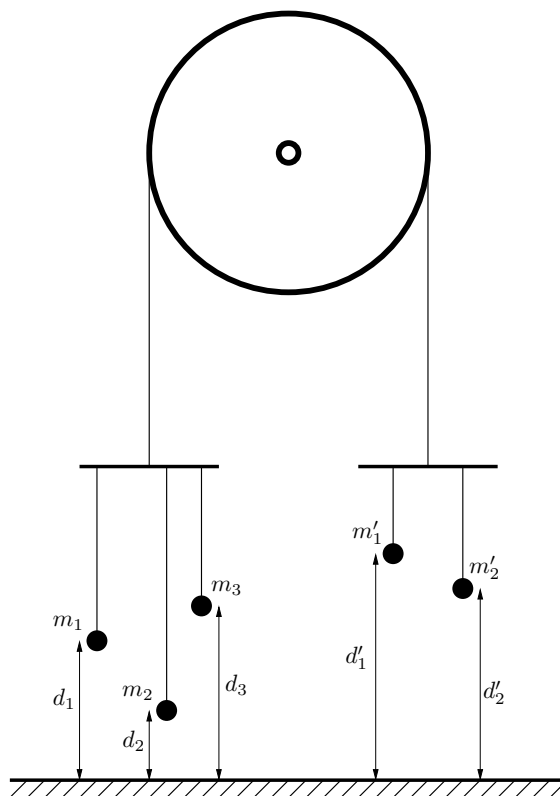
11.8. Uvažujme Příklad 11.4 takto pozměněný: pán má pomocníka, kterému platí 10 Kč za každý kg vyrobeného zboží (je jedno, zda to jsou lupínky nebo hranolky). Ovšem pokud se toho vyrobí hodně, chce pomocník větší plat, protože musí zůstat přesčas. Tak za každý kg nad 20 kg vyrobeného zboží si nechá připlatit 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.

- 11.9. Armáda má ve dvou skladech uskladněno 6 a 5 tun střeliva. Střelivo je nutno přepravit ke třem střelnicím s požadavky 3, 2 a 2 tuny střeliva. Kvůli minimalizaci rizika je nutné minimalizovat maximální množství střeliva, které se veze po kterékoliv cestě od skladu ke střelnici. Formulujte jako lineární program.

- 11.10. 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, 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 .
 - Formulujte hledání minima potenciální energie soustavy jako lineární program.
- 11.11. 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 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.

11.12. Spočtěte limitu $\lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \lim_{p \rightarrow \infty} (|x_1|^p + \dots + |x_n|^p)^{1/p}$.

11.13. Máme dvě konečné množiny $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$ a $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ bodů z \mathbb{R}^d . Najděte

- a) nadrovinu procházející počátkem (tj. lineární podprostor dimenze $d - 1$),
- b) nadrovinu (tj. afinní podprostor dimenze $d - 1$),

která množiny odděluje (tedy všechny body z první množiny jsou na její jedné straně a všechny body z druhé množiny jsou na její druhé straně).

11.14. Rozhodněte (a odpověď dokažte), pro jaká n je následující funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ norma:

- a) $f(\mathbf{x}) = |\max\{x_1, \dots, x_n\}|$
- b) $f(\mathbf{x}) = \|\mathbf{x}\|_1 + \|\mathbf{x}\|_2$
- c) $f(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\|_1 + \|\mathbf{y}\|_2$ (argument této funkce je vektor $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n$, kde $\mathbf{x} \in \mathbb{R}^k$ a $\mathbf{y} \in \mathbb{R}^{n-k}$)

11.15. (★) Najděte co nejjednodušší algoritmus, který spočte minimum funkce $\mathbf{a}^T \mathbf{X} \mathbf{b}$ na množině dvojité stochastických matic \mathbf{X} , kde \mathbf{a}, \mathbf{b} jsou dané vektory.

Nápověda a řešení

11.2.b) Úloha už je v rovnicovém tvaru, takže v prvním kroku nemusíme dělat nic. Úloha jde napsat v maticovém tvaru (11.10), to ale není požadovaný tvar, protože proměnné jsou soustředěné do matice \mathbf{X} a nikoliv do vektoru. Požadovaný tvar získáme volbou

$$\mathbf{u} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \in \mathbb{R}^{mn}, \quad \mathbf{P} = \begin{bmatrix} \mathbf{I}_n & \cdots & \mathbf{I}_n \\ \mathbf{1}_m^T & & \\ & \ddots & \\ & & \mathbf{1}_m^T \end{bmatrix} \in \mathbb{R}^{(m+n) \times (mn)}, \quad \mathbf{q} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \in \mathbb{R}^{m+n}, \quad \mathbf{r} = \begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_n \end{bmatrix} \in \mathbb{R}^{mn}$$

kde $\mathbf{x}_1, \dots, \mathbf{x}_n$ jsou sloupce matice $\mathbf{X} = [x_{ij}]$ a $\mathbf{c}_1, \dots, \mathbf{c}_n$ jsou sloupce matice $\mathbf{C} = [c_{ij}]$.

11.3.a) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\} = \sum_{i=1}^n \max\{0, c_i\}$, tedy optimální hodnota je součet kladných čísel c_i .

Pro důkaz si stačí uvědomit, že optimum se nabývá pro vektor \mathbf{x} takový, 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 $c_i > 0$.

11.3.b) $\sum_{i=1}^n |c_i|$. Dokáže se podobně.

11.3.c) $\max_{i=1}^n c_i$

11.3.d) $\max_{i=1}^n \max\{0, c_i\} = \max\{0, \max_{i=1}^n c_i\}$

11.3.e) Když $c_i = a$ pro každé i (tj. všechna c_i jsou stejná), tak optimální hodnota je $|a|$. Jinak je úloha neomezená.

11.3.i) Nápověda: substituuje $y_i = x_i - x_{i-1}$

11.4.a) $\min\{z_1 + z_2 \mid x_1, x_2, z_1, z_2 \in \mathbb{R}, 2x_1 - x_2 \geq 1, -x_1 + 2x_2 \geq 1, x_1 \leq z_1, x_2 \leq z_2, -x_1 \leq z_1, -x_2 \leq z_2\}$

11.4.b) Nejde.

11.4.c) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A} \mathbf{x} = \mathbf{b}, \mathbf{d}^T \mathbf{x} \leq 1, -\mathbf{d}^T \mathbf{x} \leq 1, \mathbf{x} \geq \mathbf{0}\}$

11.4.d) $\min\{\mathbf{1}^T \mathbf{z} \mid \mathbf{c}_{kl}^T \mathbf{x} + d_{kl} \leq z_l \ (\forall k, l), -\mathbf{c}_{kl}^T \mathbf{x} - d_{kl} \leq z_l \ (\forall k, l), \mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{R}^L\}$ (analogické §11.1.1)

- 11.4.f) $\min\{\mathbf{1}^T \mathbf{z} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{R}^m, -\mathbf{z} \leq \mathbf{Ax} - \mathbf{b} \leq \mathbf{z}, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}\}$
- 11.4.i) $\min\{\mathbf{1}^T \mathbf{y} + z \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m, z \in \mathbb{R}, -\mathbf{z} \leq \mathbf{Ax} - \mathbf{b} \leq \mathbf{z}, -z\mathbf{1} \leq \mathbf{x} \leq z\mathbf{1}\}$
- 11.5. 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 $A = \max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ a $B = \min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$.
- 11.6. Inspirujte se úvahou v §11.1.1.
- 11.7. Funkce f je po částech afinní, je tedy diferencovatelná všude kromě bodů y_i . Seřad'te si čísla y_1, \dots, y_m a najděte hodnotu derivace funkce f pro každý interval (y_{i-1}, y_i) . Z toho usud'te, kde je funkce klesající/rostoucí/konstantní a kde nabývá minimum.
- 11.8. Stejná úloha jako v Příkladu 11.4, jen účelová funkce se změnila na $120l + 76h - f(l + h)$ kde $f(t) = \max\{10t, 200 + 20(t - 20), 400 + 40(t - 30)\}$. To převedeme na LP dle §11.1.1.
- 11.9. Necht' x_{ij} označuje množství střeliva vezeného od skladu i do střelnice j . Minimalizujeme funkci $\max\{x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23}\}$ za podmínek $x_{11} + x_{12} + x_{13} \leq 6$, $x_{21} + x_{22} + x_{23} \leq 5$, $x_{11} + x_{21} = 3$, $x_{12} + x_{22} = 2$, $x_{13} + x_{23} = 2$, a $x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23} \geq 0$. To převedeme na LP dle §11.1.1.
- 11.10.a) $E(x) = \sum_{i=1}^n m_i \max(d_i + x, 0) + \sum_{i=1}^{n'} m'_i \max(d'_i - x, 0)$
- 11.10.b) $\min\{\sum_{i=1}^n m_i z_i + \sum_{i=1}^{n'} m'_i z'_i \mid x, z_i, z'_i \in \mathbb{R}, z_i \geq d_i + x, z'_i \geq d'_i - x, z_i \geq 0, z'_i \geq 0\}$
- 11.12. Stačí spočítat $\lim_{p \rightarrow \infty} (x_1^p + \dots + x_n^p)^{1/p}$ pro $x_1, \dots, x_n \geq 0$. Označte $a = \max\{x_1, \dots, x_n\}$ a pište $\lim_{p \rightarrow \infty} (x_1^p + \dots + x_n^p)^{1/p} = a \lim_{p \rightarrow \infty} ((x_1/a)^p + \dots + (x_n/a)^p)^{1/p}$. Zbytek je snad jasný.
- 11.13.a) Viz Příklad 16.10.
- 11.14.a) Je to norma pro $n = 1$, protože pak $f(x) = |x|$. Pro $n \geq 2$ to norma není, protože z $f(\mathbf{x}) = 0$ neplyne $\mathbf{x} = \mathbf{0}$ (první axiom), např. pro $\mathbf{x} = (-1, 0, 0, \dots)$.
- 11.15. Dle Věty 11.1 se optimální hodnota úlohy nezmění, když budeme minimalizovat přes permutační matice. Tedy hledáme minimum funkce $\sum_i a_i b_{\pi(i)}$ přes všechny permutace π .

Kapitola 12

Konvexní množiny a mnohostěny

12.1 Konvexní množiny

Množina $X \subseteq \mathbb{R}^n$ se nazývá **konvexní**, jestliže

$$\mathbf{x} \in X, \mathbf{y} \in X, 0 \leq \alpha \leq 1 \implies (1 - \alpha)\mathbf{x} + \alpha\mathbf{y} \in X. \quad (12.1)$$

Množina $\{(1 - \alpha)\mathbf{x} + \alpha\mathbf{y} \mid 0 \leq \alpha \leq 1\}$ 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 :



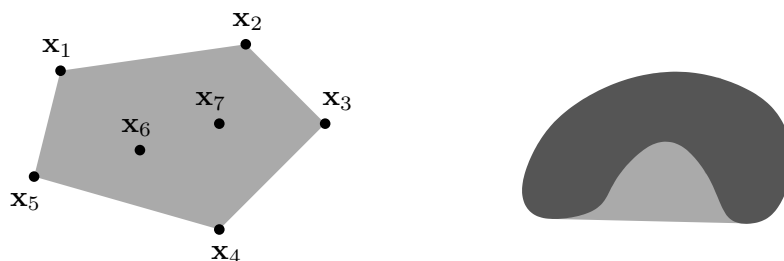
Konvexní kombinace vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ je jejich lineární kombinace $\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k$ taková, že $\alpha_1 + \dots + \alpha_k = 1$ a $\alpha_1, \dots, \alpha_k \geq 0$. Lze dokázat, že množina je konvexní právě tehdy, když je uzavřená vůči konvexním kombinacím (neboli každá konvexní kombinace vektorů z množiny leží v množině). Všimněte si, že $(1 - \alpha)\mathbf{x} + \alpha\mathbf{y}$ pro $0 \leq \alpha \leq 1$ je konvexní kombinací dvou vektorů \mathbf{x}, \mathbf{y} , neboť $(1 - \alpha) + \alpha = 1$, $1 - \alpha \geq 0$, $\alpha \geq 0$.

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. na 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 množiny** $X \subseteq \mathbb{R}^n$ (konečné či nekonečné) definujeme jako průnik všech konvexních množin, které množinu X obsahují. Značíme jej $\text{conv } X$.

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 $(1 - \alpha)\mathbf{x} + \alpha\mathbf{y}$ také v X i Y , tedy je v $X \cap Y$. \square

Sjednocení konvexních množin ale nemusí být konvexní množina.

12.2 Čtyři druhy kombinace vektorů

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 §3.3). 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 (konečné či nekonečné) 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

12.3 Konvexní mnohostěny

Zopakujme (viz §3.3), že *nadrovina* v \mathbb{R}^n je afinní podprostor dimenze $n - 1$. Je to tedy množina

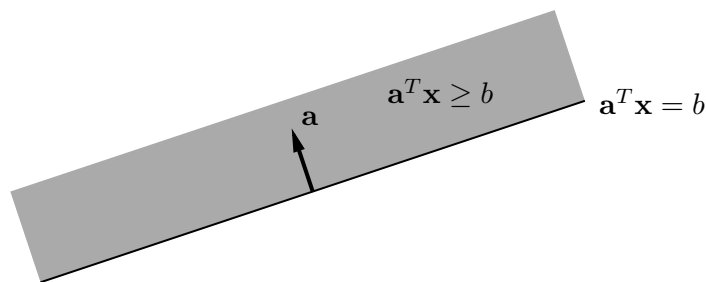
$$\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}, \quad (12.3)$$

kde $\mathbf{a} \in \mathbb{R}^n$ ($\mathbf{a} \neq \mathbf{0}$) je normála nadroviny a $b \in \mathbb{R}$. **Uzavřený¹ poloprostor** (dále jen **poloprostor**) je množina

$$\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} \geq b\}. \quad (12.4)$$

Hranice (viz §8.9) poloprostoru je nadrovina (12.3). Obrázek ilustruje tyto pojmy pro $n = 2$:

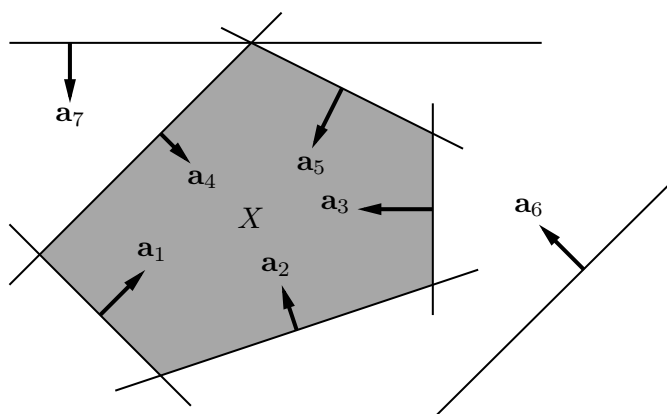
¹Otevřený poloprostor je množina $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} > b\}$.



Konvexní mnohostěn (krátce jen **mnohostěn**, angl. *polyhedron*) je průnik konečně mnoha uzavřených poloprostorů. Je to tedy množina

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T \mathbf{x} \geq b_i \quad \forall i = 1, \dots, m \} = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} \geq \mathbf{b} \}, \quad (12.5)$$

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 $\mathbf{b} \in \mathbb{R}^m$. Obrázek ukazuje příklad pro $n = 2$ a $m = 7$ (všimněte si, že poloprostory 6 a 7 jsou nadbytečné, jejich vynecháním se mnohostěn nezmění):



Poloprostor je očividně konvexní množina, proto dle Věty 12.1 je konvexní mnohostěn konvexní množina. Všimněte si, že konvexní mnohostěn nemusí být omezený.

Dimenze mnohostěnu je dimenze jeho afinního obalu, $\dim X = \dim \text{aff } X$ (afinní obal viz §12.2, dimenze afinního podprostoru viz §3.3). Jinými slovy, je to největší počet afinně nezávislých bodů, které lze do mnohostěnu umístit, minus jedna.

Příklad 12.2. Množina X z Příkladu 11.1 je konvexní mnohostěn. □

Příklad 12.3. Příklady ‘jednoduchých’ konvexních mnohostěnu v \mathbb{R}^n :

- prázdná množina \emptyset
- celý prostor \mathbb{R}^n
- intervaly typu $(-\infty, a]$, $[a, \infty)$ (tj. poloprostory v \mathbb{R}) a jejich průniky $[a, b]$
- každý afinní podprostor (např. bod, přímka, rovina, nadrovina)
- polopřímka $\{ \mathbf{x} + \alpha \mathbf{v} \mid \alpha \geq 0 \}$
- poloprostor
- množina \mathbb{R}_+^n (kde \mathbb{R}_+ značí množinu nezáporných reálných čísel)
- panel $\{ \mathbf{x} \in \mathbb{R}^n \mid b_1 \leq \mathbf{a}^T \mathbf{x} \leq b_2 \}$

- hyperkrychle $[-1, 1]^n = \{ \mathbf{x} \in \mathbb{R}^n \mid -1 \leq x_i \leq 1 \ \forall i = 1, \dots, n \} = \{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_\infty \leq 1 \}$
- hyperkvádr ('box') $[a_1, b_1] \times \dots \times [a_n, b_n] = \{ \mathbf{x} \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i \ \forall i = 1, \dots, n \}$
- 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)
- hyper-osmistěn (také zvaný křížový polytop) $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i| \leq 1 \}$. \square

Příklad 12.4. Příklady konvexních množin, které nejsou mnohostěny:

- Koule v \mathbb{R}^n pro $n \geq 2$ je průnikem nekonečně mnoha poloprostorů $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} \leq 1 \}$ pro všechna $\mathbf{a} \in \mathbb{R}^n$ splňující $\|\mathbf{a}\|_2 = 1$ (nakreslete a promyslete!). Není to mnohostěn, protože počet poloprostorů není konečný.
- otevřený poloprostor $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} > b \}$
- interval $[0, 1)$ \square

12.3.1 (★) Extrémní body

Bod $\mathbf{x} \in X$ se nazývá **extrémní bod** konvexní množiny X , jestliže neexistují dva různé body $\mathbf{z} \in X$ takové, že \mathbf{x} je střed úsečky spojující tyto dva body, tj. jestliže platí implikace

$$\mathbf{x}_1, \mathbf{x}_2 \in X, \quad \mathbf{x} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2) \implies \mathbf{x}_1 = \mathbf{x}_2. \quad (12.6)$$

Pro mnohostěn (12.5) a pro množinu $I \subseteq \{1, \dots, m\}$ budeme symbolem \mathbf{A}_I označovat² matici tvořenou řádky matice \mathbf{A} s indexy I a \mathbf{b}_I bude označovat vektor tvořený složkami vektoru \mathbf{b} s indexy I . Tedy $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ je jen jiné označení soustavy $\mathbf{a}_i^T \mathbf{x} = b_i, i \in I$.

Věta 12.2. Pro každý bod \mathbf{x} mnohostěnu (12.5) jsou následující tvrzení ekvivalentní:

- Bod \mathbf{x} je extrémní bod mnohostěnu.
- Existuje $I \subseteq \{1, \dots, m\}$ tak, že $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ a sloupce matice \mathbf{A}_I jsou lineárně nezávislé.

Důkaz. (★) Necht' X označuje mnohostěn (12.5) a $\mathbf{x} \in X$ je jeho extrémní bod. Označme jako $I = \{ i \mid \mathbf{a}_i^T \mathbf{x} = b_i \}$ množinu indexů nerovností v definici (12.5), které jsou v bodě \mathbf{x} aktivní. Dokážeme, že sloupce \mathbf{A}_I jsou lineárně nezávislé. Kdyby totiž nebyly, měla by matice \mathbf{A}_I netriviální nulový prostor (dle Věty 3.7), tedy $\mathbf{A}_I \mathbf{v} = \mathbf{0}$ pro nějaké $\mathbf{v} \neq \mathbf{0}$. Pro každé $i \notin I$ je $\mathbf{a}_i^T \mathbf{x} > b_i$, tedy $\mathbf{a}_i^T \mathbf{x} \pm \varepsilon \geq b_i$ pro nějaké $\varepsilon > 0$. Tedy by bylo $\mathbf{A}(\mathbf{x} \pm \varepsilon \mathbf{v}) \geq \mathbf{b}$ neboli $\mathbf{x} \pm \varepsilon \mathbf{v} \in X$. Dle (12.6) by proto \mathbf{x} nebyl extrémní bod.

Necht' I je taková, že sloupce \mathbf{A}_I jsou lineárně nezávislé. Necht' $\mathbf{x} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2)$ pro nějaké $\mathbf{x}_1, \mathbf{x}_2 \in X$. Pak

$$\begin{aligned} \mathbf{A}_I \mathbf{x}_1 &\geq \mathbf{b}_I, \\ \mathbf{A}_I \mathbf{x}_2 &\geq \mathbf{b}_I, \\ \mathbf{A}_I \mathbf{x} &= \frac{1}{2}(\mathbf{A}_I \mathbf{x}_1 + \mathbf{A}_I \mathbf{x}_2) = \mathbf{b}_I. \end{aligned}$$

²K tomu předpokládáme, že množina I je uspořádaná.

Z toho plyne $\mathbf{A}_I \mathbf{x}_1 \geq \frac{1}{2}(\mathbf{A}_I \mathbf{x}_1 + \mathbf{A}_I \mathbf{x}_2) \leq \mathbf{A}_I \mathbf{x}_2$. Ale pro libovolná čísla $\alpha_1, \alpha_2 \in \mathbb{R}$ platí implikace

$$\alpha_1 \geq \frac{1}{2}(\alpha_1 + \alpha_2) \leq \alpha_2 \implies \alpha_1 = \alpha_2. \quad (12.7)$$

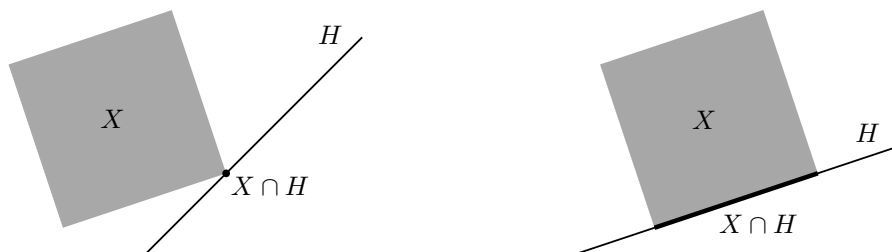
Tedy platí $\mathbf{A}_I \mathbf{x}_1 = \mathbf{A}_I \mathbf{x}_2$. Protože \mathbf{A}_I má lineárně nezávislé sloupce, plyne z toho $\mathbf{x}_1 = \mathbf{x}_2$ (dle Věty 3.1). Tedy \mathbf{x} je extrémní bod. \square

Připomeňme, že lineární soustava má právě jedno řešení právě tehdy, když má aspoň jedno řešení a její matice má lineárně nezávislé sloupce. Věta tedy vlastně říká, že $\mathbf{x} \in \mathbb{R}^n$ je extrémní bod mnohostěnu, právě když soustava $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ má právě jedno řešení \mathbf{x} a toto řešení navíc splňuje $\mathbf{A} \mathbf{x} \geq \mathbf{b}$. Geometricky to znamená, že hranice poloprostorů s indexy I se protínají v právě jednom bodě a ten bod leží v mnohostěnu. To nám dává návod, jak vypsat všechny extrémní body mnohostěnu. Procházíme všechny podmnožiny $I \subseteq \{1, \dots, m\}$. Jestliže soustava $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ má právě jedno řešení \mathbf{x} a to navíc splňuje $\mathbf{A} \mathbf{x} \geq \mathbf{b}$, pak \mathbf{x} je extrémní bod.

Protože podmnožin I je konečně mnoho, má každý konvexní mnohostěn konečně mnoho extrémních bodů. Mnohostěn popsaný malým množstvím lineárních nerovnic ovšem může mít extrémních bodů velmi mnoho. Tomu se není co divit, protože počet všech podmnožin množiny $\{1, \dots, m\}$ je 2^m . Např. hyperkrychle $\{\mathbf{x} \in \mathbb{R}^n \mid -1 \leq \mathbf{x} \leq 1\}$ je popsána $2n$ lineárními nerovnicemi ale má 2^n extrémních bodů. Není tomu tak vždy, např. simplex (viz Příklad 12.3) je popsán $n + 1$ lineárními nerovnicemi a má $n + 1$ extrémních bodů.

12.3.2 (★) Extrémy lineární funkce na mnohostěnu

Opěrná nadrovina konvexní množiny $X \subseteq \mathbb{R}^n$ je nadrovina $H = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}$ (kde $\mathbf{a} \neq \mathbf{0}$) taková, že $X \cap H \neq \emptyset$ a pro všechna $\mathbf{x} \in X$ platí $\mathbf{a}^T \mathbf{x} \geq b$. Neboli je to nadrovina, která obsahuje množinu X v jednom ze svých dvou poloprostorů a má s množinou X neprázdný průnik. Obrázky ukazují dvě různé opěrné nadroviny mnohostěnu:



Je-li H opěrná rovina mnohostěnu X , pak množina $X \cap H$ se nazývá **stěna** mnohostěnu. Dle Věty 12.1 je každá stěna konvexního mnohostěnu sama o sobě konvexní mnohostěn. Stěny některých dimenzí mají jméno:

- Stěna dimenze 0 se nazývá **vrchol** (na obrázku vlevo).
- Stěna dimenze 1 se nazývá **hrana** (na obrázku vpravo).
- Stěna dimenze $\dim X - 1$ se nazývá **faseta** (pro množinu X dimenze 2 na obrázku je tedy faseta totéž co hrana).

Lze dokázat (důkaz neuvádíme), že bod mnohostěnu je vrchol, právě když je to extrémní bod.

Každým hraničním bodem (viz §8.9) konvexního mnohostěnu prochází alespoň jedna opěrná rovina. To snadno dokážeme: Mnohostěn je průnik konečně mnoha poloprostorů. Hranice každého z těchto poloprostorů je opěrná nadrovina mnohostěnu. Zároveň každý hraniční bod mnohostěnu leží na hranici aspoň jednoho poloprostoru.

Věta 12.3. Je-li H opěrná nadrovina konvexního mnohostěnu X , pak každý extrémní bod mnohostěnu $X \cap H$ je extrémní bod mnohostěnu X .

Věta tedy říká, že extrémní bod stěny mnohostěnu je zároveň extrémní bod mnohostěnu.

Důkaz. (*) Necht' $H = \{ \mathbf{x} \mid \mathbf{a}^T \mathbf{x} = b \}$, tedy $\mathbf{a}^T \mathbf{x} \geq b$ pro všechna $\mathbf{x} \in X$. Necht' $\mathbf{x} \in X \cap H$ není extrémní bod mnohostěnu X , takže $\mathbf{x} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2)$ pro nějaké $\mathbf{x}_1, \mathbf{x}_2 \in X$ kde $\mathbf{x}_1 \neq \mathbf{x}_2$. Platí

$$\begin{aligned}\mathbf{a}^T \mathbf{x}_1 &\geq b, \\ \mathbf{a}^T \mathbf{x}_2 &\geq b, \\ \mathbf{a}^T \mathbf{x} &= \frac{1}{2}(\mathbf{a}^T \mathbf{x}_1 + \mathbf{a}^T \mathbf{x}_2) = b.\end{aligned}$$

Díky (12.7) z toho plyne $\mathbf{a}^T \mathbf{x}_1 = \mathbf{a}^T \mathbf{x}_2 = b$, tedy $\mathbf{x}_1, \mathbf{x}_2 \in H$. Tedy \mathbf{x} není extrémní bod mnohostěnu $X \cap H$. \square

Ne každý mnohostěn má extrémní bod, např. poloprostor nebo afinní podprostor (což je konvexní mnohostěn) nemá žádný. Pro následující větu připomeňme, že *přímka* je afinní podprostor dimenze 1, tedy množina $\{ \mathbf{x} + \alpha \mathbf{v} \mid \alpha \in \mathbb{R} \}$ pro nějaké $\mathbf{x} \in \mathbb{R}^n$ a $\mathbf{0} \neq \mathbf{v} \in \mathbb{R}^n$.

Věta 12.4. Pro každý neprázdný konvexní mnohostěn jsou následující tvrzení ekvivalentní:

- Mnohostěn má aspoň jeden extrémní bod.
- Mnohostěn neobsahuje přímku.

Důkaz. (*) Nejprve dokážeme, že když mnohostěn X obsahuje přímku, tak nemá ani jeden extrémní bod. Víme, že existují $\mathbf{x} \in X$ a $\mathbf{v} \neq \mathbf{0}$ takové, že $\mathbf{x} + \alpha \mathbf{v} \in X$ pro všechna $\alpha \in \mathbb{R}$. Je-li X ve tvaru (12.5), platí $\mathbf{A}\mathbf{x} \geq \mathbf{b}$ a tedy $\mathbf{A}(\mathbf{x} + \alpha \mathbf{v}) \geq \mathbf{b}$ pro všechna $\alpha \in \mathbb{R}$. Z toho plyne $\alpha \mathbf{A}\mathbf{v} \geq \mathbf{0}$ pro $\alpha \in \mathbb{R}$, z čehož plyne $\mathbf{A}\mathbf{v} = \mathbf{0}$. To ale znamená, že pro každé $\mathbf{x} \in X$ a každé $\alpha \in \mathbb{R}$ platí $\mathbf{A}(\mathbf{x} + \alpha \mathbf{v}) \geq \mathbf{b}$ neboli $\mathbf{x} + \alpha \mathbf{v} \in X$. Tedy každým bodem mnohostěnu prochází přímka, která leží v mnohostěnu. Proto žádný bod mnohostěnu nemůže být extrémní bod.

Nyní dokážeme, že jestliže mnohostěn neobsahuje přímku, pak má aspoň jeden extrémní bod. Dokážeme to indukcí podle dimenze mnohostěnu. Tvrzení triviálně platí pro mnohostěny dimenze 0, což jsou pouhé body. Předpokládejme, že tvrzení platí pro všechny mnohostěny dimenze menší než n . Pokud neprázdný mnohostěn X dimenze n neobsahuje přímku, pak každá přímka protínající X narazí na hranici X . Jestliže vezmeme přímku, která leží v afinním obalu X , pak v obdrženém hraničním bodě mnohostěnu X existuje opěrná nadrovina H , která mnohostěn neobsahuje (tj. $X \not\subseteq H$). Mnohostěn $X \cap H$ má proto dimenzi menší než n a neobsahuje přímku (protože X ji neobsahuje), tedy podle indukčního předpokladu má aspoň jeden extrémní bod. Ten je dle Věty 12.3 extrémní bod mnohostěnu X . \square

Věta 12.5. Mějme konvexní mnohostěn, který neobsahuje přímku. Jestliže lineární funkce má na tomto mnohostěnu minimum, pak tato funkce nabývá na mnohostěnu minima aspoň v jednom z jeho extrémních bodů.

Důkaz. Necht' $\mathbf{x}^* \in X$ je bod mnohostěnu X , ve kterém lineární funkce $f(\mathbf{x}) = \mathbf{c}^T \mathbf{x}$ nabývá minima, tedy $\mathbf{c}^T \mathbf{x}^* \leq \mathbf{c}^T \mathbf{x}$ pro všechna $\mathbf{x} \in X$. To znamená, že $H = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{c}^T \mathbf{x} = \mathbf{c}^T \mathbf{x}^* \}$ je opěrná rovina mnohostěnu X v bodě \mathbf{x}^* (H je vlastně vrstevnice funkce f výšky $\mathbf{c}^T \mathbf{x}^*$). Tedy f nabývá na X minima ve všech bodech množiny $X \cap H$ (což je stěna mnohostěnu X). Mnohostěn $X \cap H$ neobsahuje přímku (protože X ji neobsahuje), tedy dle Věty 12.4 má aspoň jeden extrémní bod. Dle Věty 12.3 je tento bod extrémní bod mnohostěnu X . \square

Hledejme nyní minimum lineární funkce na mnohostěnu, neboli řešme lineární program

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} \geq \mathbf{b}\}.$$

Jestliže mnohostěn neobsahuje přímku a funkce na něm má minimum, pak dle Věty 12.5 stačí jednoduše projít všechny extrémní body mnohostěnu a vybrat z nich ten, ve kterém má funkce nejmenší hodnotu. Tento algoritmus má několik zádrhelů. Nevíme, co dělat s mnohostěny, které obsahují přímku. Nevíme, jak poznat, že účelová funkce má na mnohostěnu minimum. Zdaleka největší zádrhel je ovšem to, že prakticky není možné projít všechny extrémní body, neboť jich je příliš mnoho. V příští kapitole popíšeme algoritmus, který téměř všechny tyto potíže řeší.

12.4 Cvičení

12.1. Odpovězte, zda následující množiny jsou konvexní a odpověď dokažte z definice konvexní množiny:

- a) interval $[a, b] \subseteq \mathbb{R}$, kde $a \leq b$
- b) $\{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$
- c) $\{(x, y) \in \mathbb{R}^2 \mid y \geq x^2\}$
- d) $\{(x, y) \in \mathbb{R}^2 \mid y^2 \geq x^2\}$
- e) $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \leq \mathbf{b}, \mathbf{Cx} = \mathbf{d}\}$
- f) $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{x} \leq 1\}$
- g) \mathbb{Z} (množina celých čísel)
- h) $\{\mathbf{x} \in \mathbb{R}^n \mid \max\{x_1, \dots, x_n\} \geq 0\}$

12.2. Jsou následující množiny konvexní? Odpověď nemusíte dokazovat z definice konvexní množiny, stačí uvést přesvědčivý argument. Jestliže množina není konvexní, napište její konvexní obal jako množinu řešení soustavy (co možná nejjednodušší) rovnic a nerovnic.

- a) $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}, \sum_{i=1}^n x_i = 1\}$
- b) $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 = 1\}$
- c) $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 < 1\}$
- d) $\{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$
- e) $\{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, xy = 1\}$
- f) $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 2\} \cap \{(x, y) \in \mathbb{R}^2 \mid (x-1)^2 + y^2 \leq 2\}$
- g) $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1, x \geq 0, y \geq 0\}$
- h) $\{-1, 0, 1\}$
- i) $\{(1, 1), (2, 2)\}$
- j) $\{(1, 1), (1, 2), (3, 1)\}$

12.3. (★) Operace konvexního obalu je speciální případ operace zvané *uzávěr*, což je každá operace splňující tyto axiomy:

- $X \subseteq \text{conv } X$ (extensivita)
- $\text{conv conv } X = \text{conv } X$ (idempotence)
- $X \subseteq Y \implies \text{conv } X \subseteq \text{conv } Y$ (monotonicita)

Dokažte tyto vztahy.

- 12.4. (★) Navrhnete algoritmus na nalezení konvexního obalu konečné množiny bodů v rovině. Výstupem bude uspořádný seznam bodů, které tvoří vrcholy konvexního obalu.
- 12.5. Nakreslete lineární, afinní, nezáporný a konvexní obal náhodně zvolených k vektorů v \mathbb{R}^n pro všech devět případů $k, n \in \{1, 2, 3\}$.
- 12.6. Bude Věta 12.1 platit, pokud v ní výraz ‘konvexní množina’ nahradíme výrazem ‘lineární podprostor’ (příp. ‘afinní podprostor’, ‘konvexní kužel’)? Odpověď dokažte.
- 12.7. Jsou následující množiny konvexní mnohostěny? Zápornou odpověď odůvodněte přesvědčivým argumentem. Kladnou odpověď dokažte tak, že množinu napíšete jako množinu řešení soustavy lineárních nerovnic (tj. jako průnik poloprostorů).

- a) $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{C} \mathbf{x} \leq 1 \}$, kde \mathbf{C} je pozitivně definitní
- b) $\{ \alpha \mathbf{v} \mid \alpha \in \mathbb{R} \}$, kde $\mathbf{v} \in \mathbb{R}^n$
- c) (★) $\{ \mathbf{C} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\|_2 \leq 1 \}$, kde $\mathbf{C} \in \mathbb{R}^{m \times n}$
- 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

- 12.8. (★) Které z následujících mnohostěnů mají aspoň jeden extrémní bod? Pokud mnohostěn nemá extrémní bod, najděte přímku, která v něm leží.
- a) všechny mnohostěny z Příkladu 12.3
- b) $\{ (x, y, z) \in \mathbb{R}^3 \mid x + y \leq 1, x \geq 0, y \geq 0 \}$
- c) $\{ (x, y) \in \mathbb{R}^2 \mid -1 \leq x + y \leq 1 \}$

- 12.9. (★) Chceme vypsat všechny extrémní body mnohostěnu $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} \geq \mathbf{b} \}$.

- a) Vypište extrémní body pro

$$\mathbf{A} = \begin{bmatrix} -1 & -2 \\ 3 & -1 \\ -1 & 1 \\ 2 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -14 \\ 0 \\ -2 \\ 4 \end{bmatrix}.$$

- b) Napište program v Matlabu, který vypíše všechny extrémní body pro libovolné \mathbf{A}, \mathbf{b} .

- 12.10. (★) Pro mnohostěn $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} \geq \mathbf{b} \}$ dokažte:

- a) Má-li \mathbf{A} lineárně závislé sloupce (tj. $\text{rank } \mathbf{A} < n$), mnohostěn nemá extrémní bod.
- b) Má-li mnohostěn aspoň jeden extrémní bod, je $\text{rank } \mathbf{A} = n$.

- 12.11. Mějme vektory $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$. Pro každé $i = 1, \dots, m$ definujeme množinu

$$X_i = \{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}_i\|_2 \leq \|\mathbf{x} - \mathbf{a}_j\|_2 \ \forall j \neq i \}.$$

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 $\mathbf{a}_1, \dots, \mathbf{a}_4$.

- 12.12. Hledáme největší (hyper)kouli $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}\|_2 \leq r \}$, která se vejde do mnohostěnu $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} \leq \mathbf{b} \}$. Zformulujte úlohu jako lineární program. Optimalizujeme přes proměnné $\mathbf{a} \in \mathbb{R}^n$ a $r > 0$.

Nápověda a řešení

- 12.1.a) Konvexní, protože pro libovolné $\alpha \in [0, 1]$ je $(1 - \alpha)a + \alpha b \in [a, b]$.
- 12.1.b) Není konvexní. Např. $\mathbf{x} = (1, 1)$ a $\mathbf{y} = (-1, 1)$ patří do množiny, ale $\mathbf{x}/2 + \mathbf{y}/2 = (0, 1)$ nepatří.
- 12.1.c) Je konvexní.
- 12.1.e) Konvexní.
- 12.1.f) Konvexní.
- 12.1.g) Nekonvexní. Např. pro $x = 1, y = 2, \alpha = \frac{1}{2}$ číslo $(1 - \alpha)x + \alpha y = 1.5$ není celé.
- 12.1.h) Nekonvexní.
- 12.2.a) průnik poloprostorů a nadroviny, konvexní mnohostěn
- 12.2.b) Sféra, nekonvexní. Konv. obal je koule $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 \leq 1\}$.
- 12.2.c) koule bez hranice, konvexní
- 12.2.d) Graf paraboly, nekonvexní. Konvexní obal je $\{(x, y) \in \mathbb{R}^2 \mid y \geq x^2\}$.
- 12.2.e) Jedna větev grafu hyperboly, není konvexní. Konv. obal je $\{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, xy \geq 1\}$.
- 12.2.f) průnik dvou koulí, konvexní
- 12.2.g) Čtvrt kružnice, nekonvexní. Konv. obal je $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1, x + y \geq 1\}$.
- 12.2.h) Tři body v \mathbb{R} , nekonvexní protože např. $(-1 + 0)/2$ nepatří do množiny. Konv. obal je interval $[-1, 1]$.
- 12.2.i) Dva body v \mathbb{R}^2 , nekonvexní. Konv. obal je úsečka spojující ty dva body. Tuto úsečku máme napsat jako množinu řešení rovnic a nerovnic, což jde např. takto: $\{(x, y) \in \mathbb{R}^2 \mid 1 \leq x \leq 2, x = y\}$.
- 12.2.j) Tři body v \mathbb{R}^2 , nekonvexní. Konv. obal je trojúhelník $\{(x, y) \in \mathbb{R}^2 \mid x \geq 1, y \geq 1, x + 2y \leq 5\}$.
- 12.7.a) Je to elipsoid, tedy je to konvexní množina. Mnohostěn je to jen pro $n = 1$ (v tom případě je to úsečka).
- 12.7.b) Přímka jdoucí počátkem (tj. lineární podprostor dimenze 1) se směrovým vektorem \mathbf{v} . V uvedeném tvaru není vyjádřena jako průnik poloprostorů. Je-li ale $\mathbf{A} \in \mathbb{R}^{(n-1) \times n}$ matice taková, že $\{\alpha \mathbf{v} \mid \alpha \in \mathbb{R}\} = \text{span}\{\mathbf{v}\} = \text{null } \mathbf{A}$, přímku jsme vyjádřili jako množinou řešení lineární soustavy $\mathbf{A}\mathbf{x} = \mathbf{0}$, což už je průnik poloprostorů.
- 12.8.b) nemá extrémní bod
- 12.8.c) nemá extrémní bod
- 12.9.a) Extrémní body a odpovídající množiny $I \subseteq \{1, \dots, 4\}$:
- | | | | | |
|--------------|------------|------------|-------------------------------|------------|
| I | $\{1, 2\}$ | $\{1, 3\}$ | $\{2, 4\}$ | $\{3, 4\}$ |
| \mathbf{x} | $(2, 6)$ | $(6, 4)$ | $(\frac{4}{5}, \frac{12}{5})$ | $(2, 0)$ |
- 12.10.a) Je-li $\text{rank } \mathbf{A} < n$, je $\mathbf{A}\mathbf{v} = \mathbf{0}$ pro nějaké $\mathbf{v} \neq \mathbf{0}$. Tedy pro každé \mathbf{x} splňující $\mathbf{A}\mathbf{x} \geq \mathbf{b}$, mnohostěn obsahuje přímku $\{\mathbf{x} + \alpha \mathbf{v} \mid \alpha \in \mathbb{R}\}$. Tedy každým bodem mnohostěnu prochází přímka, která celá leží v mnohostěnu. Tedy žádný bod mnohostěnu není extrémní bod. Srov. Věta 12.4.

Kapitola 13

Simplexová metoda

V této kapitole popíšeme slavný algoritmus na řešení lineárních programů zvaný **simplexová metoda**. Pro začátek zapomeňme na účelovou funkci a zkoumejme pouze mnohostěn přípustných řešení LP. Místo mnohostěnu v obecném tvaru (12.5) algoritmus pracuje s mnohostěnem v rovnicovém tvaru (11.3)

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}, \quad (13.1)$$

kde navíc matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ má lineárně nezávislé řádky (tj. $\text{rank } \mathbf{A} = m$). Množina řešení soustavy $\mathbf{Ax} = \mathbf{b}$ je tedy afinní podprostor dimenze nejvýše $n - m$. Definujme tyto pojmy:

- Množina $J \subseteq \{1, 2, \dots, n\}$ se nazývá **báze**, 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 $\mathbf{Ax} = \mathbf{b}$ 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 $\mathbf{Ax} = \mathbf{b}$ dána tabulkou (blokovou maticí)

$$[\mathbf{A} \quad \mathbf{b}] = \left[\begin{array}{cccccc|c} -1 & 1 & 3 & 1 & 0 & 2 & 1 \\ 1 & 0 & 4 & 0 & 1 & 4 & 4 \\ -1 & 0 & 4 & 1 & 1 & 4 & 2 \end{array} \right]. \quad (13.2)$$

- $J = \{2, 3, 5\}$ 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 nepřípustné, 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) neobsahuje přímku, protože každá přímka ‘narazí’ na omezení $\mathbf{x} \geq \mathbf{0}$ (dokažte přesně!). Tedy jestliže je neprázdný, má dle Věty 12.4 aspoň jeden extrémní bod. Věta 12.2 má jednoduchý důsledek:

Důsledek 13.1. *Bod mnohostěnu (13.1) je extrémní bod (tj. vrchol), právě když je to přípustné bázové řešení.*

Důkaz. (★) Mnohostěn (13.1) lze napsat ve tvaru (12.5) jako $X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}'\mathbf{x} \geq \mathbf{b}'\}$, kde

$$\mathbf{A}' = \begin{bmatrix} \mathbf{A} \\ -\mathbf{A} \\ \mathbf{I}_n \end{bmatrix} \in \mathbb{R}^{(2m+n) \times n}, \quad \mathbf{b}' = \begin{bmatrix} \mathbf{b} \\ -\mathbf{b} \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^{2m+n}.$$

Dle Věty 12.2 je $\mathbf{x} \in \mathbb{R}^n$ extrémní bod mnohostěnu X právě tehdy, když existuje množina $I \subseteq \{1, \dots, 2m+n\}$ tak, že soustava $\mathbf{A}'_I \mathbf{x} = \mathbf{b}'_I$ má právě jedno řešení \mathbf{x} a toto řešení navíc splňuje $\mathbf{A}'\mathbf{x} \geq \mathbf{b}'$. Množina I musí obsahovat indexy $1, \dots, 2m$, protože jinak by neplatilo $\mathbf{A}\mathbf{x} = \mathbf{b}$. Výběr z posledních n řádků matice \mathbf{A}' (tedy z řádků matice \mathbf{I}_n) odpovídá rovnicím $x_j = 0$. Podmnožiny řádků matice \mathbf{I}_n , pro které soustava $\mathbf{A}'_I \mathbf{x} = \mathbf{b}'_I$ má právě jedno řešení, odpovídají bázím J . \square

Dále lze ukázat (důkaz vynecháme), že dvojice sousedních bází odpovídají buď jedinému (degenerovanému) vrcholu nebo dvojici vrcholů spojených hranou. Simplexová metoda přechází mezi sousedními bázemi tak, že bázové řešení jsou stále přípustná (tedy přechází po hranách mnohostěnu) a účelová funkce se zlepšuje nebo aspoň nezhoršuje.

V §13.1 popíšeme stavební kameny metody, které pak v §13.2 spojíme do celého algoritmu.

13.1 Stavební kameny algoritmu

13.1.1 Přejít k sousední standardní bázi

Simplexová metoda pracuje pouze se *standardními* bázemi, tj. sloupce J jsou 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 $\mathbf{A}\mathbf{x} = \mathbf{b}$: libovolný řádek tabulky $[\mathbf{A} \mid \mathbf{b}]$ 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' .

Říkáme, že jsme provedli *ekvivalentní úpravu kolem pivotu* s indexy (i, j) .

Příklad 13.2. Mějme soustavu

$$[\mathbf{A} \quad \mathbf{b}] = \left[\begin{array}{cccccc|c} 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & \boxed{1} & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \end{array} \right] \quad (13.4)$$

se (standardní) bází $J = \{1, 4, 5\}$. Vidíme ihned odpovídající bázevé řešení, $\mathbf{x} = (3, 0, 0, 4, 1, 0)$.

Chceme nahradit bázevý sloupec $j' = 1$ nebázevým 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:

$$[\mathbf{A} \quad \mathbf{b}] = \left[\begin{array}{cccccc|c} -2 & 0 & 0 & 1 & 0 & 0 & -2 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 1 & 0 & 4 & 0 & 1 & 4 & 4 \end{array} \right].$$

Nyní sloupce $\{2, 4, 5\}$ tvoří standardní bázi. □

13.1.2 Kdy je nové bázevé řešení přípustné?

Uvedeným způsobem můžeme od aktuální báze přejít k libovolné sousední bázi. Přitom nové bázevé řešení může nebo nemusí být přípustné. Je-li aktuální bázevé řešení přípustné, jak poznáme, zda i nové bázevé řešení bude přípustné?

Protože nenulové složky bázevého řešení \mathbf{x} jsou rovny složkám vektoru \mathbf{b} , bázevé řešení je přípustné právě tehdy, když $\mathbf{b} \geq \mathbf{0}$. Necht' v aktuální tabulce je $\mathbf{b} \geq \mathbf{0}$. Proved'me 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 zůstat 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. V tabulce (13.4):

- Ekvivalentní úprava okolo pivotu $(i, j) = (3, 2)$ nepovede k přípustnému báзовému řešení, neboť $a_{ij} = -1 < 0$, což porušuje podmínku (13.5a).
- Ekvivalentní úprava okolo pivotu $(i, j) = (2, 2)$ nepovede k přípustnému báзовému 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 báзовému 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 báзовým. 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 báзовé řešení \mathbf{x} . Když se x_2 bude libovolně zvětšovat, změnu lze kompenzovat současným zvětšováním báзовých proměnných x_1, x_4, x_5 tak, že vektor \mathbf{Ax} zůstane nezměněn a tedy roven \mathbf{b} . Konkrétně, 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 nevšímalí. Tyto úpravy lze rozšířit na účelovou funkci. Nebudeme účelovou funkci 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?

Nechť 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 *redukované ceny*). Protože bázevé řešení \mathbf{x} je v nebázeových sloupcích nulové, znamená to $\mathbf{c}^T \mathbf{x} = 0$. Tedy hodnota účelové funkce $\mathbf{c}^T \mathbf{x} - d$ v bázeovém řešení \mathbf{x} je rovna jednoduše $-d$.

Navíc je snadno vidět, co udělá přechod k nové bázi s účelovou funkcí. Nechť 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

$$\begin{bmatrix} \mathbf{c}^T & d \\ \mathbf{A} & \mathbf{b} \end{bmatrix} = \left[\begin{array}{cccccc|c} 1 & -2 & -3 & -1 & 2 & 1 & 4 \\ 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \end{array} \right],$$

kde $J = \{1, 4, 5\}$. Složky vektoru \mathbf{c} v bázeových sloupcích vynulujeme tak, že k účelovému řádku přičteme první řádek, odečteme druhý řádek, a odečteme dvojnásobek třetího řádku:

$$\begin{array}{cccccc|c} 0 & 1 & -2 & 0 & 0 & -1 & 3 \\ \hline 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \\ \hline \mathbf{x} = & 3 & 0 & 0 & 4 & 1 & 0 \end{array}$$

Pod tabulku jsme napsali bázevé řešení \mathbf{x} . Nyní je $\mathbf{c}^T \mathbf{x} = 0$, a tedy hodnota účelové funkce v bázeovém řešení je $\mathbf{c}^T \mathbf{x} - d = -d = -3$.

Dejme tomu, že chceme přidat do báze nebázeový sloupec 2 a vyloučit z ní některý z bázeových 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 $c_2 x_2$. Kritérium tedy stoupne nebo zůstane stejné, protože $c_2 = 1 > 0$. \square

Jestliže v některém sloupci j je $c_j < 0$ a $a_{ij} \leq 0$ pro všechna i , proměnnou x_j můžeme 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 bázevé 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 ,
- bázevé řešení odpovídající této bázi je přípustné, $\mathbf{b} \geq \mathbf{0}$,
- složky vektoru \mathbf{c} v bázeových 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!)

$$i \in \operatorname{argmin}_{i' \mid a_{i'j} > 0} \frac{b_{i'}}{a_{i'j}}, \quad (13.8)$$

kde $\operatorname{argmin}_{i' \mid a_{i'j} > 0}$ označuje, že se minimalizuje přes všechna i' splňující $a_{i'j} > 0$.

3. Udělej ekvivalentní úpravu okolo pivotu (i, j) (§13.1.1).
4. Ekvivalentní úpravou účelového řádku vynuluj c_j v novém báзовém sloupci j (§13.1.5).

Algoritmus, který opakuje uvedenou iteraci, nazveme **základní simplexový algoritmus**. Algoritmus končí, když už nelze vybrat pivot (i, j) splňující podmínky 1 a 2 výše. To nastane z jednoho z těchto důvodů:

- Všechny koeficienty c_j jsou nezáporné. Účelovou funkci nelze zlepšit a jsme v optimu.
- V každém sloupci j , kde $c_j < 0$, máme $a_{ij} \leq 0$ 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 množinu 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 prvý, druhý atd. První iterace simplexového algoritmu se provede v těchto krocích:

1. Vybereme sloupec j , který vstoupí do báze. To může být libovolný sloupec, který má v nultém řádku záporné číslo. Můžeme vzít např. nejmenší takové číslo, zde -3 , tedy $j = 6$.
2. 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.
- 3, 4. 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 $J = \{1, 4, 6\}$, bazové 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

$$\begin{array}{cccccc|c} -2 & 6 & 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 1 & 0 & 2 & 2 \\ \boxed{2} & -1 & -2 & 0 & 1 & 1 & 1 \end{array}$$

Tabulka po první iteraci je

$$\begin{array}{cccccc|c} 0 & 5 & -1 & 0 & 1 & 1 & 1 \\ 0 & -1.5 & -2 & 1 & 0.5 & 2.5 & 2.5 \\ 1 & -0.5 & -1 & 0 & 0.5 & 0.5 & 0.5 \end{array}$$

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 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 má simplexovou tabulku

$$\begin{bmatrix} \mathbf{c}^T & \mathbf{0} & \mathbf{0} \\ \mathbf{A} & \mathbf{I} & \mathbf{b} \end{bmatrix},$$

ve které sloupce příslušné proměnným \mathbf{u} tvoří standardní bázi.

Příklad 13.8. Vyřešte simplexovým algoritmem:

$$\begin{array}{ll} \min & -3x_1 - x_2 - 3x_3 \\ \text{za podmínek} & 2x_1 + x_2 + x_3 \leq 2 \\ & x_1 + 2x_2 + 3x_3 \leq 5 \\ & 2x_1 + 2x_2 + x_3 \leq 6 \\ & x_1, x_2, x_3 \geq 0 \end{array}$$

Přidáme slackové proměnné $u_1, u_2, u_3 \geq 0$, abychom omezení uvedli do tvaru rovností:

$$\begin{array}{llllll} \min & -3x_1 - x_2 - 3x_3 & & & & \\ \text{za podmínek} & 2x_1 + x_2 + x_3 + u_1 & & & & = 2 \\ & x_1 + 2x_2 + 3x_3 & + u_2 & & & = 5 \\ & 2x_1 + 2x_2 + x_3 & & + u_3 & & = 6 \\ & x_1, x_2, x_3, u_1, u_2, u_3 & \geq & 0 & & \end{array}$$

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

$$\min\{\mathbf{1}^T \mathbf{u} \mid \mathbf{Ax} + \mathbf{u} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}, \mathbf{u} \geq \mathbf{0}\} \quad (13.10)$$

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 když $\mathbf{u} = \mathbf{0}$. Tedy úloha (13.9) je přípustná, právě když je 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 bázev proměnné kladné. Protože $\mathbf{u} = \mathbf{0}$, proměnné \mathbf{u} budou tedy nebázové. Proto mezi sloupce příslušnými proměnným \mathbf{x} existuje standardní báze.

Pokud je optimální řešení (\mathbf{x}, \mathbf{u}) úlohy (13.10) degenerované, některé proměnné \mathbf{u} mohou být na konci algoritmu bázev. Pak je nutno udělat dodatečné úpravy kolem pivotů ve sloupcích příslušných bázevým proměnným \mathbf{u} , abychom tyto sloupce dostali z báze ven. Toto podrobně popisovat nebudeme.

Nalezení nějakého přípustného řešení v pomocné úloze (13.10) se nazývá **první fáze** a řešení původní úlohy pak **druhá fáze** algoritmu, mluvíme tedy o **dvoufázové 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

$$\begin{array}{cccccc|c} 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 3 & 2 & 1 & 1 & 0 & 0 & 10 \\ 1 & 2 & 2 & 0 & 1 & 0 & 15 \end{array}$$

Sloupce nad přidanými proměnnými tvoří standardní bázi, můžeme tedy na pomocnou úlohu pustit základní simplexový algoritmus. Po vynulování ceny nad bázevými proměnnými budou kroky algoritmu vypadat takto:

$$\begin{array}{cccccc|c} -4 & -4 & -3 & 0 & 0 & 0 & -25 \\ 3 & \boxed{2} & 1 & 1 & 0 & 0 & 10 \\ 1 & 2 & 2 & 0 & 1 & 0 & 15 \\ \hline 2 & 0 & -1 & 2 & 0 & 0 & -5 \\ 1.5 & 1 & 0.5 & 0.5 & 0 & 0 & 5 \\ -2 & 0 & \boxed{1} & -1 & 1 & 0 & 5 \\ \hline 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 2.5 & 1 & 0 & 1 & -0.5 & 0 & 2.5 \\ -2 & 0 & 1 & -1 & 1 & 0 & 5 \end{array}$$

Optimum je rovno 0, tedy původní úloha je přípustná. Proměnné u_1, u_2 jsou nebázové a tedy rovny nule, bážové proměnné jsou x_2, x_3 . Ted' tedy můžeme začít druhou fází (řešení původní úlohy) s počáteční tabulkou

$$\begin{array}{ccc|c} -20 & -30 & -40 & 0 \\ \hline 2.5 & 1 & 0 & 2.5 \\ -2 & 0 & 1 & 5 \end{array}$$

□

13.4 Cvičení

13.1. Najděte všechny báze a bážová řešení mnohostěnu $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ pro

$$[\mathbf{A} \quad \mathbf{b}] = \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 2 & -1 & 4 & 1 & 4 \end{array} \right].$$

Která z nich jsou přípustná? Která z nich jsou degenerovaná?

13.2. Protože $\text{rank } \mathbf{A} = m$, dimenze mnohostěnu (13.1) je $\dim X \leq n - m$. Může být $\dim X$ libovolné číslo mezi 0 a $n - m$? Jestliže ano, dokažte to tak, že pro každou dvojici (n, k) , kde $n \in \mathbb{N}$ a $k = 0, \dots, n - m$, najdete matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ a vektor $\mathbf{b} \in \mathbb{R}^m$ tak, aby $\text{rank } \mathbf{A} = m$ a $\dim X = k$.

13.3. V tabulce zakroužkujte všechny pivoty takové, že ekvivalentní úprava kolem nich povede k přípustnému bážovému řešení:

$$[\mathbf{A} \quad \mathbf{b}] = \left[\begin{array}{cccccc|ccc} 0 & 2 & 6 & 1 & 0 & -4 & 3 & 0 & 4 \\ 1 & 1 & -3 & 0 & 0 & 2 & 3 & 0 & 3 \\ 0 & -1 & 1 & 0 & 1 & -2 & -3 & 0 & 1 \\ 0 & -2 & 2 & 0 & 0 & 2 & -1 & 1 & 1 \end{array} \right]$$

13.4. Zapište lineární program

$$\begin{array}{llll} \min & -x_1 & & -x_4 - 3x_5 \\ \text{za podmíněk} & 2x_1 & & + x_4 + x_5 + x_6 = 2 \\ & -x_1 + x_2 & & + 2x_4 + 3x_5 = 4 \\ & 2x_1 & + x_3 + 2x_4 - x_5 & = 6 \\ & & & x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{array}$$

do simplexové tabulky. Předpokládejte, že aktuální báze je $\{2, 3, 6\}$. Jaké je aktuální bážové řešení? Je toto bážové řešení přípustné. Je degenerované? Pokud je to možné, udělejte jeden krok simplexového algoritmu. Pokud to možné není, vysvětlete proč.

13.5. Vyřešte simplexovou metodou:

$$\begin{array}{ll} \max & 2x_1 - x_2 - 3x_3 \\ \text{za podmíněk} & -2x_1 - x_2 + x_3 \leq 2 \\ & -x_1 + 2x_2 - 3x_3 \leq 5 \\ & -2x_1 - 4x_2 + x_3 \leq 6 \\ & x_1, x_2, x_3 \geq 0 \end{array}$$

13.6. Vyřešte simplexovou metodou (i když lze řešit úvahou):

$$\begin{aligned} \max \quad & 6x_1 + 9x_2 + 5x_3 + 9x_4 \\ \text{za podmínek} \quad & x_1 + x_2 + x_3 + x_4 = 1 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

13.7. Necht' úloha (13.6) má více optimálních řešení. Jak se to pozná v simplexové tabulce? Navrhněte algoritmus, který vypíše všechna optimální bázeová řešení.

13.8. Mějme lineární program

$$\begin{aligned} \min \quad & 2x_1 - 3x_3 + x_4 \\ \text{za podmínek} \quad & x_1 - x_2 - x_3 \geq 0 \\ & -x_1 + 2x_2 - 3x_3 \leq 5 \\ & 2x_1 - x_2 - x_3 + 2x_4 = 6 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

Inicializujte co nejjednodušším způsobem základní simplexový algoritmus. Vyřešte tímto algoritmem. Nepoužívejte dvoufázovou metodu.

13.9. Vyřešte dvoufázovou simplexovou metodou:

$$\begin{aligned} \max \quad & 3x_1 - 4x_2 \\ \text{za podmínek} \quad & -2x_1 - 5x_2 \leq 10 \\ & 3x_1 + x_2 \leq 3 \\ & -2x_1 + x_2 \leq -2 \\ & x_1 \geq 0 \\ & x_2 \leq -1 \end{aligned}$$

13.10. Necht' úloha (13.6) má simplexovou tabulku

$$\begin{bmatrix} \mathbf{c}^T & d \\ \mathbf{A} & \mathbf{b} \end{bmatrix} = \left[\begin{array}{cccccc|c} 3 & 0 & 0 & 1 & -3 & 0 & 0 \\ 7 & 0 & 0 & -3 & 3 & 2 & 1 \\ 1 & 1 & 0 & 0 & -1 & 0 & 3 \\ -1 & 0 & 2 & 1 & -1 & 0 & 1 \end{array} \right]$$

Je tato úloha omezená?

Nápověda a řešení

13.1.

J	$\{1, 2\}$	$\{1, 3\}$	$\{1, 4\}$	$\{2, 3\}$	$\{3, 4\}$
\mathbf{x}	$(1, -2)$	$(0, 1)$	$(1, 2)$	$(0, 1)$	$(1, 0)$
příp.	ne	ano	ano	ano	ano
degen.	ne	ano	ne	ano	ano

13.5. Úloha je neomezená kvůli prvnímu sloupci.

13.8. Např. otočíme znaménko prvního omezení, třetí omezení vydělíme dvěma, přidáme slackové proměnné pro první a druhé omezení. Pak vynulujeme ceny nad bázeovými sloupci. Iterace algoritmu:

$$\begin{array}{cccccc|c} 1 & 0.5 & -2.5 & 0 & 0 & 0 & -3 \\ -1 & 1 & \boxed{1} & 0 & 1 & 0 & 0 \\ -1 & 2 & -3 & 0 & 0 & 1 & 5 \\ 1 & -0.5 & -0.5 & 1 & 0 & 0 & 3 \end{array} \quad \begin{array}{cccccc|c} -1.5 & 3 & 0 & 0 & 2.5 & 0 & -3 \\ -1 & 1 & 1 & 0 & 1 & 0 & 0 \\ -4 & 5 & 0 & 0 & 3 & 1 & 5 \\ \boxed{0.5} & 0 & 0 & 1 & 0.5 & 0 & 3 \end{array} \quad \begin{array}{cccccc|c} 0 & 3 & 0 & 3 & 4 & 0 & 6 \\ 0 & 1 & 1 & 2 & 2 & 0 & 6 \\ 0 & 5 & 0 & 8 & 7 & 1 & 29 \\ 1 & 0 & 0 & 2 & 1 & 0 & 6 \end{array}$$

Výsledek: $(x_1, x_2, x_3, x_4) = (6, 0, 6, 0)$, hodnota optima -6 .

13.9. Optimum je $(x_1, x_2) = (25, -36)/13$.

13.10. Úloha je neomezená. Mohlo by se zdát, že úloha je omezená, protože neexistuje sloupec j takový, že $a_{ij} \leq 0$ a $c_j > 0$. Ale po jedné či několika iteracích simplexové metody se takový sloupec objeví.

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:

$$\begin{array}{llll} \min & \sum_{j \in J} c_j x_j & \max & \sum_{i \in I} b_i y_i \\ \text{za podm.} & \sum_{j \in J} a_{ij} x_j = b_i & \text{za podm.} & y_i \in \mathbb{R}, \quad i \in I_0 \\ & \sum_{j \in J} a_{ij} x_j \geq b_i & & y_i \geq 0, \quad i \in I_+ \\ & \sum_{j \in J} a_{ij} x_j \leq b_i & & y_i \leq 0, \quad i \in I_- \quad (14.1) \\ & x_j \in \mathbb{R} & & \sum_{i \in I} a_{ij} y_i = c_j, \quad j \in J_0 \\ & x_j \geq 0 & & \sum_{i \in I} a_{ij} y_i \leq c_j, \quad j \in J_+ \\ & x_j \leq 0 & & \sum_{i \in I} a_{ij} y_i \geq c_j, \quad j \in J_- \end{array}$$

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 dvojice úloh: i -tému primárnímu omezení $\sum_j a_{ij} x_j \geq b_i$ odpovídá duální proměnná¹ $y_i \geq 0$. Opačně, j -tá primární proměnná $x_j \geq 0$ odpovídá j -tému duálnímu omezení $\sum_i a_{ij} y_i \leq c_j$. Tak je to pro všechny řádky s tím, že lineární rovnici v primáru odpovídá neomezená proměnná v duálu, nerovnici typu \geq v primáru odpovídá nezáporná proměnná v duálu, a nerovnici typu \leq v primáru odpovídá nekladná proměnná v duálu.

¹Poznamenejme, že tato proměnná je vlastně *Lagrangeův multiplikátor* příslušného omezení. Podobně, j -tá primární proměnná x_j je Lagrangeův multiplikátor j -tého duálního omezení $\sum_i a_{ij} x_j \leq c_j$.

Příklad 14.1. Následující dvojice lineárních programů je navzájem duální:

$$\begin{array}{ll}
\min & 2x_1 \quad - 3x_3 + x_4 \\
\text{za podm.} & 2x_1 - x_2 + x_3 + 2x_4 = 6 \\
& -x_1 + 2x_2 - 3x_3 \leq 5 \\
& x_1 - x_2 - x_3 - 3x_4 \geq 0 \\
& x_1 \geq 0 \\
& x_2 \in \mathbb{R} \\
& x_3 \geq 0 \\
& x_4 \leq 0
\end{array}
\qquad
\begin{array}{ll}
\max & 6y_1 + 5y_2 \\
\text{za podm.} & y_1 \in \mathbb{R} \\
& y_2 \leq 0 \\
& y_3 \geq 0 \\
& 2y_1 - y_2 + y_3 \leq 2 \\
& -y_1 + 2y_2 - y_3 = 0 \\
& y_1 - 3y_2 - y_3 \leq -3 \\
& 2y_1 - 3y_3 \geq 1
\end{array}
\quad \square$$

Pro speciální tvary LP se dvojice duálních úloh přehledněji napíše v maticové formě. Např. pro $I_0 = I_- = J_0 = J_- = \emptyset$ obdržíme

$$\begin{array}{ll}
\min & \mathbf{c}^T \mathbf{x} \\
\text{za podm.} & \mathbf{Ax} \geq \mathbf{b} \\
& \mathbf{x} \geq \mathbf{0}
\end{array}
\qquad
\begin{array}{ll}
\max & \mathbf{b}^T \mathbf{y} \\
\text{za podm.} & \mathbf{y} \geq \mathbf{0} \\
& \mathbf{A}^T \mathbf{y} \leq \mathbf{c}
\end{array}
\quad (14.2)$$

14.2 Věty o dualitě

Následující věty platí pro obecný tvar (14.1), ale důkazy uděláme jen pro speciální tvar (14.2).

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 $\mathbf{y}^T \mathbf{Ax} \leq \mathbf{c}^T \mathbf{x}$. Podobně, díky přípustnosti \mathbf{x} a \mathbf{y} platí $\mathbf{Ax} \geq \mathbf{b}$ a $\mathbf{y} \geq \mathbf{0}$, z čehož plyne $\mathbf{y}^T \mathbf{Ax} \geq \mathbf{y}^T \mathbf{b}$. Z toho

$$\mathbf{c}^T \mathbf{x} \geq \mathbf{y}^T \mathbf{Ax} \geq \mathbf{y}^T \mathbf{b}. \quad (14.3)$$

□

Uved'me jeden okamžitý důsledek slabé duality.

Důsledek 14.2. *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} je optimální pro primární úlohu a \mathbf{y} je optimální pro duální úlohu.*

Důkaz. Pro libovolné primární přípustné řešení \mathbf{x}' je dle věty o slabé dualitě $\mathbf{c}^T \mathbf{x}' \geq \mathbf{b}^T \mathbf{y} = \mathbf{c}^T \mathbf{x}$. Toto platí pro každé přípustné \mathbf{x}' , tedy řešení \mathbf{x} musí být optimální pro primární úlohu.

Optimalita \mathbf{y} pro duální úlohu se dokáže symetricky. □

Věta 14.3 (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:*

$$\left(\sum_{j \in J} a_{ij} x_j = b_i \right) \text{ nebo } (y_i = 0) \quad \forall i \in I, \quad (14.4a)$$

$$(x_j = 0) \text{ nebo } \left(\sum_{i \in I} a_{ij} y_j = c_j \right) \quad \forall j \in J. \quad (14.4b)$$

Podmínky (14.4) budeme nazývat *podmínky komplementarity*. Říkají, že na každém řádku ve dvojici duálních úloh je vždy alespoň jedno omezení aktivní, buď primární nebo duální (přičemž omezení typu rovnosti považujeme vždy za aktivní).

Důkaz. Z nerovnosti (14.3) je jasné, že pro přípustná \mathbf{x}, \mathbf{y} platí ekvivalence

$$(\mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{b}) \iff (\mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{b}). \quad (14.5)$$

Dvě rovnosti na pravé straně této ekvivalence jde napsat jako

$$\mathbf{y}^T (\mathbf{A} \mathbf{x} - \mathbf{b}) = 0, \quad (14.6a)$$

$$(\mathbf{c}^T - \mathbf{y}^T \mathbf{A}) \mathbf{x} = 0. \quad (14.6b)$$

Levé strany těchto rovností jsou skalární součiny nezáporných (díky přípustnosti \mathbf{x}, \mathbf{y}) vektorů. Nyní si stačí uvědomit, že pro libovolné nezáporné vektory $\mathbf{u}, \mathbf{v} \geq \mathbf{0}$ platí

$$\mathbf{u}^T \mathbf{v} = 0 \iff \forall i (u_i v_i = 0) \iff \forall i (u_i = 0 \text{ nebo } v_i = 0). \quad \square$$

Uvědomte si, že Důsledek 14.2 a Věta 14.3 neříkají, že rovnost $\mathbf{c}^T \mathbf{x} = \mathbf{b}^T \mathbf{y}$ vůbec někdy nastane. To je předmětem nejdůležitější věty o dualitě:

Věta 14.4 (o silné dualitě). *Primární úloha má optimální řešení, právě 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 této věty 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:

$$\begin{array}{c} \text{---} \overline{\{\mathbf{b}^T \mathbf{y} \mid \mathbf{y} \geq \mathbf{0}, \mathbf{A}^T \mathbf{y} \leq \mathbf{c}\}} \quad \bigg| \quad \overline{\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A} \mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}} \text{---} \\ \text{---} \quad \quad \quad \uparrow \quad \quad \quad \text{---} \\ \text{společné optimum } \mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{b} \end{array}$$

Z věty o silné dualitě ihned plyne věta o slabé dualitě (proč?), proto se někdo může ptát, proč větu o slabé dualitě vůbec uvádíme. Je to proto, že její důkaz je o mnoho jednodušší.

Příklad 14.2. Mějme dvojici navzájem duálních úloh LP:

$\begin{array}{ll} \min & 2x_1 + 5x_2 + 6x_3 = \mathbf{5.4} \\ \mathbf{3} = & 2x_1 + x_2 + 2x_3 \geq 3 \\ \mathbf{2.4} = & x_1 + 2x_2 + 2x_3 \geq 1 \\ \mathbf{3} = & x_1 + 3x_2 + x_3 \geq 3 \\ -\mathbf{0.6} = & -x_1 + x_2 - 2x_3 \geq -1 \\ \mathbf{1.2} = & x_1 \geq 0 \\ \mathbf{0.6} = & x_2 \geq 0 \\ \mathbf{0} = & x_3 \geq 0 \end{array}$	$\begin{array}{ll} \max & 3y_1 + y_2 + 3y_3 - y_4 = \mathbf{5.4} \\ \mathbf{0.2} = & y_1 \geq 0 \\ \mathbf{0} = & y_2 \geq 0 \\ \mathbf{1.6} = & y_3 \geq 0 \\ \mathbf{0} = & y_4 \geq 0 \\ \mathbf{2} = & 2y_1 + y_2 + y_3 - y_4 \leq 2 \\ \mathbf{5} = & y_1 + 2y_2 + 3y_3 + y_4 \leq 5 \\ \mathbf{2} = & 2y_1 + 2y_2 + y_3 - 2y_4 \leq 6 \end{array}$
--	--

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 14.3. Zkuste dokázat bez použití simplexové metody, že $\mathbf{x} = (x_1, x_2, x_3) = (1.2, 0.6, 0)$ je optimální řešení úlohy z Příkladu 14.2 (přičemž optimální duální řešení \mathbf{y} není známo).

Pomocí věty o komplementaritě zkusíme z daného optimálního \mathbf{x} 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

$$\begin{aligned} 2y_1 + y_3 &= 2 \\ y_1 + 3y_3 &= 5 \end{aligned} \tag{14.7}$$

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 měla nekonečně mnoho optimálních řešení, soustava (14.7) 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

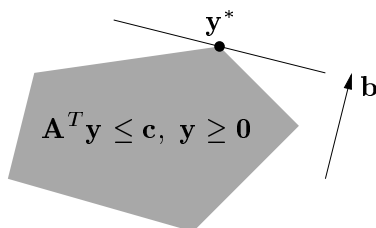
Zkoumejme, jak se změní optimální hodnota úlohy $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, jestliže nepatrně změním pravé strany omezení \mathbf{b} . Odpověď je snadno vidět v duálu.

Věta 14.5 (o stínových cenách). Necht' funkce $f: \mathbb{R}^m \rightarrow \mathbb{R}$ je definována jako

$$f(\mathbf{b}) = \min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\} = \max\{\mathbf{b}^T \mathbf{y} \mid \mathbf{A}^T \mathbf{y} \leq \mathbf{c}, \mathbf{y} \geq \mathbf{0}\}.$$

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 $\{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{A}^T \mathbf{y} \leq \mathbf{c}, \mathbf{y} \geq \mathbf{0}\}$, 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 $f(\mathbf{b}) = \mathbf{b}^T \mathbf{y}^*$, v malém okolí bodu \mathbf{b} je tedy funkce f lineární. Její derivace je $f'(\mathbf{b}) = \mathbf{y}^{*T}$. \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 $\mathbf{Ax} \geq \mathbf{b}$. Interpretujeme-li naše LP jako optimální výrobní plán (11.7) (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 dovoleno $\mathbf{a}_i^T \mathbf{x} < b_i$ a tedy malá změna b_i nemá na optimum vliv.

Příklad 14.4. Necht' je známo, že duální úloha v Příkladu 14.2 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:

$$\begin{array}{ll}
 \min & 2x_1 + 5x_2 + 6x_3 = \mathbf{5.402} \\
 & 2x_1 + x_2 + 2x_3 \geq 3.01 \\
 & x_1 + 2x_2 + 2x_3 \geq 1 \\
 & x_1 + 3x_2 + x_3 \geq 3 \\
 & -x_1 + x_2 - 2x_3 \geq -1 \\
 & x_1 \geq 0 \\
 & x_2 \geq 0 \\
 & x_3 \geq 0 \\
 \max & 3.01y_1 + y_2 + 3y_3 - y_4 = \mathbf{5.402} \\
 & \mathbf{0.2} = y_1 \geq 0 \\
 & \mathbf{0} = y_2 \geq 0 \\
 & \mathbf{1.6} = y_3 \geq 0 \\
 & \mathbf{0} = y_4 \geq 0 \\
 & \mathbf{2} = 2y_1 + y_2 + y_3 - y_4 \leq 2 \\
 & \mathbf{5} = y_1 + 2y_2 + 3y_3 + y_4 \leq 5 \\
 & \mathbf{2} = 2y_1 + 2y_2 + y_3 - 2y_4 \leq 6
 \end{array}$$

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 dobré pochopit význam nejen primární úlohy, ale i duální úlohy a vztahy mezi primární a duální úlohou (tj. věty o dualitě). Často se nám podaří dokázat platnost silné duality pro náš konkrétní problém.

Příklad 14.5. Demonstrujme nyní dualitu na velmi jednoduchém lineárním programu

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0} \} = \min \{ c_1 x_1 + \cdots + c_n x_n \mid x_1 + \cdots + x_n = 1, x_i \geq 0 \},$$

kde čísla $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{R}^n$ jsou dána. Chceme přesně rozumět, proč v tomto případě platí věty o silné dualitě a o komplementaritě.

Úvahou snadno vidíme (viz Cvičení 11.3), 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 (14.1) sestrojíme duální úlohu

$$\max\{y \in \mathbb{R} \mid y\mathbf{1} \leq \mathbf{c}\} = \max\{y \in \mathbb{R} \mid y \leq c_i, i = 1, \dots, n\}.$$

Neboli hledá se největší číslo y , které je menší než všechna čísla c_i . Takové číslo y je zřejmě nejmenší 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 proto 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 14.6. Pro daná $\mathbf{c} \in \mathbb{R}^n$ a $k \in \{1, \dots, n\}$ máme úlohu

$$\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{1}^T \mathbf{x} \leq k\} \quad (14.8)$$

Duální úlohu sestrojíme dle předpisu (14.1):

$$\begin{array}{ll} \max & c_1 x_1 + \dots + c_n x_n \\ \text{za podm.} & x_1 + \dots + x_n \leq k \\ & x_i \leq 1 \\ & x_i \geq 0 \end{array} \quad \begin{array}{ll} \min & ky + z_1 + \dots + z_n \\ \text{za podm.} & y \geq 0 \\ & z_i \geq 0 \quad i = 1, \dots, n \\ & y + z_i \geq c_i \quad i = 1, \dots, n \end{array}$$

Pravá úloha odpovídá levé úloze z dvojice (14.1), protože v úloze (14.8) se maximalizuje. Sestrojít duální úlohu je zde obtížnější než např. v Příkladě 14.2 kvůli přítomnosti indexů $i = 1, \dots, n$. Chce to cvik. Doporučujeme např. napsat dvojici (14.1) podrobně pro nějaké konkrétní malé n (např. $n = 3$) a pak ji přepsat do obecného tvaru.

Podmínky komplementarity:

- $\sum_i x_i = k$ nebo $y = 0$
- Pro každé i je $x_i = 1$ nebo $z_i = 0$
- Pro každé i je $x_i = 0$ nebo $y + z_i = c_i$

Primární úloha (14.8) se snadno vyřeší úvahou (viz Cvičení 11.3): její optimální hodnota je součet k největších kladných čísel c_i . Ovšem vůbec není patrné, proč duální úloha (??) má stejnou optimální hodnotu. Zkuste takovou úvahu vymyslet! \square

Příklad 14.7 (Ekonomická interpretace duality). Vraťme se k Příkladu 11.4 o výrobci lupínků a hranolků z brambor a oleje. Napišme k této úloze duální úlohu:

$$\begin{array}{ll} \max & 120l + 76h \\ \text{za podm.} & \begin{array}{l} 2l + 1.5h \leq 100 \\ 0.4l + 0.2h \leq 16 \\ l \geq 0 \\ h \geq 0 \end{array} \end{array} \quad \begin{array}{ll} \min & 100a + 16b \\ \text{za podm.} & \begin{array}{l} a \geq 0 \\ b \geq 0 \\ 2a + 0.4b \geq 120 \\ 1.5a + 0.2b \geq 76 \end{array} \end{array}$$

Přijde překupník a chce koupit od výrobce jeho zásoby brambor a oleje. Překupník řeší tuto otázku: Jaké nejnižší ceny musím nabídnout, aby mi výrobce ještě 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 14.8. Z §11.5 víme, že optimální argument úlohy

$$\min_{x \in \mathbb{R}} \sum_{i=1}^n |x - a_i| = \min \{ z_1 + \dots + z_n \mid z_i \in \mathbb{R}, x \in \mathbb{R}, -z_i \leq x - a_i \leq z_i \} \quad (14.9)$$

je medián z čísel a_1, \dots, a_n . Chceme sestavit duální úlohu a zjednodušit ji. Chceme pro tuto úlohu dokázat platnost silné duality.

Duální úlohu je možné sestavit podle předpisu (14.1):

$$\begin{array}{ll} \min & \sum_{i=1}^n z_i \\ \text{za podm.} & \begin{array}{l} x + z_i \geq a_i \\ -x + z_i \geq -a_i \\ z_i \in \mathbb{R} \\ x \in \mathbb{R} \end{array} \end{array} \quad \begin{array}{ll} \max & \sum_{i=1}^n (p_i - q_i)a_i \\ \text{za podm.} & \begin{array}{l} p_i \geq 0 \\ q_i \geq 0 \\ p_i + q_i = 1 \\ \sum_{i=1}^n (p_i - q_i) = 0 \end{array} \end{array} \quad \begin{array}{l} i = 1, \dots, n \\ i = 1, \dots, n \\ i = 1, \dots, n \end{array}$$

Napsat duální úlohu takto přímo je ovšem obtížné. Proto je lépe postupovat zdlouhavějším ale bezpečnějším způsobem přes matice. Dvojici úloh lze psát v maticovém tvaru

$$\begin{array}{ll} \min & \mathbf{1}^T \mathbf{z} \\ \text{za podm.} & \begin{array}{l} \mathbf{1}x + \mathbf{z} \geq \mathbf{a} \\ -\mathbf{1}x + \mathbf{z} \geq -\mathbf{a} \\ \mathbf{z} \in \mathbb{R}^n \\ x \in \mathbb{R} \end{array} \end{array} \quad \begin{array}{ll} \max & \mathbf{a}^T (\mathbf{p} - \mathbf{q}) \\ \text{za podm.} & \begin{array}{l} \mathbf{p} \geq \mathbf{0} \\ \mathbf{q} \geq \mathbf{0} \\ \mathbf{p} + \mathbf{q} = \mathbf{1} \\ \mathbf{1}^T (\mathbf{p} - \mathbf{q}) = 0 \end{array} \end{array}$$

neboli

$$\begin{array}{ll} \min & \mathbf{h}^T \mathbf{u} \\ \text{za podm.} & \begin{array}{l} \mathbf{F} \mathbf{u} \geq \mathbf{g} \\ \mathbf{u} \in \mathbb{R}^{1+n} \end{array} \end{array} \quad \begin{array}{ll} \max & \mathbf{g}^T \mathbf{v} \\ \text{za podm.} & \begin{array}{l} \mathbf{v} \geq \mathbf{0} \\ \mathbf{F}^T \mathbf{v} = \mathbf{h} \end{array} \end{array}$$

kde

$$\mathbf{F} = \begin{bmatrix} \mathbf{1} & \mathbf{I} \\ -\mathbf{1} & \mathbf{I} \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} \mathbf{a} \\ -\mathbf{a} \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} x \\ \mathbf{z} \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix}.$$

Vektor duálních proměnných \mathbf{v} jsme zde rozdělili na dva bloky \mathbf{p}, \mathbf{q} , odpovídající blokům matic \mathbf{F} a \mathbf{g} . Při konstrukci duálu tedy nejprve napíšeme primární úlohu v maticové formě, k ní napíšeme duál v maticové formě, a ten pak zjednodušíme vynásobením matic.

Duální úlohu dále zjednodušíme substitucí

$$2p_i = 1 + t_i, \quad 2q_i = 1 - t_i.$$

Po této substituci 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

$$\max\{a_1 t_1 + \cdots + a_n t_n \mid t_i \in \mathbb{R}, t_1 + \cdots + t_n = 0, -1 \leq t_i \leq 1\}. \quad (14.10)$$

Primární úloha (14.9) a duální úloha (14.10) spolu zdánlivě vůbec nesouvisejí – avšak podle silné duality jejich optimální hodnoty musejí být stejné. Zkusme dokázat, že tomu tak je.

Nejprve si všimneme, že optimální hodnota primární úlohy (14.9) 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 (14.10), 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

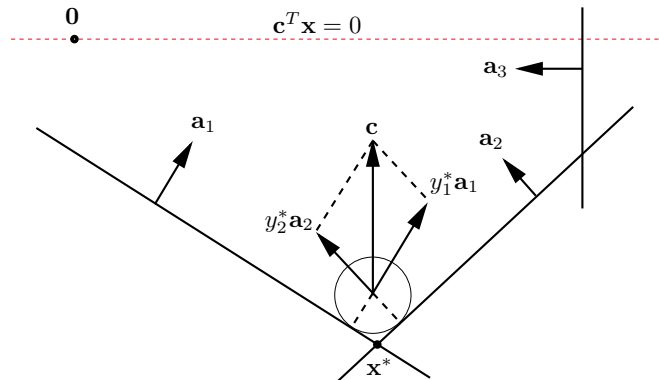
14.3.1 Mechanické modely

Lineární programy lze modelovat mechanickými (obecněji fyzikálními) modely. Můžeme jim říkat mechanické či analogové počítače. Takové modely se dobře hodí na ilustraci duality.

Příklad 14.9. Uvažujme dvojici duálních úloh

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \in \mathbb{R}^n\} = \max\{\mathbf{b}^T \mathbf{y} \mid \mathbf{A}^T \mathbf{y} = \mathbf{c}, \mathbf{y} \geq \mathbf{0}\}. \quad (14.11)$$

Mějme mnohostěn tvořený poloprostory $\mathbf{a}_i^T \mathbf{x} \geq b_i$ a vektor \mathbf{c} mířící svisle vzhůru:



Vhodíme do mnohostěnu malíčky míček, na který působí tíhová síla $-\mathbf{c}$. Pro míček se středem v bodě \mathbf{x} je číslo $\mathbf{c}^T \mathbf{x}$ přímo úměrné výšce míčku nad vodorovnou rovinou procházející počátkem,

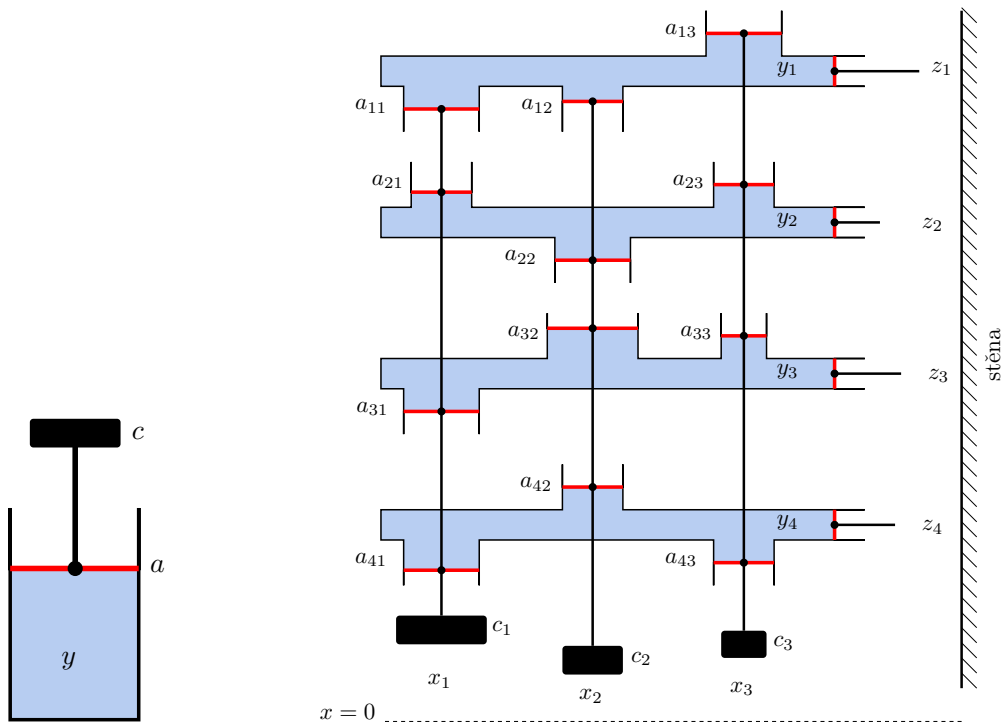
tedy potenciální energii míčku. Míček se zastaví v místě s nejmenší potenciální energií, což je nejnižší vrchol \mathbf{x}^* . Proto \mathbf{x}^* je řešením primární úlohy.

Podívejme se teď na duální úlohu. 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. Vidíme, že \mathbf{y}^* je přípustné řešení duální úlohy.

Když $\mathbf{a}_i^T \mathbf{x}^* > b_i$, míček se stěny i nedotýká a tedy síla stěny na míček je nulová, $y_i^* = 0$. Proto $y_i^*(\mathbf{a}_i^T \mathbf{x}^* - b_i) = 0$, což je podmínka komplementarity. Sečtením těchto podmínek získáme $\sum_i y_i^* \mathbf{a}_i^T \mathbf{x}^* - \sum_i b_i \mathbf{x}^* = 0$ neboli $\mathbf{c}^T \mathbf{x}^* = \mathbf{b}^T \mathbf{y}^*$ (to je vlastně Věta 14.3). Tedy platí silná dualita.

Rovnost $\mathbf{c}^T \mathbf{x}^* = \mathbf{b}^T \mathbf{y}^*$ jde vidět i jinak. Potenciální energie $\mathbf{c}^T \mathbf{x}^*$ míčku v bodě \mathbf{x}^* se rovná práci, která by se vykonala posunutím míčku do počátku. Ukážeme, že tato práce je rovna $\mathbf{b}^T \mathbf{y}^*$. Odtraňme nejprve všechny stěny, kterých se míček nedotýká. Posuňme nyní některou stěnu i rovnoběžně tak, aby procházela počátkem. Při tomto posouvání se síly stěn na míček nemění. Vzdálenost stěny od počátku je $b_i / \|\mathbf{a}_i\|_2$ (viz Cvičení 5.14). Síla stěny na míček je $y_i^* \mathbf{a}_i$ a působí 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 do počátku vykonáme práci $\sum_i b_i y_i^* = \mathbf{b}^T \mathbf{y}^*$. \square

Příklad 14.10. Známý fyzikální zákon říká, že v nádobě s nestlačitelnou kapalinou uzavřenou pístem s povrchem a , na který působí síla c , je tlak $y = c/a$ (obrázek vlevo):



Uvažujme nyní stroj na obrázku vpravo. Stroj se skládá z m nádob s nestlačitelnou kapalinou, z nichž každá je uzavřena n svislými písty a jedním vodorovným pístem. Každá m -tice svislých pístů je spojena tyčí, zakončenou dole závažím. Povrch svislého pístu v nádobě i spojeného se závažím j je $|a_{ij}|$. Pro $a_{ij} > 0$ je píst umístěn nahoře a pro $a_{ij} < 0$ dole. Výška závaží j nad referenční rovinou je x_j . Každý vodorovný píst má jednotkový povrch a je zakončen tyčí, která nemůže projít stěnou vpravo. Šířka mezery mezi tyčí a stěnou je z_i tak, že při $\mathbf{x} = \mathbf{0}$ je $z_i = -b_i$. Tíha závaží j je c_j . Tlak v nádrži i je y_i .

Ukážeme, že stroj 'řeší' levou (primární) úlohu z dvojice (14.11). Ze zachování objemu kapaliny v nádobě i plyne $z_i = a_{i1}x_1 + \dots + a_{in}x_n - b_i$. Protože vodorovné tyče nemohou projít

stěnou, musí vždy platit $z_i \geq 0$, tedy $\mathbf{z} = \mathbf{b} - \mathbf{Ax} \geq \mathbf{0}$. Potenciální energie závaží j je $c_j x_j$. Když bude stroj v rovnováze, závaží budou v takových výškách, že jejich celková potenciální energie $c_1 x_1 + \dots + c_n x_n = \mathbf{c}^T \mathbf{x}$ bude minimální.

Jaký je význam duálních omezení? Protože povrch vodorovných pístů je jednotkový, tlak y_i v nádrži i se rovná síle vodorovné tyče i na stěnu. Jelikož stěna působí silou vždy od sebe, je $y \geq \mathbf{0}$. Rovnováha sil pro svislou tyč j zní $a_{1j}y_1 + \dots + a_{mj}y_m = c_j$, tedy $\mathbf{A}^T \mathbf{y} = \mathbf{c}$.

Dle věty o silné dualitě má být v ustáleném stavu duální kritérium $\mathbf{y}^T \mathbf{b} = y_1 b_1 + \dots + y_m b_m$ rovno potenciální energii $\mathbf{c}^T \mathbf{x}$ závaží. Proč tomu tak je? Potenciální energie závaží je rovna práci, nutné na jejich posunutí do roviny $x = 0$. Tato práce se dá vykonat buď přímo posunutím závaží (což odpovídá primárnímu kritériu $\mathbf{c}^T \mathbf{x}$) nebo posunutím vodorovných tyčí do vzdáleností $-b_i$ od stěny. Druhý způsob odpovídá duálnímu kritériu. Zafixujeme-li totiž všechny vodorovné tyče kromě tyče i , při odtlačování tyče i se síla, kterou působíme na tyč, nemění. Tedy vykonáme práci $y_i b_i$. Když takto odtlačíme od stěny postupně všechny tyče, vykonáme práci $\mathbf{y}^T \mathbf{b}$.

Dle věty o komplementaritě v ustáleném stavu pro každé i platí $z_i = 0$ nebo $y_i = 0$. To je jasné, protože když se některá vodorovná tyč nedotýká stěny, je tlak v její nádobě nulový.

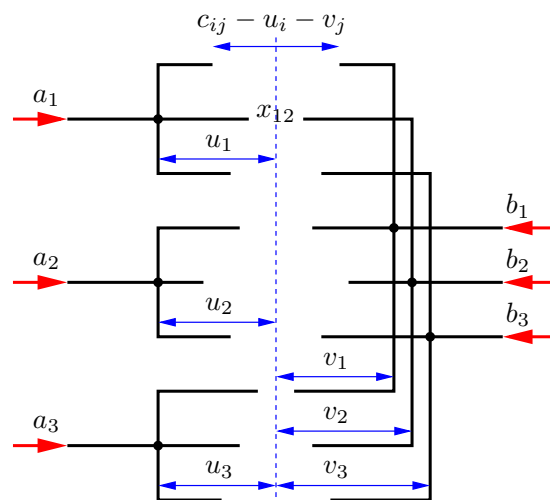
Věta o stínových cenách říká, že se změnou b_i se hodnota minimální potenciální energie mění tím více, čím je větší tlak y_i . To je ale jasné, protože čím je větší tlak, tím větší práce je třeba na posunutí tyče od stěny do vzdálenosti b_i . \square

Mohlo by se zdát, že mechanické modely z předchozích dvou příkladů dokazují větu o silné dualitě pro dvojici úloh (14.11). Dříve jsme ale řekli, že důkaz této věty je složitý. Jak je to možné? Naše fyzikální úvahy větu o silné dualitě *nedokazují*, předpokládají totiž platnost fyzikálních zákonů, které nelze matematicky dokázat ale pouze experimentálně pozorovat. Naproti tomu matematický důkaz žádné fyzikální zákony nepředpokládá.

Příklad 14.11. Zkoumejme duální úlohu k dopravní úloze (11.9). Konstrukci duálu už nebudeme popisovat podrobně, výsledná dvojice navzájem duálních úloh je

$$\begin{array}{ll} \min & \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \\ \text{za podm.} & \sum_{j=1}^n x_{ij} = a_i \\ & \sum_{i=1}^n x_{ij} = b_j \\ & x_{ij} \geq 0 \end{array} \quad \begin{array}{ll} \max & \sum_{i=1}^m a_i u_i + \sum_{j=1}^n b_j v_j \\ \text{za podm.} & u_i \in \mathbb{R} \\ & v_j \in \mathbb{R} \\ & u_i + v_j \leq c_{ij} \end{array} \quad \begin{array}{l} i = 1, \dots, m \\ j = 1, \dots, n \\ i = 1, \dots, m, j = 1, \dots, n \end{array}$$

Význam dvojice úloh ilustrujeme analogovým počítačem na obrázku (pro $m = n = 3$):



Stroj se skládá z $m + n$ pevných vidlic, z nichž každá se může vodorovně posouvat. Proměnné u_i a v_j jsou posunutí vidlic vzhledem k referenční svislé rovině² (na obrázku čárkovaně). Je-li $u_i = v_j = 0$, vzdálenost hrotů vidlic i a j je c_{ij} . Posunutí vidlic je omezeno kontakty dvojic hrotů, neboli podmínkami $c_{ij} - u_i - v_j \geq 0$. Na levé vidlice působí konstantní síly a_i , na pravé b_j (červené šipky). Působením sil se levé vidlice budou přibližovat k pravým, ale jen do té doby, než do sebe některé dvojice hrotů narazí. Proměnné x_{ij} odpovídají silám, které na sebe působí hroty vidlic (na obrázku jsou všechny tyto síly nulové, neboť žádné dva hroty se dosud nedotýkají). Dumejte, čemu odpovídá optimální hodnota úloh a proč platí silná dualita a podmínky komplementarity! \square

14.4 Cvičení

- 14.1. Dokažte, že duál duálu se rovná původní úloze. Udělejte pro (a) dvojici (14.2), (b) dvojici (14.1).
- 14.2. Pro daná čísla $c_1, \dots, c_n \in \mathbb{R}$ chceme maximalizovat $\sum_{i=1}^n c_i x_i$ za podmínek $-1 \leq x_i \leq 1$.
- Vyřešte úvahou.
 - Sestrojte duální úlohu a upravte ji do co nejjednoduššího tvaru. Vyřešte duální úlohu úvahou (musí vám vyjít stejná optimální hodnota jako u primární úlohy).
 - Napište podmínky komplementarity.
 - Najděte číselné hodnoty optimálních primárních a duálních proměnných (které si odpovídají přes podmínky komplementarity) pro $n = 3$ a $(c_1, c_2, c_3) = (-2, 3, 4)$.
- 14.3. 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. Kde to má smysl, pokuste se interpretovat duální úlohu a věty o dualitě, podobně jako v Příkladu 14.5.
- lineární program ze Cvičení 13.8
 - $\min_{x \in \mathbb{R}} \max_{i=1}^n |a_i - x|$ (střed intervalu)
 - úloha (11.13) (přibližné řešení pře určené lin. soustavy ve smyslu max-normy)
 - úloha (11.15) (přibližné řešení pře určené lin. soustavy ve smyslu 1-normy)
 - všechny úlohy ze Cvičení 11.3
 - úloha vzniklá ve Cvičení 11.10 o kladce se závažími
 - minimalizace maxima afinních funkcí (viz §11.1.1):
 - $\min_{x_1, x_2 \in \mathbb{R}} \max \{ 2x_1 - x_2 - 3, 1 - x_1, x_2 - 2, x_1 + x_2 \}$
 - $\min_{\mathbf{x} \in \mathbb{R}^n} \max_{i=1}^m (\mathbf{a}_i^T \mathbf{x} + b_i)$

²Chování stroje samozřejmě není polohou této referenční roviny ovlivněno. Jak se to projevuje algebraicky v primární a duální úloze?

14.4. Dokažte bez užití algoritmu na řešení LP, že $\mathbf{x} = (1, 1, 1, 1)$ je optimální řešení úlohy

$$\begin{aligned} \min \quad & [47 \ 93 \ 17 \ -93] \mathbf{x} \\ \text{za podm.} \quad & \begin{bmatrix} -1 & -6 & 1 & 3 \\ -1 & -2 & 7 & 1 \\ 0 & 3 & -10 & -1 \\ -6 & -11 & -2 & 12 \\ 1 & 6 & -1 & -3 \end{bmatrix} \mathbf{x} \leq \begin{bmatrix} -3 \\ 5 \\ -8 \\ -7 \\ 4 \end{bmatrix} \end{aligned}$$

Nápověda a řešení

14.1. Pravou úlohu nejdříve převedeme na tvar levé úlohy a pak k ní napíšeme duální úlohu. Ukážeme jen pro (14.2):

$$\begin{aligned} -\min \quad & (-\mathbf{b})^T \mathbf{y} \\ \text{za podm.} \quad & (-\mathbf{A}^T) \mathbf{y} \geq -\mathbf{c} \\ & \mathbf{y} \geq \mathbf{0} \end{aligned} \qquad \begin{aligned} -\max \quad & (-\mathbf{c})^T \mathbf{x} \\ \text{za podm.} \quad & \mathbf{x} \geq \mathbf{0} \\ & (-\mathbf{A}) \mathbf{x} \leq -\mathbf{b} \end{aligned} \quad (14.12)$$

Levá úloha (14.12) je ekvivalentní pravé úloze (14.2), pravá úloha (14.12) je ekvivalentní levé úloze (14.2).

14.2.a) Optimální hodnota je $\sum_{i=1}^n |c_i|$ (viz Cvičení 11.3)

14.2.b) Duální úloha je $\min \{ \sum_i (u_i + v_i) \mid u_i, v_i \geq 0, v_i - u_i = c_i \}$. Vyřešme ji úvahou. Nejdříve uvažujme každé i zvlášť a ukažme, že $\min \{ u + v \mid u, v \geq 0, v - u = c \} = |c|$. Podmínka $v - u = c$ zůstává v platnosti, odečteme-li od u, v libovolné číslo. Pokud $u + v$ má být minimální, musí se od u, v odečíst co největší číslo tak, aby platilo $u, v \geq 0$. Tedy jedno z čísel u, v bude nulové a tedy optimální hodnota bude $|c|$. Optimální hodnota celé duální úlohy bude tedy $\sum_i |c_i|$.

14.2.c) Pro každé i platí $x_i = -1$ nebo $u_i = 0$. Pro každé i platí $x_i = 1$ nebo $v_i = 0$.

14.2.d) $(x_1, x_2, x_3) = (-1, 1, 1)$, $(u_1, u_2, u_3) = (2, 0, 0)$, $(v_1, v_2, v_3) = (0, 3, 4)$.

14.3.f) Duál: $\max \{ \sum_{i=1}^n y_i d_i + \sum_{i=1}^{n'} y'_i d'_i \mid \sum_{i=1}^n y_i = \sum_{i=1}^{n'} y'_i, y_i \leq m_i, y'_i \leq m'_i, y_i, y'_i \geq 0 \}$.
Podmínky komplementarity: $z_i(y_i - m_i) = 0$, $z'_i(y'_i - m'_i) = 0$, $(z_i - d_i - x)y_i = 0$, $(z'_i - d'_i + x)y'_i = 0$.

14.3.g) (i) Nejprve převedeme na LP: $\min z$ z.p. $2x_1 - x_2 - 3 \leq z$, $1 - x_1 \leq z$, $x_2 - 2 \leq z$, $x_1 + x_2 \leq z$. Duál k tomuto LP je: $\max -3u_1 + u_2 - 2u_3$ z.p. $2u_1 - u_2 + u_4 = 0$, $-u_1 + u_3 + u_4 = 0$, $u_1 + \dots + u_4 = 1$, $u_1, \dots, u_4 \geq 0$.

14.4. Postupujte podle Příkladu 14.3.

Část IV

Konvexní optimalizace

Kapitola 15

Konvexní funkce

Funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je **konvexní** na konvexní množině $X \subseteq \mathbb{R}^n$, jestliže

$$\mathbf{x} \in X, \mathbf{y} \in X, 0 \leq \alpha \leq 1 \implies f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) \leq (1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{y}). \quad (15.1)$$

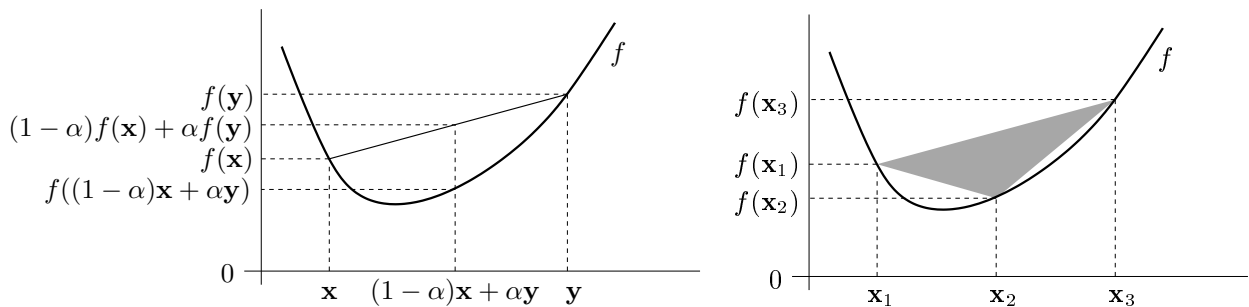
Funkce f je **konkávní** na množině X , jestliže je funkce $-f$ konvexní na množině X . Rozlišujte pojmy *konvexní množina* a *konvexní funkce*, jde o různé věci. Všimněte si, že množina X musí být konvexní (pojem konvexní funkce na nekonvexní množině není definován). Pokud X je celý definiční obor funkce f , odkaz na X můžeme vynechat a říkáme jen, že funkce f je konvexní.

Podmínku (15.1) lze zobecnit pro více než dva body: funkce f je konvexní, právě když

$$\left. \begin{array}{l} \mathbf{x}_1, \dots, \mathbf{x}_k \in X \\ \alpha_1, \dots, \alpha_k \geq 0 \\ \alpha_1 + \dots + \alpha_k = 1 \end{array} \right\} \implies f(\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k) \leq \alpha_1 f(\mathbf{x}_1) + \dots + \alpha_k f(\mathbf{x}_k). \quad (15.2)$$

Podmínka (15.2) je známá jako **Jensenova nerovnost**. Podmínka (15.1) je očividně speciální případ podmínky (15.2). Naopak lze dokázat, že (15.1) implikuje (15.2). 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 mnohoúhelník 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 podmínky (15.1), že funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ daná jako $f(\mathbf{x}) = \max_{i=1}^n x_i$ (tedy funkční hodnota je maximum ze složek vektoru \mathbf{x}) je konvexní. Máme dokázat, že pro

každé \mathbf{x}, \mathbf{y} a $0 \leq \alpha \leq 1$ platí

$$f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) = \max_i((1-\alpha)x_i + \alpha y_i) \quad (15.3a)$$

$$\leq \max_i(1-\alpha)x_i + \max_i \alpha y_i \quad (15.3b)$$

$$= (1-\alpha) \max_i x_i + \alpha \max_i y_i \quad (15.3c)$$

$$= (1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{y}) \quad (15.3d)$$

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 podmínky (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}_{++} (kde \mathbb{R}_{++} označuje množinu kladných reálných čísel, viz §1.1.1). Napišme pro tuto funkci Jensenovu nerovnost (15.2) (jelikož funkce je konkávní a ne konvexní, musíme v Jensenově nerovnosti obrátit směr 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 nikdy menší než geometrický. \square

Příklad 15.4. Uved'me 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. Funkce $f(x) = x \log x$ je konvexní na \mathbb{R}_{++} (nebo i na \mathbb{R}_+ , pokud dodefinujeme $0 \log 0 = 0$, což se často dělá, protože $\lim_{x \rightarrow 0+} x \log x = 0$). Tato funkce se vyskytuje např. jako jeden člen ve vzorci pro Shannonovu entropii náhodné veličiny, $H(\mathbf{x}) = -\sum_i x_i \log x_i$.

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 $f(\mathbf{x}) = (x_1 \cdots x_n)^{1/n}$ je konkávní na \mathbb{R}_+^n .
11. Každá norma je konvexní funkce, neboť pro každé $\alpha, \beta \geq 0$ máme

$$\|\alpha \mathbf{x} + \beta \mathbf{y}\| \leq \|\alpha \mathbf{x}\| + \|\beta \mathbf{y}\| = \alpha \|\mathbf{x}\| + \beta \|\mathbf{y}\|,$$

kde nerovnost plyne z trojúhelníkové nerovnosti a rovnost z homogenity (viz §11.5.1).

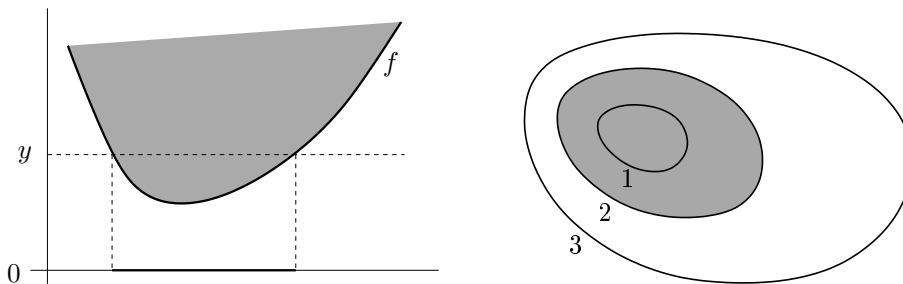
Načrtněte vrstevnice a grafy těchto funkcí (v případě více proměnných jen pro $n \in \{1, 2\}$)! \square

15.1 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 rovnosti nerovností. 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**¹ 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.

Věta 15.1. Každá subkontura konvexní funkce je konvexní množina.

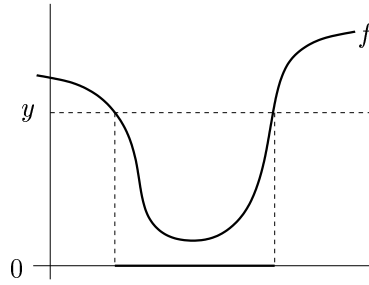
Důkaz. Předpokládejme, že body \mathbf{x}_1 a \mathbf{x}_2 patří do subkontury konvexní funkce f , tedy $f(\mathbf{x}_1) \leq y$ a $f(\mathbf{x}_2) \leq y$. Pro každé $0 \leq \alpha \leq 1$ platí

$$f((1 - \alpha)\mathbf{x}_1 + \alpha\mathbf{x}_2) \leq (1 - \alpha)f(\mathbf{x}_1) + \alpha f(\mathbf{x}_2) \leq (1 - \alpha)y + \alpha y = y,$$

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 $(1 - \alpha)\mathbf{x}_1 + \alpha\mathbf{x}_2$ patří do subkontury. Dle (12.1) je tedy subkontura konvexní množina. \square

¹Slovo 'subkontura' je pokus o český překlad anglického 'sublevel set'.

Obrácená implikace ve Větě 15.1 neplatí: existuje funkci, která není konvexní a jejíž každá subkontura je konvexní množina². Příklad je na obrázku:



Věta 15.2. Funkce je konvexní, právě 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

$$f((1 - \alpha)\mathbf{x}_1 + \alpha\mathbf{x}_2) \leq (1 - \alpha)f(\mathbf{x}_1) + \alpha f(\mathbf{x}_2) \leq (1 - \alpha)y_1 + \alpha y_2,$$

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 $(1 - \alpha)(\mathbf{x}_1, y_1) + \alpha(\mathbf{x}_2, y_2)$ patří do epigrafu, který je proto konvexní množina.

Obráceně předpokládejme, že epigraf funkce f je konvexní množina. Tedy pokud body (\mathbf{x}_1, y_1) a (\mathbf{x}_2, y_2) patří do epigrafu, pak také bod $(1 - \alpha)(\mathbf{x}_1, y_1) + \alpha(\mathbf{x}_2, y_2)$ 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

$$f((1 - \alpha)\mathbf{x}_1 + \alpha\mathbf{x}_2) \leq (1 - \alpha)y_1 + \alpha y_2 = (1 - \alpha)f(\mathbf{x}_1) + \alpha f(\mathbf{x}_2),$$

proto je funkce f konvexní. □

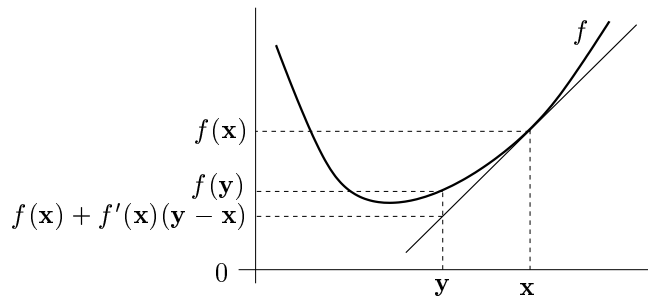
15.2 Konvexita diferencovatelných funkcí

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 15.3 (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 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 (8.14b)) je všude (tj. pro každé \mathbf{y}) menší nebo roven funkci f :



²Funkce, jejíž každá subkontura je konvexní množina, se nazývá *kvazikonvexní*. Kvazikonvexní funkce jsou užitečné, ale zdaleka ne tak jako konvexní funkce.

Věta 15.4 (Podmínka druhého řádu). Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je dvakrát diferencovatelná na konvexní množině $X \subseteq \mathbb{R}^n$. Funkce f je konvexní na množině X , právě 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 15.4. To je triviální, protože Hessián je $f''(\mathbf{x}) = 2\mathbf{A}$ a tedy je pozitivně semidefinitní.
- Dokažme konvexitu z Věty 15.3. 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 $\mathbf{x}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{A} \mathbf{y} + \mathbf{y}^T \mathbf{A} \mathbf{y} \geq 0$. Platí³

$$\mathbf{x}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{A} \mathbf{y} + \mathbf{y}^T \mathbf{A} \mathbf{y} = (\mathbf{x} - \mathbf{y})^T \mathbf{A} (\mathbf{x} - \mathbf{y}), \quad (15.5)$$

což je nezáporné pro každé \mathbf{x}, \mathbf{y} , protože \mathbf{A} je pozitivně semidefinitní.

- Dokážme konvexitu z podmínky (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

$$[(1 - \alpha)\mathbf{x} + \alpha\mathbf{y}]^T \mathbf{A} [(1 - \alpha)\mathbf{x} + \alpha\mathbf{y}] \leq (1 - \alpha)\mathbf{x}^T \mathbf{A} \mathbf{x} + \alpha\mathbf{y}^T \mathbf{A} \mathbf{y}$$

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

15.3 Operace zachovávající konvexitu funkcí

Operace zachovávající konvexitu funkcí umožňují z jednoduchých konvexních funkcí získat složitější. Konvexitu složitější funkce je často snadnější dokázat pomocí těchto operací než z podmínky (15.1) nebo Vět 15.3 a 15.4. Dále uvedeme příklady takových operací.

15.3.1 Nezáporná lineární kombinace

Věta 15.5. Jsou-li $g_1, \dots, g_k: \mathbb{R}^n \rightarrow \mathbb{R}$ konvexní funkce a $\alpha_1, \dots, \alpha_k \geq 0$, pak funkce

$$f = \alpha_1 g_1 + \dots + \alpha_k g_k \quad (15.6)$$

je konvexní.

Důkaz je snadný z podmínky (15.1) (ponecháváme jako cvičení).

Speciálně, jsou-li f a g konvexní funkce, pak $f + g$ je konvexní. Obráceně to ale neplatí: může se stát, že f nebo g nejsou konvení a $f + g$ konvexní je. Např. funkce x^3 není konvexní, ale funkce $x^3 - x^3$ (tedy konstantní nulová funkce) konvexní je.

³Všimněte si, že pro $n = 1$ a $\mathbf{A} = 1$ se rovnost (15.5) zjednoduší na známé $x^2 - 2xy + y^2 = (x - y)^2$.

15.3.2 Skládání funkcí

Složení konvexních funkcí nemusí být konvexní funkce.

Příklad 15.6. Funkce $f, g: \mathbb{R} \rightarrow \mathbb{R}$ dané jako $f(x) = |x - 1|$ a $g(x) = |x|$ jsou konvexní. Funkce $(g \circ f)(x) = g(f(x)) = ||x| - 1|$ ale není konvexní (nakreslete si její graf!). \square

Věta 15.6. Necht' funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ je konvexní. Necht' funkce $g: \mathbb{R} \rightarrow \mathbb{R}$ je konvexní a neklesající. Pak složená funkce $g \circ f$ (daná předpisem $(g \circ f)(x) = g(f(x))$) je konvexní.

Důkaz. Pro každé $x, y \in \mathbb{R}$ a $0 \leq \alpha \leq 1$ máme

$$g(f((1 - \alpha)x + \alpha y)) \leq g((1 - \alpha)f(x) + \alpha f(y)) \leq (1 - \alpha)g(f(x)) + \alpha g(f(y)).$$

První nerovnost platí, protože f je konvexní a g je neklesající. Druhá nerovnost platí, protože g je konvexní. \square

Obecněji můžeme zkoumat složené zobrazení $(g \circ \mathbf{f})(\mathbf{x}) = g(\mathbf{f}(\mathbf{x}))$, kde $\mathbb{R}^n \xrightarrow{\mathbf{f}} \mathbb{R}^m \xrightarrow{g} \mathbb{R}$. Existuje analogie Věty 15.6 pro tento případ, je ale dosti komplikovaná a nebudeme ji uvádět. Uvedeme jen důležitý případ, kdy $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ je afinní zobrazení.

Věta 15.7. Necht' funkce $g: \mathbb{R}^m \rightarrow \mathbb{R}$ je konvexní. Necht' $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{b} \in \mathbb{R}^m$. Pak funkce $h(\mathbf{x}) = g(\mathbf{A}\mathbf{x} + \mathbf{b})$ je konvexní.

Důkaz. Pro každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ a $0 \leq \alpha \leq 1$ platí

$$\begin{aligned} h((1 - \alpha)\mathbf{x} + \alpha\mathbf{y}) &= g(\mathbf{A}[(1 - \alpha)\mathbf{x} + \alpha\mathbf{y}] + \mathbf{b}) \\ &= g((1 - \alpha)(\mathbf{A}\mathbf{x} + \mathbf{b}) + \alpha(\mathbf{A}\mathbf{y} + \mathbf{b})) \\ &\leq (1 - \alpha)g(\mathbf{A}\mathbf{x} + \mathbf{b}) + \alpha g(\mathbf{A}\mathbf{y} + \mathbf{b}) \\ &= (1 - \alpha)h(\mathbf{x}) + \alpha h(\mathbf{y}). \end{aligned}$$

\square

Příklad 15.7. 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 15.7 vezmeme $\mathbf{A} = \begin{bmatrix} \mathbf{I} & -\mathbf{I} \end{bmatrix} \in \mathbb{R}^{n \times (2n)}$ a $\mathbf{b} = \mathbf{0}$. \square

15.3.3 Maximum

Nejzajímavější operace zachovávající konvexitu funkcí je ovšem maximum.

Věta 15.8. Necht' I je libovolná množina a $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$, $i \in I$, jsou konvexní funkce. Pak

$$f(\mathbf{x}) = \max_{i \in I} g_i(\mathbf{x}) \tag{15.7}$$

je konvexní funkce, kde předpokládáme, že pro každé \mathbf{x} maximum existuje⁴.

⁴Pokud 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 průnik epigrafů funkcí g_i , neboť

$$\begin{aligned} \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n, \max_{i \in I} g_i(\mathbf{x}) \leq y\} &= \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n, (\forall i \in I)(g_i(\mathbf{x}) \leq y)\} \\ &= \bigcap_{i \in I} \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n, g_i(\mathbf{x}) \leq y\}, \end{aligned}$$

kde jsme využili ekvivalence (11.6). Protože funkce g_i jsou konvexní, dle Věty 15.2 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 15.2 je tedy funkce f konvexní. \square

Příklad 15.8. Konvexitu funkce $f(\mathbf{x}) = \max_{i=1}^n x_i$ jsme již v Příkladu 15.1 dokázali z podmínky (15.1). Ovšem je mnohem pohodlnější použít Větu 15.8. Máme $g_i(\mathbf{x}) = x_i$. Funkce g_i jsou lineární, tedy konvexní. Takže funkce $f(\mathbf{x}) = \max_{i=1}^n g_i(\mathbf{x})$ je konvexní. \square

Příklad 15.9. 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 15.10. 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 nejvzdálenějšího bodu množiny C (zde předpokládáme, že maximum existuje). Dle Věty 15.7 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 15.11. Mějme funkci

$$f(\mathbf{c}) = \max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b}\},$$

která 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 15.12. 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 minimum lineárních funkcí. \square

15.4 Cvičení

15.1. Pro každou funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ dokažte z podmínky (15.1), které z těchto čtyř tvrzení platí (a pro jaké n): funkce je konvexní, konkávní, konvexní i konkávní, ani konvexní ani konkávní.

- a) $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b$
- b) $f(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$
- c) $f(\mathbf{x}) = \text{aritmetický průměr čísel } x_1, \dots, x_n$
- d) $f(\mathbf{x}) = \text{median}_{i=1}^n x_i$ (medián čísel x_1, \dots, x_n)
- e) $f(\mathbf{x}) = \min_{i=1}^n |x_i|$
- f) $f(\mathbf{x}) = \text{součet dvou nejmenších čísel z čísel } x_1, \dots, x_n$.

15.2. Dokažte konvexitu či konkavitu funkcí z Příkladu 15.4. Můžete použít podmínku (15.1) a věty z §15.2 a §15.3.

15.3. Pro každou funkci dokažte, které z těchto čtyřech tvrzení platí: funkce je konvexní, konkávní, konvexní i konkávní, ani konvexní ani konkávní. Můžete použít podmínku (15.1) a věty z §15.2 a §15.3.

- a) $f(x) = e^{x^2}$
- b) $f(x) = e^{-x^2}$
- c) $f(x, y) = |x - y|$
- d) $f(x, y) = -y$
- e) $f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2^2$
- f) $f(\mathbf{x}) = \sum_{i=1}^n x_i \log x_i$ na množině \mathbb{R}_{++}^n
- g) $f(\mathbf{x}) = \sum_{i=1}^k \log(b_i - \mathbf{a}_i^T \mathbf{x})$ na množině $X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T \mathbf{x} < b_i, i = 1, \dots, k\}$
- h) $f(\mathbf{x}) = \max_{i=1}^n x_i + \min_{i=1}^n x_i$
- i) $f(\mathbf{x}) = \max_{i=1}^n x_i - \min_{i=1}^n x_i$
- j) $f(\mathbf{x}) = \max \left\{ \left| \sum_{i=1}^n x_i - 1 \right|, \sum_{i=1}^n |x_i| - 1 \right\}$
- k) $(\star) f(\mathbf{x}) = \text{součet } k \text{ největších čísel } x_1, \dots, x_n \text{ (kde } k \leq n \text{ je dáno)}$

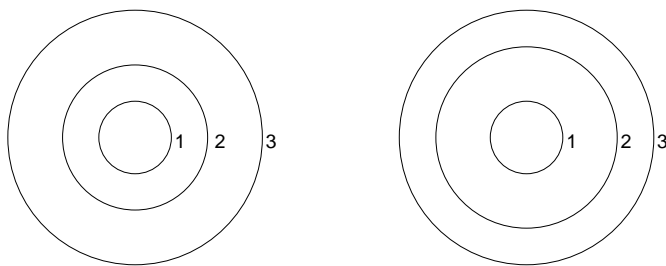
15.4. Máme funkci jedné proměnné $f(x) = (x^2 - a)^2$. Pro jaké hodnoty parametru $a \in \mathbb{R}$ je tato funkce konvexní? Načrtněte graf funkce pro nějaké a , pro které funkce není konvexní.

15.5. Robustní prokládání přímky množinou bodů $(\mathbf{x}_i, y_i) \in \mathbb{R}^n \times \mathbb{R}$ (pro $i = 1, \dots, m$) 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.

15.6. Každý z obrázků zobrazuje některé vrstevnice funkce dvou proměnných a jejich výšky. Je možné, aby funkce, která má tyto vrstevnice, byla konvexní? Dokažte z podmínky (15.1).



- 15.7. Dokažte, že elipsoid $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{A} \mathbf{x} \leq 1 \}$ (kde \mathbf{A} je pozitivně definitní, viz §6.3.1) je konvexní množina.
- 15.8. Dokažte, že množina $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}, \prod_{i=1}^n x_i \geq 1 \}$ je konvexní (symbol $\prod_i x_i$ značí součin čísel x_1, \dots, x_n).
- 15.9. Co je subkontura výšky 2 funkce jedné proměnné $f(x) = x^2 - x$?
- 15.10. Dokažte, že účelové funkce vystupující v následujících úlohách jsou nekonvexní:
- Příklad 9.10
 - Cvičení 9.5

Nápověda a řešení

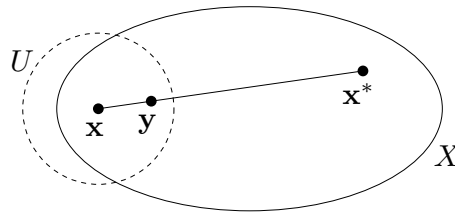
- 15.1.a) Konvexní i konkávní, nerovnost (15.1) platí s rovností.
- 15.1.b) Je konvexní, není konkávní.
- 15.1.c) Konvexní i konkávní, nerovnost (15.1) platí s rovností.
- 15.1.d) Pro $n \leq 2$ konvexní i konkávní, pro $n > 2$ ani konvexní ani konkávní.
- 15.1.e) Pro $n = 1$ je funkce konvexní. Pro $n = 2$ není konvexní, neboť (15.1) není splněna např. pro $\mathbf{x} = (1, 0)$, $\mathbf{y} = (0, 1)$, $\alpha = \frac{1}{2}$. Pro $n > 2$ také není konvexní, což plyne z nekonvexity pro $n = 2$, protože můžeme zvolit $\mathbf{x} = (1, 0, 0, 0, \dots)$ a $\mathbf{y} = (0, 1, 1, 1, \dots)$.
Funkce není konkávní pro žádné n . Pro $n = 1$ dokážeme z (15.1) volbou $\mathbf{x} = -1$, $\mathbf{y} = 1$, $\alpha = \frac{1}{2}$. Pro $n > 1$ můžeme vektory opět doplnit opakováním poslední číslice.
- 15.1.f) Dokážeme, že pro $n = 3$ funkce není konvexní. Vezmeme $\mathbf{x} = (1, 2, 3)$, $\mathbf{y} = (3, 2, 1)$, $\alpha = \frac{1}{2}$. Pak $f(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}) = f(2, 2, 2) = 4 \not\leq (3 + 3)/2 = \frac{1}{2}f(1, 2, 3) + \frac{1}{2}f(3, 2, 1) = \frac{1}{2}f(\mathbf{x}) + \frac{1}{2}f(\mathbf{y})$.
- 15.3.j) Je konvexní. Absolutní hodnota je konvexní funkce, jejich součet také, maximum konvexních funkcí je konvexní funkce.
- 15.6. V podmínce (15.1) zvolte \mathbf{x}, \mathbf{y} na vrstevnicích výšky 1 a 3. Zvolte chytře α . Odpovědi: ne, ano.
- 15.7. Funkce $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ je konvexní, viz Příklad 15.5. Elipsoid je subkontura této funkce, tedy je to konvexní množina.
- 15.8. Po zlogaritmování jde podmínka $\prod_i x_i \geq 1$ psát jako $\sum_i \log x_i \geq 0$ (to můžeme udělat, protože logaritmus je rostoucí funkce a proto nemění znaménko nerovnosti). Funkce $f(\mathbf{x}) = -\sum_i \log x_i$ na množině \mathbb{R}_{++}^n a tedy její subkontura je konvexní množina.
- 15.9. Interval $[-1, 2]$.

Kapitola 16

Konvexní optimalizační úlohy

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 lokálního minima (viz §8.10) 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 $\mathbf{y} = (1 - \alpha)\mathbf{x} + \alpha\mathbf{x}^*$ leží v U . Protože je množina X konvexní, leží bod \mathbf{y} zároveň i v X . Je

$$f(\mathbf{y}) = f((1 - \alpha)\mathbf{x} + \alpha\mathbf{x}^*) \leq (1 - \alpha)f(\mathbf{x}) + \alpha f(\mathbf{x}^*) < (1 - \alpha)f(\mathbf{x}) + \alpha f(\mathbf{x}) = f(\mathbf{x}).$$

Ale tvrzení $f(\mathbf{y}) < f(\mathbf{x})$ je ve sporu s předpokladem, že \mathbf{x} je lokální minimum. \square

Minimalizaci konvexní funkce na konvexní množině se říká **konvexní optimalizační úloha**. Pro takovou úlohu nám tedy stačí najít libovolné lokální minimum, abychom našli globální minimum.

16.1 Příklady nekonvexních úloh

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 nekonvexní úloha může mít lokálních minim velmi mnoho.

Příklad 16.1. Zopakujme úlohu (7.1): pro danou čtvercovou matici \mathbf{A} minimalizujeme funkci $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ na množině $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{x} = 1\}$. Tato úloha není konvexní, neboť množina přípustných řešení není konvexní (je to n -rozměrná sféra). My ale víme, že globální optimum úlohy lze najít pomocí spektrálního rozkladu. \square

Zde jsme měli štěstí: najít globální optimum úlohy (7.1) je snadné, i když úloha není konvexní. To je ale výjimka – typicky je nalezení globálního minima nekonvexní úlohy velmi těžké.

Příklad 16.2. Uvedme příklad, na kterém bude na první pohled vidět, že nekonvexní úloha může mít velmi mnoho lokálních minim. 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, úloha má 2^n lokálních maxim.

V tomto symetrickém případě globální maximum snadno najdeme úvahou. Uvažme však mírně obecnější úlohu

$$\min\{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x} \in [-1, 1]^n\}, \quad (16.2)$$

kde $\mathbf{A} \in \mathbb{Z}^{n \times n}$ (tedy matice má celočíselné prvky). Je známo, že vyřešení (tj. nalezení globálního maxima) této úlohy je prakticky nemožné (přesněji, je NP-těžké, srov. §11.6). \square

Příklad 16.3. Mějme m bodů v rovině $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^d$. Úkolem je rozmístit dalších n bodů $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$ tak, aby nejdelší vzdálenost bodu \mathbf{a}_i k nejbližšímu bodu \mathbf{x}_j byla nejmenší. Tedy minimalizujeme účelovou funkci

$$f(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{i=1}^m \min_{j=1}^n \|\mathbf{a}_i - \mathbf{x}_j\| \quad (16.3)$$

přes vektory $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$. Máme $f: \mathbb{R}^{dn} \rightarrow \mathbb{R}$, tedy přesněji můžeme říci, že minimalizujeme funkci f přes jediný vektor $(\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{dn}$.

Úloha je známá jako *shlukování*. Jako motivaci si představme optimální rozmístění cisteren ve vesnici, kde občas neteče voda. Zde máme $n = 2$, \mathbf{a}_i jsou souřadnice domů a \mathbf{x}_j jsou souřadnice cisteren. Chceme, aby průměrná vzdálenost obyvatele k nejbližší cisterně byla co nejmenší.

Je funkce (16.3) konvexní? Vezměme jednoduchý případ $d = 1$, $m = 1$, $n = 2$, $a_1 = 0$. Pak (16.3) má tvar $f(x_1, x_2) = \min\{|x_1|, |x_2|\}$. Snadno dokážeme (proved'tel), že toto není konvexní funkce. Bez důkazu uvedme, že (nepřekvapivě) funkce není konvexní ani pro větší d, m, n .

Je známo, že neexistuje algoritmus, který by našel optimální řešení úlohy (16.3) v čase, který je polynomiální funkcí čísel d, m, n . V praktické situaci tedy nezbyvá nic jiného, než použít algoritmus, který najde pouze přibližné optimum. Takovým algoritmem je např. *k-means*. \square

Příklad 16.4. Úloha celočíselného programování (11.20) je nekonvexní, protože množina $\{\mathbf{x} \in \{0, 1\}^n \mid \mathbf{A} \mathbf{x} \geq \mathbf{b}\}$ jejích přípustných řešení je nekonvexní (srov. Cvičení .g). \square

16.2 Konvexní optimalizační úloha ve standardním tvaru

Uvažujme nyní obecnou úlohu spojitě optimalizace ve standardním tvaru (1.9),

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\} \quad (16.4)$$

neboli

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

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 g_1, \dots, g_m jsou konvexní a zobrazení \mathbf{h} je afinní, je postačující ale nikoliv nutná pro konvexitu množiny přípustných řešení.

Příklad 16.5. Následující dva tvary definují stejnou množinu (promyslete!):

$$\{\mathbf{x} \in \mathbb{R}^2 \mid x_1/(1+x_2^2) \leq 0, (x_1+x_2)^2 = 0\} = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 \leq 0, x_1+x_2 = 0\}.$$

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. \square

Úloze tvaru (16.4), ve které jsou funkce f, g_1, \dots, g_m konvexní a zobrazení \mathbf{h} afinní, říkáme **konvexní optimalizační úloha ve standardním tvaru**.

16.3 Ekvivalentní transformace úlohy

Dvě úlohy ve tvaru (16.4) nazveme **ekvivalentní**, když se z množiny optimálních řešení jedné dá ‘snadno’ 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:

- *Změna proměnných.* Necht’ $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ je bijektivní zobrazení (viz §1.1.2). Pak úloha (16.4) 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 (viz §15.3.2).

- *Monotónní transformace účelové funkce.* Necht’ $\psi: \mathbb{R} \rightarrow \mathbb{R}$ je 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 16.6. 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 LP (viz §11.1), úloha (16.4) je ekvivalentní úloze

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{s} \in \mathbb{R}^m, \mathbf{s} \geq \mathbf{0}, \mathbf{g}(\mathbf{x}) + \mathbf{s} = \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}.$$

Tato 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 (16.4) je ekvivalentní úloze

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

V nové úloze vlastně hledáme nejnižší bod epigrafu funkce f na množině přípustných řešení. Tato transformace zachovává konvexitu úlohy. Plyne z ní, že každou konvexní úlohu lze převést na konvexní úlohu s lineární účelovou funkcí.

16.4 Třídy konvexních optimalizačních úloh

Optimalizační úlohy ve tvaru (16.4) se taxonomizují podle druhu funkcí f, g_i, h_i . Pro každou třídu existují specializované algoritmy schopné najít lokální minimum (v případě konvexní úlohy tedy globální minimum). Přehled dostupných implementací takových algoritmů je možno najít např. na <http://www.neos-guide.org>.

16.4.1 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.

16.4.2 Kvadratické programování (QP)

V *kvadratickém programování* jsou funkce g_i, h_i afinní a funkce f je kvadratická konvexní, tedy $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c$ kde \mathbf{A} je pozitivně semidefinitní (viz Příklad 15.5).

Příklad 16.7. Při řešení soustavy ve smyslu nejmenších čtverců (§5.1) počítáme konvexní QP bez omezení $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A} \mathbf{x} - \mathbf{b}\|_2^2$. 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 16.8. Hledání řešení přeuročené lineární soustavy s nejmenší normou (§5.2) vede na úlohu $\min \{ \mathbf{x}^T \mathbf{x} \mid \mathbf{A} \mathbf{x} = \mathbf{b} \}$, což je konvexní QP. \square

Příklad 16.9. Vzdálenost bodu $\mathbf{y} \in \mathbb{R}^n$ od konvexního mnohostěnu $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} \leq \mathbf{b} \}$ je optimální hodnota kvadratického programu

$$\min \{ \|\mathbf{y} - \mathbf{x}\|_2^2 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A} \mathbf{x} \leq \mathbf{b} \}. \quad \square$$

Příklad 16.10. 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

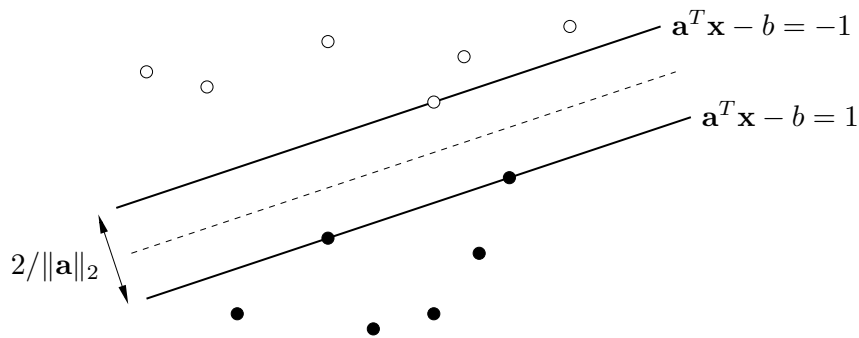
$$y_i(\mathbf{a}^T \mathbf{x}_i - b) > 0, \quad i = 1, \dots, m. \quad (16.5)$$

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 (16.5) můžeme ekvivalentně psát jako

$$y_i(\mathbf{a}^T \mathbf{x}_i - b) \geq 1, \quad i = 1, \dots, m. \quad (16.6)$$

Hledáme-li libovolnou oddělující nadrovinu, stačí nám najít libovolné řešení soustavy nerovnic (16.6), což vede na úlohu LP.

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 5.14), že šířka pásu je $2/\|\mathbf{a}\|_2$. V úloze *support vector machine* (SVM) hledáme oddělující nadrovinu která maximalizuje šířku pásu, tedy minimalizuje $\|\mathbf{a}\|_2^2 = \mathbf{a}^T \mathbf{a}$ za podmínek (16.6). To je konvexní úloha QP. \square

16.4.3 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.

16.4.4 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

$$g_i(\mathbf{x}) = \|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\|_2 - (\mathbf{c}_i^T \mathbf{x} + d_i). \quad (16.7)$$

Tedy úloha SOCP má tvar (vynecháváme afinní omezení $h_i(\mathbf{x}) = 0$)

$$\begin{aligned} \min \quad & \mathbf{e}^T \mathbf{x} \\ \text{za podmínek} \quad & \|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\|_2 \leq \mathbf{c}_i^T \mathbf{x} + d_i, \quad i = 1, \dots, m. \end{aligned}$$

Funkce g_i jsou konvexní (nebot' norma je konvexní funkce, dále viz Věta 15.7). 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 konvexní QCQP jsou speciální případy SOCP.

Příklad 16.11. Jsou dány body $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ a chceme minimalizovat funkci

$$f(\mathbf{x}) = \sum_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\|_2 \quad (16.8)$$

na množině \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 a po zavedení pomocných proměnných z_i (podobná úprava jako v §11.1.1) ji lze formulovat jako SOCP:

$$\begin{aligned} \min \quad & z_1 + \dots + z_m \\ \text{za podmínky} \quad & \|\mathbf{x} - \mathbf{a}_i\|_2 \leq z_i, \quad i = 1, \dots, m. \end{aligned}$$

Pro případ $n = 2$ má úloha jednoduchý mechanický model¹. 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 $f(\mathbf{x})$. \square

16.4.5 Semidefinitní programování (SDP)

Věta 16.2. Pro každé $n \in \mathbb{N}$ je množina všech pozitivně semidefinitních matic rozměru $n \times n$ konvexní kužel.

Důkaz. Necht' $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ jsou takové, že pro každé $\mathbf{x} \in \mathbb{R}^n$ platí $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$ a $\mathbf{x}^T \mathbf{B} \mathbf{x} \geq 0$. Pak pro každé $\mathbf{x} \in \mathbb{R}^n$ a $\alpha, \beta \geq 0$ platí $\mathbf{x}^T (\alpha \mathbf{A} + \beta \mathbf{B}) \mathbf{x} = \alpha \mathbf{x}^T \mathbf{A} \mathbf{x} + \beta \mathbf{x}^T \mathbf{B} \mathbf{x} \geq 0$. \square

Konvexní kužel je konvexní množina. To umožňuje formulovat třídu konvexních úloh známou jako *semidefinitní programování* (SDP). Jednou z možných formulací je²

$$\begin{aligned} \min \quad & \mathbf{C} \cdot \mathbf{X} \\ \text{za podmínky} \quad & \mathbf{X} \text{ je pozitivně semidefinitní} \\ & \mathbf{A}_i \cdot \mathbf{X} = b_i, \quad i = 1, \dots, m. \end{aligned} \tag{16.9}$$

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 $\mathbf{A} \cdot \mathbf{X} = \sum_i \sum_j a_{ij} x_{ij}$ označuje skalární součin matic (viz §4.7).

SDP je velmi obecná třída konvexních úloh. LP, konvexní QCQP a SOCP jsou speciální případy SDP. Pro ilustraci ukážeme, že pokud matice \mathbf{C}, \mathbf{A}_i jsou diagonální, úloha (16.9) 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 když všechny její prvky jsou nezáporné (viz Cvičení 6.15). Tedy úloha (16.9) se redukuje na

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq \mathbf{0}, \mathbf{a}_i^T \mathbf{x} = b_i \ (i = 1, \dots, m) \},$$

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 16.12. *Analytický střed mnohostěnu* $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} \geq \mathbf{b} \}$ je bod uvnitř mnohostěnu, který maximalizuje součin vzdáleností od nadrovin $\mathbf{a}_i^T \mathbf{x} = b_i$. Předpokládáme, že v každé nerovnici $\mathbf{a}_i^T \mathbf{x} \geq b_i$ je $\|\mathbf{a}_i\|_2 = 1$ (to jde vždy zařídit vydělením nerovnice číslem $\|\mathbf{a}_i\|_2$). Vzdálenost

¹Toto mechanické zařízení je známé jako *Varignon frame* a v minulosti se opravdu používalo na řešení úlohy. Úloha má bohatou historii, je známa také jako Fermat-Weberův problém.

²Namítnete, že nemůžeme mluvit o konvexitě úlohy (16.9), protože v této úloze optimalizujeme přes množinu *matic* a konvexitu jsme definovali pro množiny a funkce *vektorů*. Definice konvexity lze ovšem snadno zobecnit na množiny a funkce matic: matici $\mathbb{R}^{m \times n}$ můžeme buď přerovnat do vektoru \mathbb{R}^{mn} , nebo (lépe) můžeme konvexitu definovat místo na prostoru \mathbb{R}^n na obecném vektorovém prostoru (viz učebnice lineární algebry).

(se znaménkem) bodu \mathbf{x} od nadroviny $\mathbf{a}_i^T \mathbf{x} = b_i$ je tedy rovna $\mathbf{a}_i^T \mathbf{x} - b_i$ (viz Cvičení 4.17.c). Místo součinu vzdáleností maximalizujeme jeho logaritmus (což je rostoucí funkce), tedy funkci

$$f(\mathbf{x}) = \log \prod_{i=1}^m (\mathbf{a}_i^T \mathbf{x} - b_i) = \sum_{i=1}^m \log(\mathbf{a}_i^T \mathbf{x} - b_i). \quad (16.10)$$

Tato funkce je konkávní. Funkce je definována jen pro vnitřní body mnohostěnu (tj. body splňující $\mathbf{Ax} > \mathbf{b}$) a blíží se $+\infty$ když se bod \mathbf{x} blíží zevnitř k hranici. Tím je implicitně vynucena podmínka, že optimální bod má ležet uvnitř mnohostěnu. Lze ukázat, že pokud je mnohostěn neprázdný a omezený, úloha má maximum a toto maximum je jediné. \square

16.5 Konvexní relaxace nekonvexních úloh

Relaxace je technika, kterou lze někdy získat přibližná řešení obtížných úloh. Spočívá na očividné skutečnosti (promyslete!), že pro každou množinu $X \subseteq \mathbb{R}^n$ a funkci $f: X \rightarrow \mathbb{R}$ platí

$$Y \supseteq X \implies \min_{\mathbf{x} \in Y} f(\mathbf{x}) \leq \min_{\mathbf{x} \in X} f(\mathbf{x}). \quad (16.11)$$

Jak toho použít? Necht' úloha $\min_{\mathbf{x} \in X} f(\mathbf{x})$ je obtížná její obtížnost pramení ze složitosti množiny přípustných řešení X . Nahradíme množinu X 'jednodušší' množinou $Y \supseteq X$ a řešíme snadnější úlohu $\min_{\mathbf{x} \in Y} f(\mathbf{x})$. Když budeme mít štěstí, bude nerovnost v (16.11) platit s rovností, tedy obě optima budou stejná. Když ne, získáme alespoň dolní mez na optimální řešení.

Je-li množina X nekonvexní, můžeme ji takto nahradit konvexní množinou $Y \supseteq X$. Pokud funkce f je konvexní, získáme konvexní úlohu. Mluvíme o pak **konvexní relaxaci**. Příkladem konvexní relaxace je LP relaxace, kterou jsme už viděli: lineární program (11.21) je konvexní relaxace lineárního celočíselného programu (11.20).

16.6 Cvičení

16.1. 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:

- a) $f(x, y) = x + y$
- b) $f(x, y) = x$
- c) $f(x, y) = \min\{x, y\}$
- d) $f(x, y) = \max\{x, y\}$
- e) $f(x, y) = |x + y|$
- f) $f(x, y) = x^2 + 9y^2$

V kterých případech se jedná o konvexní optimalizační úlohu?

16.2. Dokažte, že množina optimálních řešení konvexní optimalizační úlohy je konvexní.

16.3. 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?

16.4. Najděte spojitou funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$, jejíž každé lokální minimum je zároveň globální a

- a) není konvexní
- b) (\star) žádná její subkontura není konvexní množina.

16.5. Uvažujme úlohu, známou jako *lineární lomené programování*:

$$\begin{aligned} \min \quad & (\mathbf{c}^T \mathbf{x} + d) / (\mathbf{e}^T \mathbf{x} + f) \\ \text{za podmínek} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ & \mathbf{e}^T \mathbf{x} + f > 0 \end{aligned}$$

- a) Je účelová funkce na množině přípustných řešení konvexní?
- b) Dokažte, že úloha je ekvivalentní lineárnímu programu (s proměnnými \mathbf{x}, z)

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{y} + dz \\ \text{za podmínek} \quad & \mathbf{A}\mathbf{y} \geq \mathbf{b}z \\ & \mathbf{e}^T \mathbf{y} + fz = 1 \\ & z \geq 0 \end{aligned}$$

16.6. Najděte explicitní řešení pro následující úlohy QCQP (\mathbf{A}, \mathbf{B} jsou pozitivně definitní):

- a) $\min\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{A} \mathbf{x} \leq 1 \}$
- b) $\min\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, (\mathbf{x} - \mathbf{b})^T \mathbf{A} (\mathbf{x} - \mathbf{b}) \leq 1 \}$
- c) $\min\{ \mathbf{x}^T \mathbf{B} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{A} \mathbf{x} \leq 1 \}$

16.7. Formulujte úlohu $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_4$ jako konvexní QCQP.

16.8. 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 podmínek $\mathbf{a}_i^T \mathbf{x} > b_i$. Je tato úloha (po možné jednoduché transformaci) konvexní? Vyjádřete úlohu jako SOCP pomocí dokázané ekvivalence.

16.9. Dokažte, že účelové funkce vystupující v následujících úlohách jsou nekonvexní:

- a) Minimalizace funkce $f(\mathbf{x}_1, \dots, \mathbf{x}_n) = \max_{i=1}^m \min_{j=1}^n \|\mathbf{a}_i - \mathbf{x}_j\|$. Jaký význam má tato formulace v porovnání s (16.3)?
- b) Příklad 9.10
- c) Cvičení 9.5

Nápověda a řešení

16.3. Vezměme např. funkce $g_1(x) = -e^{-(x+1)^2}$ a $g_2(x) = -e^{-(x-1)^2}$. Každá z těchto funkcí má jediné lokální minimum, ale funkce $g_1 + g_2$ má lokální minima dvě.

16.5.a) Ne.

16.5.b) Uvažujte substituci $\mathbf{y} = \mathbf{x}/(\mathbf{e}^T \mathbf{x} + f)$, $z = 1/(\mathbf{e}^T \mathbf{x} + f)$.

16.6.a) Viz Cvičení 11.6.

16.6.b) Substituuje $\mathbf{y} = \mathbf{x} - \mathbf{b}$.

16.6.c) Optimální hodnota je nula.

16.8. Místo maximalizace funkce $f(\mathbf{x})$ minimalizujme funkci $1/f(\mathbf{x})$, která je konvexní na množině přípustných hodnot. Úloha je ekvivalentní úloze

$$\begin{aligned} \min \quad & t_1 + \dots + t_m \\ \text{za podmínek} \quad & t_i(\mathbf{a}_i^T \mathbf{x} - b_i) \geq 1, \quad i = 1, \dots, m \\ & t_i \geq 0, \quad i = 1, \dots, m. \end{aligned}$$

Použitím dokázané ekvivalence převedeme na SOCP

$$\begin{aligned} \min \quad & t_1 + \dots + t_m \\ \text{za podmínek} \quad & \|(2, \mathbf{a}_i^T \mathbf{x} - b_i - t_i)\|_2 \geq \mathbf{a}_i^T \mathbf{x} - b_i + t_i, \quad i = 1, \dots, m \\ & \mathbf{a}_i^T \mathbf{x} \geq b_i, \quad i = 1, \dots, m \\ & t_i \geq 0, \quad i = 1, \dots, m. \end{aligned}$$

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 Minimaxní nerovnost

Pro libovolné množiny X a Y a libovolnou funkci $L: X \times Y \rightarrow \mathbb{R}$ platí **minimaxní nerovnost**

$$\min_{x \in X} \underbrace{\max_{y \in Y} L(x, y)}_{F(x)} \geq \max_{y \in Y} \underbrace{\min_{x \in X} L(x, y)}_{G(y)}. \quad (17.1)$$

Zde předpokládáme, že všechna minima a maxima existují¹. V nerovnosti (17.1) nastává rovnost, právě 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 15.8 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**. Rozdíl 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 $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$. □

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

$$\min_{\mathbf{x} \in \mathbb{R}^n} \max_{\mathbf{y} \geq \mathbf{0}} L(\mathbf{x}, \mathbf{y}) = \min\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b} \}. \quad (17.12)$$

Odvodíme 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 předpisu (14.1) na konstrukci duálního LP. Konkrétně, primární a duální úloha je dvojice úloh (14.2). \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 předpisu (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

$$\min\{\mathbf{x}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} = \mathbf{b}\}.$$

Máme $X = \mathbb{R}^n$ a

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{x} + \mathbf{y}^T (\mathbf{b} - \mathbf{Ax}) = \mathbf{x}^T \mathbf{x} - \mathbf{y}^T \mathbf{Ax} + \mathbf{y}^T \mathbf{b},$$

duální funkce je tedy

$$G(\mathbf{y}) = \min_{\mathbf{x} \in \mathbb{R}^n} L(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{x} \in \mathbb{R}^n} (\mathbf{x}^T \mathbf{x} - \mathbf{y}^T \mathbf{Ax} + \mathbf{y}^T \mathbf{b}).$$

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

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} = \mathbf{b}, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}.$$

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

$$G(\mathbf{y}) = \min_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{0} \leq \mathbf{x} \leq \mathbf{1}} [(\mathbf{c}^T - \mathbf{y}^T \mathbf{A})\mathbf{x} + \mathbf{y}^T \mathbf{b}] = \mathbf{b}^T \mathbf{y} + \mathbf{1}^T \max\{\mathbf{0}, \mathbf{A}^T \mathbf{y} - \mathbf{c}\}$$

nebot' $\min_{\mathbf{0} \leq \mathbf{x} \leq \mathbf{1}} \mathbf{d}^T \mathbf{x} = \min\{\mathbf{0}, \mathbf{d}\} = -\max\{\mathbf{0}, -\mathbf{d}\}$ (kde min a max se rozumí po složkách). \square

Příklad 17.6. Řešme celočíselný lineární program

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \{0, 1\}^n, \mathbf{Ax} = \mathbf{b}\}.$$

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 vyjíměčně) i pro nekonvexní úlohu. Jednou takovou třídou úloh je libovolné (tedy ne nutně konvexní) QCQP s nejvýše *jedním* omezením.

Příklad 17.7. Úloha QCQP

$$\min\{\mathbf{x}^T \mathbf{Ax} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\} \quad (17.20)$$

není konvexní, pokud matice \mathbf{A} není pozitivně semidefinitní. Řešení ale najdeme snadno pomocí spektrálního rozkladu:

$$\min_{\mathbf{x}^T \mathbf{x} = 1} \mathbf{x}^T \mathbf{Ax} = \min_{\mathbf{x}^T \mathbf{x} = 1} \mathbf{x}^T \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \mathbf{x} = \min_{\mathbf{z}^T \mathbf{z} = 1} \mathbf{z}^T \mathbf{\Lambda} \mathbf{z} = \lambda_{\min}(\mathbf{A}).$$

Ukážeme navíc, že platí silná dualita. Máme

$$L(\mathbf{x}, y) = \mathbf{x}^T \mathbf{Ax} + y(1 - \mathbf{x}^T \mathbf{x}) = \mathbf{x}^T (\mathbf{A} - y\mathbf{I})\mathbf{x} + y.$$

Tedy

$$G(y) = \begin{cases} y & \text{když } \mathbf{A} - y\mathbf{I} \text{ je pozitivně semidefinitní,} \\ -\infty & \text{jinak.} \end{cases}$$

Ale matice $\mathbf{A} - y\mathbf{I}$ je pozitivně semidefinitní, právě když nejmenší vlastní číslo matice $\mathbf{A} - y\mathbf{I}$ je nezáporné, neboli (viz Cvičení 6.4) $\lambda_{\min}(\mathbf{A}) \geq y$. Duální úloha tedy je

$$\max\{y \in \mathbb{R} \mid \lambda_{\min}(\mathbf{A}) \geq y\}.$$

Ta má zřejmé optimální řešení $y = \lambda_{\min}(\mathbf{A})$. \square

Rejstřík

úloha

- dopravní, 123
- duální, 156
- ekvivalentní, 180
- konvexní, 178
- konvexní ve standardním tvaru, 180
- Lagrangeova duální, 188
- lineárního programování, 118
- na největší nezávislou množinu, 130
- na nejmenší vrcholové pokrytí, 129
- na optimální výrobní program, 121
- přiřazovací, 128
- primární, 156
- směšovací, 122
- spojité optimalizace, 10

řešení

- bázové, 144
- degenerované, 144
- optimální, 11
- přípustné, 11, 144

řez zobrazení, 85

afinní

- funkce, 30
- kombinace, 28, 136
- nezávislost, 29
- obal, 28, 136
- podprostor, 28, 136
- zobrazení, 30

báze

- lineárního podprostoru, 23
- mnohostěnu, 144
- sousední, 144
- standardní, 18

bod, 30

- extrémní, 138
- hraniční, 90
- regulární, 108
- sedlový, 97, 187

- stacionární, 96, 97
- vnitřní, 90

cyklení simplexového algoritmu, 149

derivace

- funkce jedné proměnné, 81
- parciální, 81
- směrová, 86
- totální, 82

determinant, 16

dimenze

- afinního podprostoru, 30
- lineárního podprostoru, 24
- mnohostěnu, 137
- stěny mnohostěnu, 139

doplnění na čtverec, 62

dualita

- silná, 158
- slabá, 157, 189
- v lineárním programování, 156

dyáda, 18

elipsoid, 63

epigraf, 171

extrém, 9

- globální, 91
- lokální, 91
- vázaný, 11
- vázaný rovnostmi, 107
- volný, 11

extrémní bod, 138

fasety mnohostěnu, 139

forma

- kvadratická, 59
- lineární, 25

Frobeniova

- norma, 42

funkce, 8

- účelová, 11

- afinní, 30
- klesající, 96
- konkávní, 169
- konvexní, 169
- kvadratická, 62
- Lagrangeova, 110
- po částech afinní, 120
- rostoucí, 96
- gradient, 87
- graf funkce, 79
- Gramm-Schmidtova ortonormalizace, 40
- hodnost
 - matice, 26
 - plná, 15
- hodnost matice, 15
- hrana mnohostěnu, 139
- hraniční bod množiny, 90
- hranice množiny, 90
- inverze matice, 16
- isometrie, 37
- jednotková sféra, 124
- klesající funkce, 96
- kombinace
 - afinní, 28, 136
 - konvexní, 135, 136
 - lineární, 23, 136
 - nezáporná, 136
- komplementarita, 157
- konvexní
 - funkce, 169
 - kombinace, 136
 - kužel, 136
 - množina, 135, 136
 - mnohostěn, 137
 - obal, 135, 136
 - optimalizační úloha, 178
- kuželosečka, 63
- kvadratická
 - forma, 59
 - funkce, 62
- kvadrika, 63
- Lagrangeův multiplikátor, 110
- Lagrangeova funkce, 110
- Laplacián grafu, 65
- lineární
 - program v rovnicovém tvaru, 119
 - forma, 25
 - kombinace, 23, 136
 - nerovnice, 118
 - nezávislost, 23
 - obal, 23, 136
 - optimalizace, 118
 - podprostor, 23, 136
 - program, 118
 - rovnice, 118
 - soustava, 46
 - soustava nehomogenní nedourčená, 46
 - soustava homogenní, 46
 - soustava homogenní přeурčená, 72
 - soustava nehomogenní, 46
 - soustava nehomogenní přeурčená, 46
 - zobrazení, 24
- LP relaxace, 127
- matice, 14
 - úzká, 14
 - čtvercová, 14
 - široká, 14
 - antisymetrická, 15
 - bloková, 18
 - diagonální, 14
 - diagonalizovatelná, 59
 - Hessova, 88
 - indefinitní, 60
 - inverzní, 16
 - Jacobiho, 82
 - jednotková, 14
 - nulová, 14
 - obdélníková, 14
 - ortogonální, 37
 - permutační, 38
 - pozitivně/negativně (semi)definitní, 60
 - regulární, 16
 - rotační, 37
 - singulární, 16
 - speciální ortogonální, 37
 - symetrická, 15
 - transponovaná, 15
 - trojúhelníková, 14
- metoda

- Gauss-Newtonova, 102
- Levenberg-Marquardtova, 104
- nejmenších čtverců, 46
- Newtonova, 99
- sestupná, 98
- simplexová, 144
- simplexová dvoufázová, 152
- simplexová základní, 149
- metrika eukleidovská, 34
- minimální
 - argument funkce na množině, 9
 - hodnota funkce na množině, 9
 - prvek množiny, 9
- minimum
 - funkce na množině, 9, 91
 - funkce na množině ostré, 9
 - globální, 91
 - lokální, 91
 - lokální ostré, 91
 - ostré, 91
- množina
 - číselná, 7
 - konvexní, 135, 136
- mnohostěn konvexní, 137
- monom, 57
- nadrovina, 30, 136
 - opěrná, 139
- nejmenší
 - čtverce, 46
 - norma, 51
 - prvek množiny, 9
- nerovnost
 - Cauchy-Schwarzova, 34
 - Jensenova, 169
 - minimaxní, 187
 - trojúhelníková, 34
- nezávislost
 - afinní, 29
 - lineární, 23
- norma
 - p -norma, 124
 - eukleidovská, 34
 - Frobeniova, 42
 - matice, 41
 - vektorová, 124
- obal
 - afinní, 28, 136
 - konvexní, 136
 - lineární, 23, 136
 - nezáporný, 136
- omezení, 11
- opěrná nadrovina, 139
- optimální
 - řešení, 11
 - argument, 11
 - hodnota, 11
- ortogonální
 - doplňěk, 35
 - matice, 37
 - podprostory, 35
 - projekce, 38
 - projektor, 39
 - vektory, 34
- ortonormální množina vektorů, 36
- přímka, 30, 140
- pivot, 146
- pivotové pravidlo, 149
- podprostor, 23
 - afinní, 28, 136
 - lineární, 23, 136
- podprostory
 - ortogonální, 35
- poloprostor, 136
- polynom, 57
 - charakteristický, 58
 - homogenní, 57
 - Taylorův, 88
- projekce ortogonální, 38
- projektor ortogonální, 39
- prostor
 - lineární, 23
 - nulový, 27
 - obrazů, 26
- pseudoinverze
 - matice s l.n. řádky, 52
 - matice s l.n. sloupci, 48
 - obecné matice, 53
- regrese, 49
 - lineární, 49
 - robustní, 126
- regulární
 - bod, 108

- matice, [16](#)
- regularizace, [51](#)
- rostoucí funkce, [96](#)
- rovina, [30](#)
- rovnice
 - lineární, [118](#)
 - normální, [47](#)
- rozklad matice
 - podle hodnoty, [32](#)
 - podle vlastních čísel, [59](#)
 - QR, [41](#)
 - spektrální, [59](#)
- simplexová
 - metoda, [144](#)
 - tabulka, [147](#)
- singulární
 - matice, [16](#)
- skalár, [14](#)
- skalární součin
 - matic, [41](#)
 - vektorů, [34](#)
- slacková proměnná, [120](#)
- směr
 - Gauss-Newtonovův, [102](#)
 - Newtonův, [101](#)
 - sestupný, [98](#)
- součin
 - maticový, [15](#)
 - skalární s. matic, [41](#)
 - skalární s. vektorů, [34](#)
 - skaláru a matice, [14](#)
 - vnější, [18](#)
- souřadnice, [24](#)
- spektrální
 - rozklad, [59](#)
- spektrum matice, [58](#)
- stínové ceny, [160](#)
- stacionární bod, [97](#)
- stacionární bod, [96](#)
- stopa matice, [17](#)
 - na podprostoru, [69](#)
- subkontura, [171](#)
- transpozice, [15](#)
- věta
 - Fermatova, [96](#)
 - Frobeniova, [46](#)
 - spektrální, [59](#)
- vektor, [9](#), [17](#)
 - řádkový, [17](#)
 - normalizovaný, [36](#)
 - sloupcový, [17](#)
- vektory
 - afinně nezávislé, [29](#)
 - kolmé, [34](#)
 - lineárně nezávislé, [23](#)
 - ortogonální, [34](#)
 - ortonormální, [36](#)
- vlastní
 - číslo, [57](#)
 - vektor, [57](#)
- vnitřek množiny, [90](#)
- vnitřní bod množiny, [90](#)
- vrchol, [139](#)
- vrstevnice funkce, [79](#)
- vychýlená hodnota, [126](#)
- zobrazení, [8](#)
 - řez z., [85](#)
 - afinní, [30](#)
 - diferencovatelné, [82](#)
 - lineární, [24](#)
 - spojité, [80](#)
 - spojitě diferencovatelné, [82](#)